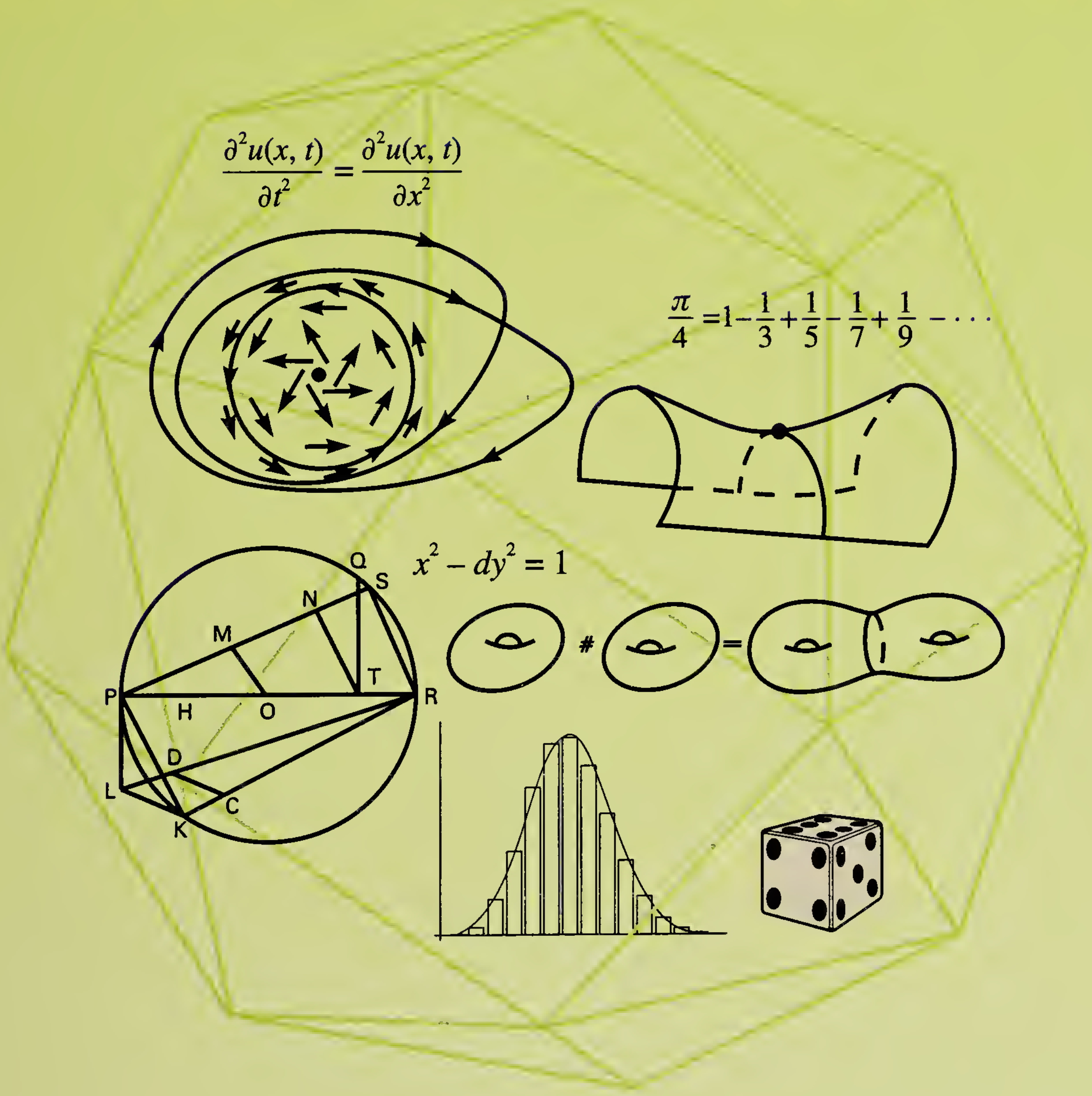
Universities Press MATHEMATICS

Echoes from

Resonance

Selected Articles from Resonance—A Journal of Science Education

Number Theory


$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{\partial^2 u(x,t)}{\partial x^2}$
$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots$
$x^2 - dy^2 = 1$
Q
S
N
M
T
P
H
O
R
D
L
C
K
#
=

Editors

Shailesh Shirali ✦ C S Yogananda

INDIAN ACADEMY OF SCIENCES

Number Theory

Number Theory

Editors

Shailesh Shirali + C S Yogananda

INDIAN ACADEMY OF SCIENCES

Universities Press

Universities Press (India) Private Limited

Registered Office
3-5-819 Hyderguda, Hyderabad 500 029 (A.P.), India

Distributed by
Orient Longman Private Limited

Registered Office
3-6-752 Himayatnagar, Hyderabad 500 029 (A.P.), India

Other Offices
Bangalore / Bhopal / Bhubaneshwar / Chennai
Ernakulam / Guwahati / Hyderabad / Jaipur / Kolkata
Lucknow / Mumbai / New Delhi / Patna

ISBN 81 7371 454 1

Typeset by
Devi Information Systems
Chennai

Printed at
Orion Printers
Hydreabad 500 004

Published by
Universities Press (India) Private Limited
3-5-819 Hyderguda, Hyderabad 500 029 (A.P.), India

Contents

Foreword

The Indian Academy of Sciences launched *Resonance* as a monthly journal devoted to science education in January 1996. *Resonance* is aimed largely at undergraduate students and teachers of science, though material of interest to somewhat younger students is also included. Each issue contains papers that span a wide area of science and mathematics, in various formats. Some are individual general articles, others consist of series with several parts. An effort is made to ensure good expository quality in all of them.

"Echoes from Resonance" is a series of books born out of *Resonance*, by putting together in a coherent manner a collection of articles (both series and single pieces) taken from *Resonance*, all written around a common theme. Typically, the individual articles would have appeared quite independently at different times. These collections should prove useful to a reader who is keen to learn about a specific subject, with accounts given by different authors from different perspectives, but all in an expository manner. We hope these volumes would be useful for students and teachers alike, and that they will complement the structure of individual issues of *Resonance* which cover different areas of science and mathematics in a balanced manner.

N. MUKUNDA

Preface

Number theory has been a subject of study by mathematicians from the most ancient of times. In the Plimpton 322 clay artefact, excavated from the ruins of ancient Babylon, one finds a systematic listing of a large number of Pythagorean triples—triples (a, b, c) of positive integers such that $a^2 + b^2 = c^2$; they appear to be listed in order of increasing c/a ratio. (One sees in the table the beginnings of trigonometry.) The Greeks had a deep interest in number theory. Euclid's great text, *The Elements*, generally considered as a book only on Geometry, actually contains a fair amount of number theory too; in particular it contains the proofs of two gems discovered by the Greeks–the irrationality of $\sqrt{2}$ and the infinitude of the primes. It also contains a description of the algorithm now known as the Euclidean algorithm, which computes the greatest common divisor of two given numbers. In ancient India too there was much interest in number theory, particularly in Diophantine equations; for instance, in the linear two-variable equation $ax + by = c$, where a, b, c are given integers, and in the equation later to be known as the Pell equation ($x^2 - Ny^2 = 1$, where N is a given positive integer). Building on the work of Brahmagupta (6th century) Bhaskara II (12th century) gave a completely general way of solving the latter equation.

In this book we offer the reader some articles in number theory that appeared in *Resonance* over the years 1996–2001. Traditionally, number theory begins with a study of congruences (Wilson's and Fermat's theorems, the Chinese remainder theorem, quadratic residues, primitive roots, ...), then proceeds to a study of prime numbers (the infinitude of various classes of primes, divergence of the sum $\sum 1/p$ taken over all primes, ...) and later to a study of Diophantine equations (solution of equations such as $x^2 + y^2 = z^2$, $ax + by = c$, where a, b, c are given integers, Pell's equation ...). The last two topics (prime numbers, Diophantine equations) are distinguished by the extraordinary diversity in terms of level of difficulty, of the problems they offer to the students. There is something in number theory for practically everyone!

The articles included within form a varied lot with the first half (articles 1 to 8) being of an elementary nature. We begin with a short essay on the axiomatic approach in modern mathematics: on how conventions sometimes need to be followed for the sake of preserving uniformity and maintaining mathematical harmony. The next two of the articles deal with elementary problems: "Find four positive integers such that the sum of any two is a square", and Bachet's problem ("100 kg with five stones"), solved using generating functions. There is a piece on mathematical induction, one of the very trustworthy and important techniques in the toolkit of any mathematician, particularly the number theorist and combinatorist. The following article describes Euler's proof of the infinitude of primes, which establishes rather more than Eulid's well-known proof of the same result. Then there is a short piece on Fermat's two-square theorem, elaborating on a "crisp and elegant proof" by Zagier of the theorem that a prime of the form 1 (mod 4) is a sum of two squares. The article also suggests an algorithm approach towards proving the theorem. In the following article, "Fermat's Two Squares Theorem Revisited", Bhaskar Bagchi proves the correctness of the algorithm. Following this is a report on recent work done on the factorization of Fermat numbers defined by $F_n = 2^{2^n} + 1$. (Fermat had conjectured, perhaps rather rashly, that the numbers F_n are all prime. Now it appears that for $n > 4$ they may never be prime!)

Articles 9–16 are of more substantive nature beginning with a two-part article (articles 9 and 10) on the class number problem ("Binary Quadratic Forms" and "Algebraic Number Theory"), a topic dealt with for the first time and in considerable detail by Gauss in his path-breaking book *Disquisitiones Arithemeticae*. The two articles which follow—"Roots are not contained in cyclotomic fields" and "Die Ganzen Zahlen hat Gott gemacht, alles andere ist Menschenwerk"—deal with cyclotomic polynomials and cyclotomic fields giving interesting applications of ideas introduced in the previous two-part article. A proof of a beautiful relation between prime representing quadratic equations and class numbers is the subject of the next article. We then have an article on congruent numbers, dealing with a problem dating from ancient times but which has intimate connections with a very modern topic - that of elliptic curves. This is followed by an expository article on one of the great mathematical achievements of the 20th century—the proof of "Fermat's Last Theorem" by Andrew Wiles. To top off the collection we have brief survey of some currently unsolved problems in number theory. (In passing, we remark briefly that references to Fermat appear surprisingly many times in this collection!)

We hope that the reader will enjoy this varied collection.

SHAILESH SHIRALI
C S YOGANANDA

1

On Provability versus Consistency in Elementary Mathematics

Shailesh A Shirali

A reader asks, "*Why is 1 not listed as a prime? After all, does it not satisfy the stated criteria for primality?*" This chapter is written in response to this question.

The layperson usually thinks that mathematics deals with absolute truths, and indeed this was how mathematics was viewed during earlier centuries. However, ever since the epochal discoveries of Bolyai, Lobachevsky and Riemann that there can be geometries (note the plural) other than the one presented in Euclid's text *The Elements*, this implicit notion had to be dropped. Even the notion that everything in mathematics is provably true or provably false had to be abandoned, after the astonishing results obtained by Gödel in 1930. Alongside this development, mathematics has seen a pioneering and extremely productive method: the axiomatic method, in which new areas of mathematics get created merely by defining suitable sets of axioms. As a result, the accent in mathematics has to some extent shifted to the study of axiomatic systems, and the essential question in such cases has become one of consistency and richness of the axiom system rather than its intrinsic truth or falsity. Much of modern algebra, starting with group theory, the theory of fields and rings and vector spaces and so on can be viewed in this light. Loosely speaking, one might say that in the modern mathematical paradigm, *true* is roughly equivalent to *consistent*, while *false* is equivalent to *self-contradictory*[1].

Here are some instances to illustrate the theme of consistency as opposed to absolute truth. In school arithmetic, one encounters the question, "Why is $-1 \times -1 = 1$?" Many 'proofs' are offered, but the plain fact is that the relation is a *convention*, not

[1] It is an interesting commentary on the psychology of modern mathematicians that, when pressed, most of them will readily say that there is no such thing as absolute truth in mathematics, and that a mathematical proposition is true or false only with reference to a particular axiomatic system. But amongst themselves most mathematicians 'know' that what they deal with does indeed refer to something 'concrete', 'real' and 'absolute'!

an absolute truth, and therefore there is no question of proving it[2]. One adopts it because of its implication for the law of distributivity of multiplication over addition (LDMA for short), according to which $a(b+c) = ab + ac$ for all a, b, c. The LDMA is too valuable an axiom to lose! Here is roughly how it happens. Starting with **N** the set of positive integers, with × and + defined on **N** in the usual manner, we enlarge the set by including 0 and imposing the following rules:

$$a + 0 = 0 + a = a, \quad a \times 0 = 0 \times a = 0.$$

Note that the two statements are consistent with one another because of the LDMA. For example, $2 \times 3 = 2 \times (3 + 0) = 2 \times 3 + 2 \times 0$, so we must have $2 \times 0 = 0$. Next, one constructs the negative numbers via the rule $a + (-a) = 0$. To do addition we call upon commutativity and associativity. For instance we have:

$$(-2) + (-3) + (2 + 3) = (-2) + 2 + (-3) + 3 = 0 + 0 = 0.$$

Therefore, $(-2) + (-3) + 5 = 0$ and $(-2) + (-3) = -5$.

Finally, multiplication is taken up, and here one invokes distributivity. We find that we are forced to adopt the convention that $-1 \times 1 = -1$ and $-1 \times -1 = 1$:

$$0 = 0 \times 1 = \{(1 + (-1)\} \times 1 = \{1 \times 1\} + \{(-1) \times 1\} = 1 + \{(-1) \times 1\},$$

therefore, $(-1) \times 1 = -1$; and,

$$0 = \{1 + (-1)\} \times (-1) = \{1 \times (-1)\} + \{(-1) \times (-1)\} = -1 + \{(-1) \times (-1)\},$$

therefore, $(-1) \times (-1) = +1$. *The point is that we need these relations if we are to preserve the LDMA, which we cannot afford to lose. The consistency of the system must be preserved at all cost*[3].

Here is another question, also asked at the school level: Why is $a^0 = 1$ for all $a > 0$? We proceed to resolve this in a similar vein. Let $x, y \in \mathbf{N}$; then $a^{x+y} = a^x \times a^y$ and $a^{x-y} = a^x/a^y$ when $x > y$. These follow from the very meaning of a^n when n is a positive integer. What do we do with a^0? If we wish to have a system of algebra that is consistent and easy to work with, then we need to adopt the convention that $a^0 = 1$. There is nothing absolute about this. Rather, we *choose* to give a^0 a meaning that makes it easy to deal with. In short, we make a^0 a well-behaved object. (Note that 0^0 cannot be given any consistent meaning, nor can $0/0$; that is, it is not possible to make these objects well-behaved.)

Finally we take up the question: "*Is 1 a prime?*" We recall the fundamental theorem of arithmetic (FTA): *Every integer $N > 1$ can be expressed in just one way as a product of primes, except possibly for the order of occurrence of the primes.* If 1 were included in the set of primes **P**, then the fact that $1^n = 1$ for all integers n would require us to rephrase the FTA by adding the clause "... except that 1 may occur to

[2] Here is a particularly preposterous proof which I encountered a few years back: the parabola $y = x^2$ is symmetric in the y-axis, therefore minus times minus equals plus!

[3] Sacrificing the LDMA would mean that we lose the ring structure of **Z**.

any arbitrary power." We would end up labelling 1 as a special prime, to be excluded from most of the interesting theorems about prime numbers. Indeed, what would in all likelihood happen is that theorems about primes would end up being phrased in terms of the set $\mathbf{P}' = \mathbf{P}\setminus\{1\}$. Thus giving 1 membership in **P** proves to be a nuisance, and it is simpler to keep it out right at the start.

The matter can be considered from another viewpoint. Let **Z** denote the set of integers, and consider the set of complex numbers of the form $a+bi$, where $a, b \in \mathbf{Z}$, and $i = \sqrt{-1}$. These are the *Gaussian integers* first studied in detail by Gauss, and the set of such numbers is denoted by $\mathbf{Z}(i)$. (Note that **Z** is a subset of $\mathbf{Z}(i)$.) Now in **Z**, the only elements that possess multiplicative inverses are ± 1 (that is, their reciprocals lie within the same set); these are the *units* of **Z**. In $\mathbf{Z}(i)$, the set of units turns out to be $\{\pm 1, \pm i\}$. (The reader is invited to verify that there are no other units in $\mathbf{Z}(i)$.) Arithmetic can be done in $\mathbf{Z}(i)$ just as it is in **Z**; for instance, we can factorize numbers:

$$9+7i = (2+3i)(3-i), \qquad 13 = (2+3i)(2-3i), \ldots.$$

Observe that 13, which is prime in **Z**, loses its primality status in $\mathbf{Z}(i)$.

We declare a number $z \in \mathbf{Z}(i)$ to be *prime* if z is not a unit and if in every factorization $z = uv$, with $u, v \in \mathbf{Z}(i)$, either u or v is a unit[4]. The reader is invited to verify that 3, 7 and $2+3i$ are Gaussian primes, whereas 2, 5 and 13 are composite (because $2 = (1+i)(1-i), 5 = (1+2i)(1-2i)$, etc.). We now have the result: *every number in* $\mathbf{Z}(i)$, *not 0 or a unit, can be written as a product of Gaussian primes; moreover, there is essentially only one way of doing this*[5]. That is, we have an analogue of the FTA for the Gaussian integers, provided that the units are not considered as primes.

Other such number systems can be constructed. Indeed, once one grasps the idea, such systems seem to be available in abundance and can be spotted in many settings. For instance, consider the set $\mathbf{Z}(\sqrt{2})$ whose elements are numbers of the form $a + b\sqrt{2}$ where $a, b \in \mathbf{Z}$. This system presents itself quite naturally when one tries to solve the equation $x^2 - 2y^2 = \pm 1$ in integers. A striking fact about $\mathbf{Z}(\sqrt{2})$ is that it has infinitely many units. (The reader is invited to show this. Hint: Show that $\sqrt{2} - 1$ and its integral powers are units; (harder) show that these are the *only* units.) What are the primes of $\mathbf{Z}(\sqrt{2})$? It turns out that $\sqrt{2}$ is prime, as are the numbers 3, 5 and 11, but not 7, because $7 = (3 - \sqrt{2}) \times (3 + \sqrt{2})$, nor 17, because $17 = (5 - 2\sqrt{2}) \times (5 + 2\sqrt{2})$. It is an interesting exercise to classify the primes of $\mathbf{Z}(i)$ and $\mathbf{Z}(\sqrt{2})$. Is there an analogue of the FTA for $\mathbf{Z}(\sqrt{2})$? The answer is "yes", though it is hard work to prove it. However there are numerous number systems that closely resemble $\mathbf{Z}(i)$ and $\mathbf{Z}(\sqrt{2})$ but which do not have the FTA property. An example is $\mathbf{Z}(\sqrt{10})$: it can be shown that $2, 3, 4 - \sqrt{10}$ and $4 + \sqrt{10}$ are primes in $\mathbf{Z}(\sqrt{10})$, yet

$$6 = 2 \times 3 = (4 - \sqrt{10}) \times (4 + \sqrt{10}),$$

[4] Since this article deals with terminology, it should be pointed out that what we refer to as 'prime' here is usually called 'irreducible' in the standard texts. In the standard definition, p is prime if we have the implication $p|ab \Longrightarrow p|a$ or $p|b$. In the class of rings known as UFD's the two notions coincide. Examples of UFD's are **Z**, $\mathbf{Z}(i)$ and $\mathbf{Z}(\sqrt{2})$. However $\mathbf{Z}(\sqrt{10})$ is not a UFD.

[5] Units may enter the picture, hence the use of the words 'essentially only one way'.

providing a counter example to the FTA. As the reader will have noted by now, the word *prime* no longer carries a fixed meaning; it acquires meaning only with reference to a particular context[6]. The interested reader can consult the well-known text by G H Hardy and E M Wright (*An Introduction to the Theory of Numbers*, Chapters XIV and XV) for further details.

Here is another example of axiomatic generalization. A rational number can be thought of as a root of the equation $mx + n = 0$, with $m, n \in \mathbf{Z}$, $m \neq 0$; here $m = 1$ gives us the integers — we call these the *rational integers*. Generalizing, we define an *algebraic number* as a root of the polynomial equation $ax^n + bx^{n-1} + cx^{n-2} + \cdots = 0$ with $a, b, c, \ldots \in \mathbf{Z}$, $a \neq 0$ and $n \in \mathbf{N}$; if $a = 1$ then we have an *algebraic integer*. It is a non-trivial fact that the set $\mathbf{A}$ of *algebraic integers* is closed under addition and multiplication but not under division. Thus $\mathbf{A}$ behaves very much like $\mathbf{Z}$, and we have at hand a genuine generalization of the notion of integer.

These examples may serve to highlight the extraordinary freedom that the axiomatic approach brings into mathematics. Some critics complain, however, that in exercising this freedom, mathematicians tend to "go too far"; but that is another matter altogether and we shall not address it here.

TAIL-PIECE. Mr T B Nagarajan of Thanjavur has sent me the following problem: *Find four distinct positive integers such that the sum of any two of them is a square.* He writes that the problem is not too hard if the restriction on positivity is removed, or if one is content with solutions having very large integers. In support of this statement, he lists the following solutions:

$$\{55967, 78722, 27554, 10082\}, \qquad \{15710, 86690, 157346, 27554\}.$$

Readers are invited to take a crack at the problem. (To find a *triple* with the stated property is much easier; an example is $\{6, 19, 30\}$. Readers may enjoy trying to list further such triples before going on to the more challenging four-number problem.)

SHAILESH A SHIRALI
Rishi Valley School
Rishi Valley 517 352
Andhra Pradesh

[6] Historically, many of these developments were a result of efforts to prove Fermat's last theorem. See *Resonance*, Volume 1, No. 1 for more details.

2

To Find Four Distinct Positive Integers such that the Sum of Any Two of them is a Square

S H Aravind

The problem is to find four distinct positive integers such that any two of them add up to a square. Let a, b, c, d with $a < b < c < d$ be four positive integers such that the sum of any two of them is a square. Observing that

$$a + b + c + d = (a + b) + (c + d),$$
$$a + b + c + d = (a + c) + (b + d),$$
$$a + b + c + d = (a + d) + (b + c),$$

we need to find a number which can be written as a sum of two non-zero squares in three different ways. We proceed to find such a number.

To begin with, note that if two numbers n and n' can each be expressed as a sum of two squares, then nn' can also be so expressed in two ways. Indeed, if

$$n = k^2 + l^2, \qquad n' = k'^2 + l'^2,$$

then

$$\begin{aligned} nn' &= (kk' + ll')^2 + (kl' - lk')^2 \\ &= (kl' + lk')^2 + (kk' - ll')^2. \end{aligned}$$

Start with 25 and 13 both of which are sums of two squares, $25 = 3^2 + 4^2$, $13 = 2^2 + 3^2$. By the identity, $25 \times 13 = 325$ can be expressed as a sum of two squares: $325 = 10^2 + 15^2 = 6^2 + 17^2 = 1^2 + 18^2$. (Note that $10^2 + 15^2 = 5^2(2^2 + 3^2)$.)

We show that 13×25^2 has three representations as a sum of two squares and gives us a solution. Consider the following representations (among others):

$$8125 = 30^2 + 85^2 = 50^2 + 75^2 = 58^2 + 69^2.$$

Thus, we take $a + b + c + d = 8125$ and look for solutions a, b, c, d in positive integers. Then $a + b$, $a + c$, $a + d$, $b + c$, $b + d$, $c + d$ are precisely 30^2, 85^2, 50^2, 75^2, 58^2, 69^2 in some order. We have

$$a + b < a + c < a + d < b + d < c + d,$$

and
$$a + b < a + c < b + c < b + d < c + d.$$

We arbitrarily take $b + c$ to be less than $a + d$. So we have $a + b = 30^2$, the least of the squares, $a + c = 50^2$, the next smallest, and $c + b = 58^2$. We get $c - b = 1600$ and solving for c, b we have $c = 2482, b = 882$. From this we get $a = 18, d = 4743$. Thus $\{18, 882, 2482, 4742\}$ is a set of four positive integers with the required property.

The same method can give large integer solutions too. For example, the following solutions can be obtained by choosing suitable squares:

$$\{4190, 10210, 39074, 83426\}, \qquad \{7070, 29794, 71330, 172706\}.$$

S H Aravind
12, First Main Road
Ponmeni Jayanagar
Madurai 625 010

3

Bachet's Problem

B Bagchi

A grocery shopkeeper keeps five stones of different weights. He is able to use a common balance and weigh out quantities ranging from 1 to 100 kg, in steps of 1 kg. What are the weights of these five stones?

The above is the problem "100 kg with five stones" posed by R Yusufzai in the "Think it Over" column of the July 1996 issue of *Resonance*. A much better problem will result if the figure 100 is replaced by 121. This is because the question "*What are the weights of these five stones*?" seems to suggest that there are uniquely determined weights to be found! However, as may easily be verified, the weights in kg of the stones might be 1, 3, 9, 27 and m, where m is any integer in the range $60 \leq m \leq 81$. In fact, there are many other solutions to the problem as posed. If, however, it was given that the grocer can weigh any object of weight between 1 kg and 121 kg (in steps of 1 kg) using his five stones, then the weights (in kg) of the stones must have been 1, 3, 9, 27 and 81. This is the case $k = 5$ of the result stated and proved below.

The problem is a well-known variation of an old problem due to Bachet (see Suggested Reading). In the original *binary* version, the grocer cannot subtract, so he must put the stones in one pan and the object in the other. Mr Yusufzai's problem is an instance of the *ternary* version where this restriction is removed. The general problem (in its *ternary* version) may be stated as follows:

> Given a positive integer k, find the largest integer N_k such that any object whose weight is an integer between 1 and N_k (ends included) can be weighed using k stones of suitable integral weights. In this notation, the problem is to show that $N_5 \geq 100$.

In fact, we have:

THEOREM. $N_k = \frac{3^k - 1}{2}$. If k stones are such that all integral weights between 1 and N_k can be measured using them, then the weights of these stones must be 3^j, $0 \leq j \leq k - 1$.

This is, essentially, Theorem 141 in the book by Hardy and Wright (see Suggested Reading).

In order to prove this, we must convert it into a precise mathematical statement. To this end, let $a_0, \ldots, a_{k-1}$ be the (positive integral) weights of k stones. In order to weigh an object of integral weight m, the grocer places the object together with some of the stones on the right pan (say) and puts some other stones on the left pan. For $0 \leq j \leq k-1$, put $\varepsilon_j = 1$ if the stone of weight a_j is placed on the left pan, $\varepsilon_j = -1$ if it is on the right pan, $\varepsilon_j = 0$ if it is not used. Since the two pans must balance, we get

$$m = \sum_{j=0}^{k-1} \varepsilon_j a_j. \quad \text{where} \quad \varepsilon_j \in \{0, 1, -1\} \quad \text{for} \quad 0 \leq j \leq k-1. \tag{1}$$

This leads us to

DEFINITION. If $A = \{a_0, \ldots, a_{k-1}\}$ is a finite set of positive integers, then the *capacity* $C(A)$ of A is the largest integer M such that for every integer m in the range $1 \leq m \leq M$, equation (1) has a solution.

Informally, the capacity $C(A)$ is the largest M such that all weights between 1 and M can be measured using k stones whose weights are in A. In terms of this definition, the above theorem may be restated as follows:

THEOREM. If A is of size k, then $C(A) \leq \frac{3^k - 1}{2}$. Equality holds here if and only if $A = \{3^j : 0 \leq j \leq k-1\}$.

To prove the theorem, note that if m can be written as in (1), then so can $-m$ (just change the signs of all ε_js); also, trivially, $m = 0$ can be written thus (take $\varepsilon_j = 0$ for all j). Therefore, if $C(A) = M$, then all the $2M + 1$ integers m in the range $-M \leq m \leq M$ can be expressed as in (1). But there are three choices for ε_j for each j, hence only 3^k choices for the right hand side of (1). Hence, $2M + 1 \leq 3^k$, or $C(A) \leq (3^k - 1)/2$. Now, if we take $A = \{3^j : 0 \leq j \leq k-1\}$, then for $1 \leq m \leq (3^k - 1)/2$ write $[(3^k - 1)/2] - m$ in base 3:

$$[(3^k - 1)/2] - m = \sum_{j=0}^{k-1} \delta_j \, 3^j,$$

where $\delta_j \in \{0, 1, 2\}$. Put $\varepsilon_j = 1 - \delta_j$. Then (1) holds. Thus $C(A) \geq (3^k - 1)/2$ for this set. Together with the previous inequality, we get $C(A) = (3^k - 1)/2$.

Only the uniqueness part of the theorem remains to be proved. In fact, this is the only non-trivial and interesting part. To prove this, let $A = \{a_0, \ldots, a_{k-1}\}$ have capacity N_k. Since, now, equality holds in the inequality $C(A) \leq (3^k - 1)/2$ which appears in the statement of the theorem, the proof of the inequality shows that every integer m in the range $[-(3^k - 1)/2] \leq m \leq [(3^k - 1)/2]$ has a *unique* representation

(1); conversely, any m of the form (1) belongs to this range. Therefore, letting X be an indeterminate, we get

$$\prod_{j=0}^{k-1}(X^{-a_j}+1+X^{a_j}) = \sum_{|m|\le \frac{3^k-1}{2}} X^m \qquad (2)$$

as may be verified by multiplying out the left-hand. Since, in particular, the largest integer (viz., $\sum_{j=0}^{k-1} a_j$) of the form (1) must be the largest integer in the range $[-\frac{3^k-1}{2}, \frac{3^k-1}{2}]$, we also have

$$\sum_{j=0}^{k-1} a_j = \frac{3^k-1}{2}. \qquad (3)$$

Using (3) and a little algebra, (2) simplifies to

$$\prod_{j=0}^{k-1} \frac{X^{3a_j}-1}{X^{a_j}-1} = \frac{X^{3^k}-1}{X-1}. \qquad (4)$$

Now fix j; $0 \le j \le k-1$. Let w be a primitive $3a_j$th root of unity. That is, w is a complex number such that $w^l = 1$ if and only if l is an integral multiple of $3a_j$. (For instance, we may take $w = \exp(2\pi\sqrt{-1}/3a_j)$.) Then w is a zero of the left-hand side, and hence also of the right-hand of (4). Thus, $w^{3^k} = 1$. So $3a_j$ divides 3^k. That is, $a_j \in \{3^i, 0 \le i \le k-1\}$. Since this holds for all j, we have $A \subseteq \{3^i : 0 \le i \le k-1\}$. Since both sets have size k, we must have $A = \{3^i : 0 \le i \le k-1\}$. This proves the uniqueness of the set of given size and maximum capacity.

The reader may like to look up the proof in the book by Hardy and Wright, which is very different from the proof given here. It is a clever use of mathematical induction.

TAIL-PIECE. Bachet is better remembered by mathematicians for another reason. It was on Bachet's edition of Diophantus' Arithmetic that Fermat scribbled his famous marginal notes. Bachet was also the first man to state, (without proof) what is now known as Lagrange's four square theorem: every natural number is the sum of at most four perfect squares.

Suggested Reading

[1] F Schuh. *The Master Book of Mathematical Recreations.* Dover. New York. pp 115–118, 1968.

[2] G H Hardy and E M Wright. *An Introduction to the Theory of Numbers.* Oxford Univ. Press. London. pp 115–117, 1971.

B BAGCHI
Statistics and Mathematics Unit
Indian Statistical Institute
Bangalore 560 059

4

Mathematical Induction

An Impresario of the Infinite

B Sury

In the natural sciences, if a certain phenomenon is observed to occur a number of times, often a general law is formulated. This process is called *empirical induction.* In general, any reasoning that draws a general conclusion based on verification of particular cases is known as induction. But, in mathematics, a statement involving a natural number n might turn out to be erroneous even if it happens to be true for the first ten, or thousand, or even million natural numbers. For instance, the numbers $2^{2^0} + 1 = 3$, $2^{2^1} + 1 = 5$, $2^{2^2} + 1 = 17$, $2^{2^3} + 1 = 257$, $2^{2^4} + 1 = 65537$ are all prime numbers and the 17th century mathematician Pierre de Fermat suggested that $2^{2^n} + 1$ must be prime for every positive integer n. However, a century later, another great mathematician Leonhard Euler showed that $2^{2^5} + 1 = 641 \times 6700417$. An even more convincing example is the following. If we evaluate the expression $991n^2+1$ for small values of n, the resulting number is not the square of a whole number. But, for $n = 12055735790331359447442538767$, the value is a perfect square. Indeed, this is the smallest value of n for which it is a square! This tells us that, in mathematics, a lot of care is needed to establish an induction procedure which proves a mathematical theorem for each of an infinite sequence of cases, without exception. The method of mathematical induction is such a procedure. Let us start with a simple example.

Suppose we want to prove the statement that $2^n > n$ for every natural number n. Clearly, this inequality holds for $n = 1$. Now, to prove the inequality for all natural numbers, we consider an *arbitrary* natural number $k \geq 1$. We *assume* that the inequality $2^k > k$ holds. Then, for the *next* natural number $k+1$, $2^{k+1} = 2 \times 2^k > 2k$ by our *assumption* that $2^k > k$. Now, $2k = k + k \geq k + 1$, so that the inequality $2^{k+1} > k + 1$ follows. Thus, we have proved that *if* the inequality is true for any particular k, *then* it is also true for $k + 1$.

The crux of the above argument rests on the points:

(0) Given an infinite sequence of statements $P_r, P_{r+1}, \ldots,$ we would like to prove that there is a 'next' to any statement, and each particular statement can

be reached in a finite number of steps starting from the 'first' statement P_r.

(1) There is a general method of proving that for any $n \geq r$, *if* P_n is true, *then* P_{n+1} is true; and

(2) The first statement P_r is *known* to be true.

It is believed that these rules of logic are as fundamental to mathematics as the classical rules of Aristotelian logic.

It is necessary to verify both steps (1) and (2) to avoid landing in absurdities. For example, if step (2) that 'starts induction' is not verified, one can 'prove' that all natural numbers are equal as follows. For, simply denote by P_n the statement '$n = n + 1$'. Then, obviously, if P_n is assumed to be true, then $n = n + 1$ and so $n + 1 = n + 2$, which means that P_{n+1} is also true.

Everybody has seen instances of mathematical induction being applied. The summing of arithmetic and geometric progressions are usually done by this method.

An important point is in order here. *Mathematical induction can be used to prove a statement that is given to begin with. As for coming up with that statement itself (as a guess, say), it is altogether a different matter. Therein lies the creative element which cannot be pinned down by any general rules.*

As we observed earlier, mathematical induction is a procedure that involves such extremely 'believable' logic that we accept it as valid reasoning. But, interestingly, we can actually prove its validity if we assume another believable principle which is that *any non-empty set of positive integers has a least number*. That this principle gives a proof of the validity of mathematical induction is left as an exercise to the reader.

We now proceed to give various instances where the method of mathematical induction appears and proves fruitful.

The following is a slight variant of the form in which induction is used:

To prove an infinite sequence $P_k, P_{k+1}, \ldots,$ of assertions, one verifies the two steps:

(i) P_k is true.

(ii) For any $n \geq k$, if we assume that all the assertions $P_k, P_{k+1}, \ldots, P_n$ hold good, then P_{n+1} also holds true.

Induction in Geometry

As an example, let us show that the sum of the interior angles of a (not necessarily convex) polygon of n sides is $180(n - 2)$ degrees for all $n \geq 3$. Call this statement P_n. P_3 is true as the sum for a triangle is 180 degrees. P_4 is also true since any quadrilateral can be split into two triangles.

Now, let $n > 4$ and we assume that P_k is true for $k = 3, 4, \ldots, n - 1$. Let $A_1, A_2, \ldots, A_n$ be the vertices of a polygon with n sides. We first notice that there is always a diagonal (i.e., a segment A_iA_j that is not a side) that splits the polygon into two with smaller numbers of sides. To see this, consider three neighbouring vertices A, B, C. Consider all the rays emanating from B and filling the interior angle ABC. We terminate any ray when it first meets a side or a vertex of the polygon. Either all

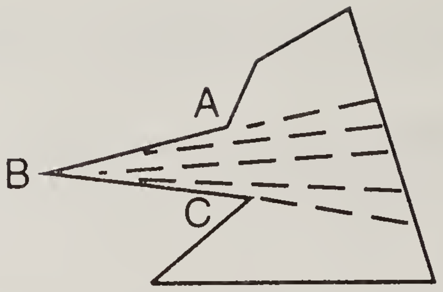


Figure 4.1

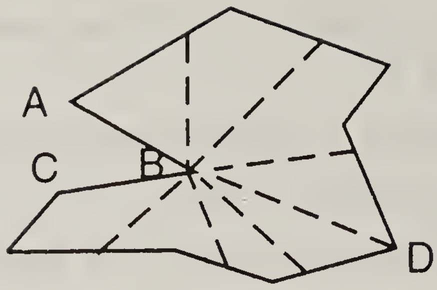


Figure 4.2

these rays intersect only one side (Figure 4.1) or they intersect more than one side (Figure 4.2). In the first case, AC is a diagonal that splits the original polygon into a triangle and a polygon with $n-1$ sides. In the second case at least one ray terminates on a vertex other than A or C. Call such a vertex, D. Then, BD is a diagonal splitting the polygon into two, with smaller numbers of sides.

Therefore, in general, let A_1A_k denote a diagonal which splits the polygon $A_1A_2 \ldots A_n$ into the polygons $A_1A_2 \ldots A_k$ and $A_kA_{k+1} \ldots A_nA_1$ of k and $n-k+2$ sides respectively. By induction hypothesis, P_k and P_{n-k+2} are true, i.e., the sum of the interior angles of the original polygon $A_1A_2 \ldots A_n$ is $180(k-2)+180(n-k) = 180(n-2)$ degrees. So, P_n is true, which proves by induction that P_r is true for every $r \geq 3$.

After this standard example, we look at an example where it may not be quite apparent that induction can be used.

The Marriage Problem

The classical 'marriage problem' can be stated as follows. Suppose that each of a set of girls is acquainted with a subset from a given set of boys. Is it possible for each girl to marry one of her acquaintances? Obviously, a necessary condition is that every set of m girls be collectively acquainted with at least m boys. *That this suffices is the assertion.* Here is a proof by induction.

Let n denote the number of girls. If $n = 1$, the assertion is trivial. If $n > 1$ and if it is true that every set of m girls, $1 \leq m < n$, has at least $m+1$ acquaintances, then an arbitrary girl is allowed her choice and the rest are referred to the induction hypothesis. If, on the other hand, some group of m girls, $1 \leq m < n$, has precisely m collective acquaintances, then this set of m girls is married off by induction and, it is indeed true that the rest of the $n-m$ girls satisfy the necessary condition with respect to the remaining boys. If this were not so, then some set of s spinsters with $1 \leq s \leq n-m$ would know fewer than s bachelors, and this set of s spinsters together with the m just-married girls would have known fewer than $s+m$ boys.

The reader is invited to apply induction to solve the following two problems.

EXERCISE. Consecutive Number Problem: Agatha and Beula are 'given' two consecutive natural numbers n and $n+1$. Both know that the numbers are consecutive but neither knows whose number is bigger. After every minute a beep is heard and each is asked to simultaneously say out aloud whether she knows the other's number. Prove by induction on the smaller number n that the person who has the number n guesses correctly after precisely the nth beep.

EXERCISE. Macaulay Expansion: Given a natural number $d \geq 2$, let us write down the d-tuples of positive integers in a strictly decreasing order. Order the tuples lexicographically. Prove that the number of tuples appearing prior to a particular tuple $(k_d, k_{d-1}, \ldots, k_1)$ is precisely $\binom{k_d}{d} + \binom{k_{d-1}}{d-1} + \cdots + \binom{k_1}{1}$.

This proves that any n has a unique expansion

$$n = \binom{k_d}{d} + \binom{k_{d-1}}{d-1} + \cdots + \binom{k_1}{1},$$

where $k_d > k_{d-1} > \cdots > k_1$. Here $\binom{n}{r}$ denotes the binomial coefficient which is 0 when $n < r$.

Induction Incognito—Use of a 'Dummy' Element

Look at the following statement:

If $a_1 < a_2 < \cdots < a_{n+1}$ are integers from the set $\{1, 2, \ldots, 2n\}$, then a_i divides a_j for some $i < j$.

This can be proved by the 'pigeon-hole principle' as follows. Write $a_i = 2^{k_i} l_i$ with l_i odd. Then, $l_1, \ldots, l_{n+1}$ being $n+1$ odd numbers between 1 and $2n$ cannot be different. If $l_i = l_j = l$ with $i < j$, then, clearly $a_i = 2^{k_i} l$ divides $a_j = 2^{k_j} l$.

In terms of economy and elegance, this is unbeatable. However, we find to our surprise that even induction works and, in fact, proves the following more general statement:

Let $r \geq 1$, and let $A \subset \{1, 2, \ldots, 2^r n\}$ be a subset of cardinality $(2^r - 1)n + 1$. Then, there exists a chain of $r + 1$ elements of A with each dividing the next.

Let us prove the original statement (corresponding to $r = 1$). Note that it is clearly true for $n = 1$. Assume it is true for n. Consider now $n + 2$ numbers $a_1 < \cdots < a_{n+2}$ among 1 to $2n + 2$. If $a_{n+1} \leq 2n$, we are done by the induction hypothesis. In the contrary case, we must have $a_{n+1} = 2n + 1$ and $a_{n+2} = 2n + 2$. If one of the a_i's is $n + 1$, we are done as it divides a_{n+2}. So, suppose $a_i \neq n + 1$ for any i. We may also assume that none of the n numbers $a_1, \ldots, a_n$ divides another or else we have nothing to prove. Now, we *put in this new number* $n + 1$ (as a 'dummy element') to get $n + 1$ numbers between 1 and $2n$. By induction hypothesis, one of these $n + 1$ numbers divides another. Since this has happened only after the advent of the new number $n + 1$, it must be that either: (i) some a_i $(i \leq n)$ divides $n + 1$, or (ii) $n + 1$ divides some a_i $(i \leq n)$. But, clearly (ii) cannot happen as $n + 1 \neq a_n \leq 2n$. Thus, some a_i $(i \leq n)$ divides $n + 1$ and, therefore, divides $2n + 2 = a_{n+2}$ also. Thus, we used $n + 1$ as a 'dummy element' in this proof.

The reader is urged to complete the proof of the general statement along the same lines.

Now, we come to a final example where induction appears in a different guise.

Backward Induction

If a statement is easily proved for a *particular* infinite subsequence of positive integers, it might be worthwhile to try and see whether 'backward induction' works. By this, we mean the following. Suppose we want to prove statements P_n for all positive integers n. Suppose, further, that it is easy to check the veracity of P_n for all n in an infinite sequence of natural numbers. Then, if we check that for any $m \geq 2$ the truth of P_m implies the truth of P_{m-1}, the statements P_n follow for all positive integers n.

An instance is the familiar arithmetic mean–geometric mean inequality

$$P_n \quad : \quad \left(\sum_{i \leq n} a_i\right)^n \geq n^n \prod_{i \leq n} a_i$$

for arbitrary non-negative real numbers a_i, where equality holds if, and only if all the numbers are equal.

On the one hand, we prove this for $n = 2^k$ by induction on k. Let $k = 1$. Then, $(a_1 + a_2)^2 \geq 4a_1a_2$ with equality exactly when $a_1 = a_2$, since the difference $(a_1 + a_2)^2 - 4a_1a_2 = (a_1 - a_2)^2$. Assume that P_n is true for $n = 2^r, r \leq k$. Let $a_i, i \leq 2^{k+1}$, be non-negative real numbers. Then, $\sum_{i \leq 2^{k+1}} a_i = \sum_{i \leq 2^k} b_i$ where $b_i = a_{2i-1} + a_{2i}$. Therefore,

$$\left(\sum_{i \leq 2^{k+1}} a_i\right)^{2^{k+1}} = \left(\sum_{i \leq 2^k} b_i\right)^{2^{k+1}} = \left(\left(\sum_{i \leq 2^k} b_i\right)^{2^k}\right)^2$$

$$\geq \left(\left(2^k\right)^{2^k} \prod_{i \leq 2^k} b_i\right)^2 = 2^{k2^{k+1}} \prod_{i \leq 2^k} b_i^2 \geq 2^{k2^{k+1}} \prod_{i \leq 2^k} (4a_{2i-1}a_{2i})$$

$$= 2^{k2^{k+1}} 4^{2^k} \prod_{i \leq 2^{k+1}} a_i = 2^{(k+1)2^{k+1}} \prod_{i \leq 2^{k+1}} a_i,$$

which proves that $P_{2^{k+1}}$ is true. Hence, by induction, P_{2^r} is valid for all $r \geq 1$. Moreover, note that the above proof also shows that the equality $(\sum_{i \leq 2^{k+1}} a_i)^{2^{k+1}} = 2^{(k+1)2^{k+1}} \prod_{i \leq 2^{k+1}} a_i$ implies that all inequalities occurring on the way are equalities, which again proves by induction that equality can hold in P_{2^r} if, and only if all the a_i's are equal.

On the other hand, for any m, the validity of P_m implies the validity of P_{m-1} as follows:

Let $a_1, \ldots, a_{m-1}$ be given. Consider $a_m = \frac{1}{m-1} \sum_{i \leq m-1} a_i$. Then,

$$\left(\sum_{i \leq m-1} a_i\right)^m = \left(\frac{m}{m-1}\right)^m \left(\sum_{i \leq m-1} a_i\right)^m$$

$$\geq \left(\frac{m}{m-1}\right)^m (m-1)^{m-1} \left(\prod_{i \leq m-1} a_i\right)\left(\sum_{i \leq m-1} a_i\right) = m^m \prod_{i \leq m} a_i .$$

Once again, by induction, equality implies that all the numbers are equal.

To end our discussion, the reader is invited to apply induction on the positive integer p below to prove the following result which solves an interesting two-player game called Euclid.

Let (p, q) be a pair of positive integers satisfying $p > q$. Each player subtracts a multiple of the smaller number from the bigger one without making the result negative. The winner is the one first hitting the highest common factor of p and q. Then, there is a winning strategy for the first player if, and only if $q < \frac{1}{2}(\sqrt{5} - 1)p$.

Suggested Reading

[1] R Courant and H Robbins. *What is Mathematics?* Oxford University Press, 1941.

[2] L I Golovina and I M Yaglom. *Induction in Geometry.* Little Mathematics Library. Mir Publishers. Moscow, 1979.

B SURY
Statistics and Mathematics Unit
Indian Statistical Institute
Bangalore 560 059

5

On the Infinitude of the Prime Numbers

Euler's Proof

Shailesh A Shirali

Euclid's elegant proof that there are infinitely many prime numbers is well known. Euler proved the same result, in fact a stronger one, by *analytical* methods. This article gives an exposition of Euler's proof introducing the necessary concepts along the way.

Introduction

In this article, we present Euler's very beautiful proof that there are infinitely many prime numbers. In an earlier era, Euclid had proved this result in a simple yet elegant manner. His idea is easy to describe. Denoting the prime numbers by $p_1, p_2, p_3, \ldots$, such that $p_1 = 2, p_2 = 3, p_3 = 5, \ldots$, he supposes that there are n primes in all, the largest being p_n. He then considers the number N where

$$N = p_1 p_2 p_3 \ldots p_n + 1,$$

and asks what the prime factors of N could be. It is clear that N is indivisible by each of the primes $p_1, p_2, p_3, \ldots, p_n$ (indeed, $N \equiv 1 \pmod{p_i}$ for each $i, 1 \leq i \leq n$). Since every integer greater than 1 has a prime factorization, this forces into existence prime numbers other than the p_i. Thus there can be no largest prime number, and so the number of primes is infinite.

The underlying idea of Euler's proof is very different from that of Euclid's proof. In essence, he proves that the *sum of the reciprocals of the primes is infinite*; that is,

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} + \cdots = \infty.$$

In technical language, the series $\sum_i 1/p_i$ *diverges*. Obviously, this cannot possibly happen if there are only finitely many prime numbers. The infinitude of the primes thus follows as a corollary. Note that Euler's result is stronger than Euclid's.

Convergence and Divergence

A few words are necessary to explain the concepts of convergence and divergence of infinite series. A series $a_1 + a_2 + a_3 + \cdots$, is said to *converge* if the sequence of partial sums,

$$a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots,$$

approaches some limiting value, say L; we write, in this case, $\sum_1^\infty a_i = L$. If, instead, the sequence of partial sums grows without any bound, we say that the series *diverges*, and we write, in short[1], $\sum_1^\infty a_i = \infty$.

EXAMPLES.

- The series $1/1 + 1/2 + 1/4 + \cdots + 1/2^n + \cdots$ converges (the sum is 2, as is easily shown).
- The series $1/1 + 1/3 + 1/9 + \cdots + 1/3^n + \cdots$ converges (the sum in this case is 3/2).
- The series $1 + 1 + 1 + \cdots$ diverges (rather trivially).
- The series $1 - 1 + 1 - 1 + 1 - 1 + 1 - \cdots$ also fails to converge, because the partial sums assume the values 1, 0, 1, 0, 1, 0, ... and this sequence clearly does not possess a limit.
- A more interesting example: $1 - 1/2 + 1/3 - 1/4 + \cdots$ a careful analysis shows that it too is convergent, the limiting sum being ln 2 (the natural logarithm of 2).

Divergence of the Harmonic Series $\Sigma 1/i$

In order to prove Euler's result, namely, the divergence of $\sum 1/p_i$, we need to establish various subsidiary results. Along the way, we shall meet other examples of divergent series. To start with, we present the proof of the statement that

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty.$$

This rather non-obvious result is usually referred to as *the divergence of the harmonic series*. The proof given below is due to the Frenchman Nicolo Oresme and it dates to about 1350. We note the following sequence of equalities and inequalities:

$$\frac{1}{1} = \frac{1}{1},$$
$$\frac{1}{2} = \frac{1}{2},$$

[1] A statement of the form $\Sigma a_i = \infty$ is to be regarded as merely a short form for the statement that the sums $a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots$, do not possess any limit. It is important to note that ∞ is *not* to be regarded as a number! We shall however frequently use phrases of the type '$x = \infty$' (for various quantities x) during the course of this article. The meaning should be clear from the context.

$$\frac{1}{3}+\frac{1}{4}>\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$$

$$\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}>\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{1}{2},$$

$$\frac{1}{9}+\frac{1}{10}+\cdots+\frac{1}{16}>\frac{1}{16}+\frac{1}{16}+\cdots+\frac{1}{16}=\frac{1}{2},$$

and so on. We see that it is possible to group consecutive sets of terms of the series $1/1 + 1/2 + 1/3 + \cdots$, in such a manner that each group has a sum exceeding 1/2. Since the number of such groups is infinite, it follows that the sum of the whole series is itself infinite. (Note the crisp and decisive nature of the proof!).

Based on this proof, we make a more precise statement. Let $S(n)$ denote the sum

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}.$$

e.g., $S(3) = 11/6$. Generalizing from the reasoning just used, we find that

$$S(2^n) > 1+\frac{n}{2}. \tag{1}$$

(Please fill in the details of the proof on your own.) This means that by choosing n to be large enough, the value of $S(2^n)$ can be made to exceed any given bound. For instance, if we wanted the sum to exceed 100, then (1) assures us that a mere 2^{198} terms would suffice! This suggests the extreme slowness of growth of $S(n)$ with n. Nevertheless it does grow without bound; loosely stated, $S(\infty) = \infty$.

The result obtained above, (1), can also be written in the form,

$$S(n) > 1+\frac{1}{2}\log_2 n.$$

EXERCISE. Write out a proof of the above inequality.

A much more accurate statement can be made, but it involves calculus. We consider the curve Ω whose equation is $y = 1/x, x > 0$. The area of the region enclosed by Ω, the x-axis and the ordinates $x = 1$ and $x = n$ is equal to $\int_1^n \frac{1}{x}dx$, which simplifies to $\ln n$. Now let the region be divided into $(n-1)$ strips of unit width by the lines $x = 1, x = 2, x = 3, \ldots, x = n$ (see Figure 5.1).

Consider the region enclosed by Ω, the x-axis, and the lines $x = i - 1$, $x = i$. The area of this region lies between $1/i$ and $1/(i-1)$, because it can be enclosed between two rectangles of dimensions $1 \times 1/i$ and $1 \times 1/(i-1)$, respectively. (A quick examination of the graph will show why this is true.) By letting i take the values $2, 3, 4, \ldots n$, and adding the inequalities thus obtained, we find that

$$\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n} < \ln n < \frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{n-1}. \tag{2}$$

Relation (2) implies that

$$\ln n+\frac{1}{n} < \frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n} < \ln n+1, \tag{3}$$

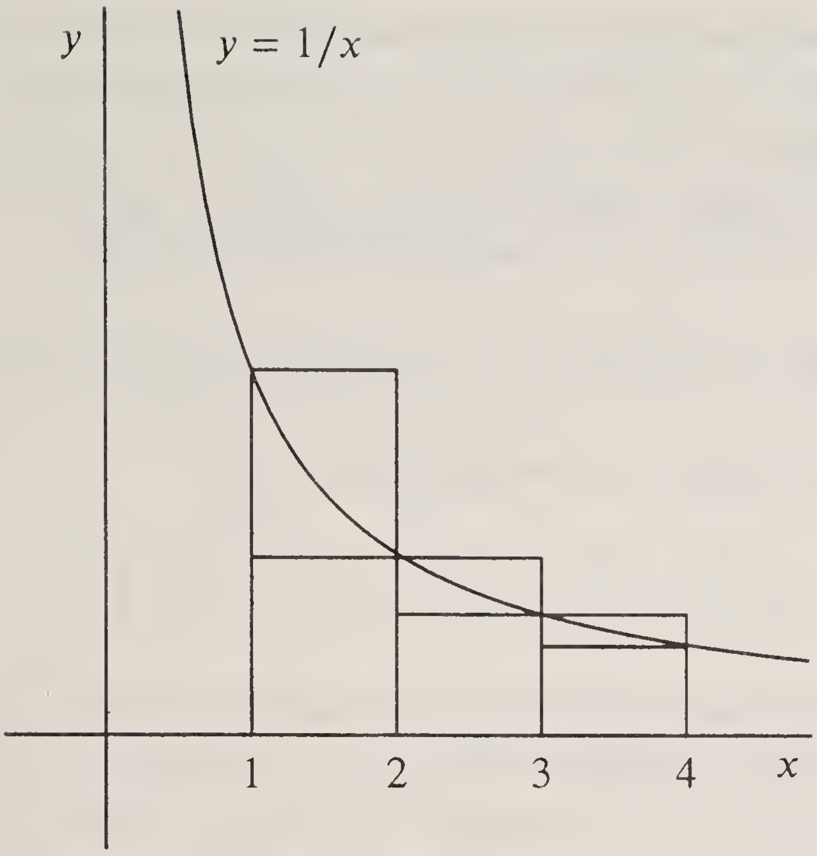


Figure 5.1 The figure shows how to bound $\ln n$ by observing that $\ln n$ is the area enclosed by the curve $y = 1/x$, the x-axis and the ordinates $x = 1$ and $x = n$.

and this means that we have an estimate for $S(n)$ (namely, $\ln n + 0.5$) that differs from the actual value by no more than 0.5. A still deeper analysis shows that for large values of n, an excellent approximation for $S(n)$ is $\ln n + 0.577$, but we shall not prove this result here. It is instructive, however, to check the accuracy of this estimate. Write $f(n)$ for $\ln n + 0.577$. We now find the following:

n =	10	100	1000	10000	100000
$S(n)$ =	2.92897	5.18738	7.48547	9.78761	12.0902
$f(n)$ =	2.87959	5.18217	7.48476	9.78734	12.0899

The closeness of the values of $f(n)$ and $S(n)$ for large values of n is striking. (The constant 0.577 is related to what is known as the Euler–Mascheroni constant.)

In general, when mathematicians find that a series $\sum a_i$ diverges, they are also curious to know how *fast* it diverges. That is, they wish to find a function, say $f(n)$, such that the ratio $(\sum_1^n a_i)/f(n)$ tends to 1 as $n \to \infty$. For the harmonic series $\sum 1/i$, we see that one such function is given by $f(n) = \ln n$. This is usually expressed by saying that the harmonic series diverges like the logarithmic function. We note in passing that this is a very slow rate of divergence, because $\ln n$ diverges more slowly than n^ε for any $\varepsilon > 0$, *no matter how small* ε *is*, in the sense that $\ln n/n^\varepsilon \to 0$ as $n \to \infty$ for every $\varepsilon > 0$. Obviously the function $\ln \ln n$ diverges still more slowly.

EXERCISE. Prove that if $a > 1$, then the series

$$\frac{1}{1^a} + \frac{1}{2^a} + \frac{1}{3^a} + \cdots$$

converges. (The conclusion holds no matter how close a is to 1, but it does not hold for $a = 1$ or $a < 1$, a curious state of affairs!) Further, use the methods of integral calculus (and the fact that for $a \neq 1$, the integral of $1/x^a$ is $x^{(1-a)}/(1-a)$ to show that the sum of the series lies between $1/(a-1)$ and $a/(a-1)$.

The fact that the sum $1/1 + 1/2^2 + 1/3^2 + \cdots$ is finite can be shown in another manner that is both elegant and elementary. We start with the inequalities, $2^2 > 1 \times 2, 3^2 > 2 \times 3, 4^2 > 3 \times 4, \ldots,$ and deduce from these that

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots < 1 + \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \frac{1}{3 \times 4} + \cdots.$$

The sum on the right side can be written in the form,

$$1 + \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots, \tag{4}$$

which (after a whole feast of cancellations) simplifies to $1 + 1/1$, that is, to 2. (This is sometimes described by stating that the series 'telescopes' to 2.) Therefore the sum $1 + 1/2^2 + 1/3^2 + 1/4^2 + \cdots$ is less than 2. We now call upon a theorem of analysis which states that if the partial sums of any series form an increasing sequence and are at the same time bounded, that is, they do not exceed some fixed number, then they possess a limit. We conclude, therefore, that the series $\sum 1/i^2$ does possess a finite sum which lies between 1 and 2.

The divergence of the harmonic series was independently proved by Johann Bernoulli in 1689 in a completely different manner. His proof is worthy of deep study, as it shows the counter-intuitive nature of infinity.

Bernoulli starts by assuming that the series $1/2 + 1/3 + 1/4 + \cdots$ (note that he starts with 1/2 rather 1/1) does have a finite sum, which he calls S. He now proceeds to derive a contradiction in the following manner. He rewrites each term occurring in S thus:

$$\frac{1}{3} = \frac{2}{6} = \frac{1}{6} + \frac{1}{6}, \quad \frac{1}{4} = \frac{3}{12} = \frac{1}{12} + \frac{1}{12} + \frac{1}{12}, \ldots,$$

and more generally,

$$\frac{1}{n} = \frac{n-1}{n(n-1)} = \frac{1}{n(n-1)} + \frac{1}{n(n-1)} + \cdots + \frac{1}{n(n-1)},$$

with $(n-1)$ fractions on the right side. Next he writes the resulting fractions in an array as shown below:

$$\begin{array}{cccccccc}
1/2 & 1/6 & 1/12 & 1/20 & 1/30 & 1/42 & 1/56 & \cdots \\
 & 1/6 & 1/12 & 1/20 & 1/30 & 1/42 & 1/56 & \cdots \\
 & & 1/12 & 1/20 & 1/30 & 1/42 & 1/56 & \cdots \\
 & & & 1/20 & 1/30 & 1/42 & 1/56 & \cdots \\
 & & & & 1/30 & 1/42 & 1/56 & \cdots \\
 & & & & & 1/42 & 1/56 & \cdots \\
 & & & & & & 1/56 & \cdots
\end{array}$$

Note that the column sums are just the fractions $1/2, 1/3, 1/4, 1/5, \ldots$; thus, S is the sum of all the fractions occurring in the array. Bernoulli now sums the rows using the telescoping technique used above (see equation (4)). Assigning symbols to the row sums as shown below,

$$\begin{aligned}
A &= \frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\frac{1}{42}+\frac{1}{56}+\cdots,\\
B &= \frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\frac{1}{42}+\frac{1}{56}+\cdots,\\
C &= \frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\frac{1}{42}+\frac{1}{56}+\cdots,\\
D &= \frac{1}{20}+\frac{1}{30}+\frac{1}{42}+\frac{1}{56}+\cdots,
\end{aligned}$$

he finds that:

$$\begin{aligned}
A &= \left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\left(\frac{1}{4}-\frac{1}{5}\right)+\cdots\\
&= 1,\\
B &= \left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\left(\frac{1}{4}-\frac{1}{5}\right)+\left(\frac{1}{5}-\frac{1}{6}\right)+\cdots\\
&= \frac{1}{2},\\
C &= \frac{1}{3}, \qquad \text{(arguing likewise)},\\
D &= \frac{1}{4},
\end{aligned}$$

and so on. Thus the sum S, which we had written in the form $A+B+C+D+\cdots$, turns out to be equal to

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots.$$

Now this looks disappointing—just as things were beginning to look promising! We seem to have just recovered the original series after a series of very complicated steps. But in fact something significant has happened: *an extra* '1' *has entered the series*. At the start we had defined S to be $1/2+1/3+1/4+\cdots$; now we find that S equals $1+1/2+1/3+1/4+\cdots$. This means that $S=S+1$. However, no finite number can satisfy such an equation. Conclusion: $S=\infty$!

There are many other proofs of this beautiful result, but I shall leave you with the pleasant task of coming up with them on your own. Along the way you could set yourself the task of proving that each of the following sums diverge:

- $1/1+1/3+1/5+1/7+1/9+\cdots$;
- $1/1+1/11+1/21+1/31+1/41+\cdots$;
- $1/a+1/b+1/c+1/d+\cdots$, where $a,b,c,d,\ldots,$ are the successive terms of any increasing arithmetic progression of positive real numbers.

Elementary Results

The next result that we shall need is the so-called fundamental theorem of arithmetic: *every positive integer greater than* 1 *can be expressed in precisely one way as a product of prime numbers.* We shall not prove this very basic theorem of number theory. For a proof, please refer to any of the well-known texts on number theory, e.g., the text by Hardy and Wright, or the one by Niven and Zuckermann.

We shall also need the following rather elementary results: (i) if k is any integer greater than 1, then

$$\frac{1}{1-1/k} = 1 + \frac{1}{k} + \frac{1}{k^2} + \frac{1}{k^3} + \frac{1}{k^4} + \cdots, \tag{5}$$

which follows by summing the geometric series on the right side, and (ii) if a_i, b_j are any quantities, then

$$\left(\sum_i a_i\right)\left(\sum_j b_j\right) = \sum_{i,j} a_i b_j,$$

where, in the sum on the right, each pair of indices (i, j) occurs *precisely once*.

Now consider the following two equalities, which are obtained from (5) using the values $k = 2$, $k = 3$:

$$\frac{1}{1-1/2} = 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots,$$
$$\frac{1}{1-1/3} = 1 + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \cdots.$$

We multiply together the corresponding sides of these two equations. On the left side we obtain $2 \times 3/2 = 3$. On the right side we obtain the product

$$(1 + 1/2 + 1/2^2 + 1/2^3 + \cdots) \times (1 + 1/3 + 1/3^2 + 1/3^3 + \cdots).$$

Expanding the product, we obtain:

$$\begin{aligned} &1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots + \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \cdots \\ &+ \frac{1}{6} + \frac{1}{12} + \frac{1}{24} + \cdots + \frac{1}{18} + \frac{1}{36} + \frac{1}{72} + \cdots, \end{aligned}$$

that is, we obtain the sum of the reciprocals of all the positive integers that have only 2 and 3 among their prime factors. The fundamental theorem of arithmetic assures us that each such integer occurs *precisely once* in the sum on the right side. Thus we obtain a nice corollary: If A denotes the set of integers of the form $2^a\, 3^b$, where a and b are non-negative integers, then

$$\sum_{z \in A} \frac{1}{z} = 3.$$

If we multiply the left side of this relation by $(1 + 1/5 + 1/5^2 + 1/5^3 + \cdots)$ and the right side by $3/(1 - 1/5)$, we obtain the following result:

$$\sum_{z \in B} \frac{1}{z} = \frac{3}{1 - 1/5} = \frac{15}{4},$$

where B denotes the set of integers of the form $2^a\,3^b\,5^c$, where a, b and c denote non-negative integers.

Continuing this line of argument, we see that infinitely many such statements can be made, for example:

- If C denotes the set of positive integers of the form $2^a\,3^b\,5^c\,7^d$, where a, b, c and d are non-negative integers, we then have $\sum_{z \in C} 1/z = (15/4)(7/6) = 35/8$.
- If D denotes the set of positive integers of the form $2^a\,3^b\,5^c\,7^d\,11^e$, then $\sum_{z \in D} 1/z = (35/8)(11/10) = 77/16$.

Infinitude of the Primes

Suppose now that there are only finitely many primes, say $p_1, p_2, p_3, \ldots, p_n$, where $p_1 = 2, p_2 = 3, p_3 = 5, \ldots$. We consider the product

$$\frac{1}{1 - 1/2}\,\frac{1}{1 - 1/3}\,\frac{1}{1 - 1/5}\cdots\frac{1}{1 - 1/p_n}$$

This is obviously a finite number, being the product of finitely many non-zero fractions. Now this product also equals

$$\left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\right) \times \left(1 + \frac{1}{3} + \frac{1}{3^2} + \cdots\right) \times$$
$$\left(1 + \frac{1}{5} + \frac{1}{5^2} + \cdots\right) \times \cdots \times \left(1 + \frac{1}{p_n} + \frac{1}{p_n^2} + \cdots\right).$$

When we expand this product, we find, by continuing the line of argument developed above, that we obtain *the sum of the reciprocals of all the positive integers*. To see why, we need to use the fundamental theorem of arithmetic and the assumption that $2, 3, 5, \ldots, p_n$ are *all* the primes that exist; these two statements together imply that every positive integer can be expressed *uniquely* as a product of non-negative powers of the n primes $2, 3, 5, \ldots, p_n$. From this it follows that the expression on the right side is precisely the sum

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots,$$

written in some permuted order. But by the Oresme–Bernoulli theorem, the latter sum is infinite! So we have a contradiction: the finite number

$$\frac{1}{1 - 1/2}\,\frac{1}{1 - 1/3}\,\frac{1}{1 - 1/5}\cdots\frac{1}{1 - 1/p_n}$$

has been shown to be infinite—an absurdity! The only way out of this contradiction is to drop the assumption that there are only finitely many prime numbers. Thus we reach the desired objective, namely, that of proving that there are infinitely many prime numbers.

Note that, as a bonus, there are several formulas that drop out of this analysis, more or less as corollaries. For instance, we find that

$$\frac{1}{1-1/2^2}\,\frac{1}{1-1/3^2}\,\frac{1}{1-1/5^2}\cdots = 1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots,$$

that is, the infinite product and the infinite sum both converge to the same (finite) value. By a stunning piece of reasoning, including a few daring leaps that would leave today's mathematicians gasping for breath, Euler showed that both sides of the above equation are equal to $\pi^2/6$. Likewise, we find that

$$\frac{1}{1-1/2^4}\,\frac{1}{1-1/3^4}\,\frac{1}{1-1/5^4}\cdots = 1+\frac{1}{2^4}+\frac{1}{3^4}+\frac{1}{4^4}+\cdots,$$

and this time both sides converge to $\pi^4/90$. Euler proved all this and much much more; it is not for nothing that he is at times referred to as *analysis incarnate!*

The Divergence of $\sum 1/p$

As mentioned earlier, Euler showed in addition that the sum

$$\sum_{i\geq 1}\frac{1}{p_i}=\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\cdots$$

is itself infinite. We are now in a position to obtain this beautiful result. For any positive integer $n \geq 2$, let P_n denote the set of prime numbers less than or equal to n. We start by showing that

$$\prod_{p\in P_n}\frac{1}{1-1/p} > \sum_{j=1}^{n}\frac{1}{j}. \tag{6}$$

Our strategy will be a familiar one. We write down the following inequality for each $p \in P_n$, which follows from (5):

$$\frac{1}{1-1/p} > 1+\frac{1}{p}+\frac{1}{p^2}+\frac{1}{p^3}+\cdots+\frac{1}{p^n}.$$

The '>' sign holds because we have left out all the positive terms that follow the term $1/p^n$. Multiplying together the corresponding sides of all these inequalities ($p \in P_n$), we obtain:

$$\prod_{p\in P_n}\frac{1}{1-1/p} > \prod_{p\in P_n}\left(1+\frac{1}{p}+\frac{1}{p^2}+\frac{1}{p^3}+\cdots+\frac{1}{p^n}\right).$$

When we expand the product on the right side, we obtain a sum of the form $\sum_{j\in A} 1/j$ for some set of positive integers A. This set certainly includes all the integers from 1 to n because the set P_n contains all the prime numbers between 1 and n. Inequality (6) thus follows immediately.

Next, we already know (see equation (3)) that

$$\sum_{j=1}^{n} \frac{1}{j} > \ln n + \frac{1}{n} > \ln n. \tag{7}$$

Combining (6) and (7), we obtain the following inequality:

$$\prod_{p \in P_n} \frac{1}{1 - 1/p} > \ln n.$$

Taking logarithms on both sides, this translates into the statement

$$\sum_{p \in P_n} \ln \left(\frac{1}{1 - 1/p} \right) > \ln \ln n. \tag{8}$$

Our task is nearly over. It only remains to relate the sum $\sum_{p \in P_n} 1/p$ with the sum on the left side of (8). We accomplish this by showing that the inequality

$$\frac{7x}{5} > \ln \frac{1}{1 - x} \tag{9}$$

holds for $0 < x \leq 1/2$.

To see why (9) is true, draw the graph of the curve Γ whose equation is $y = \ln(1/(1 - x))$, over the domain $-\infty < x < 1$, (see Figure 5.2). Note that Γ passes through the origin and is convex over its entire extent. (PROOF: Write $f(x) = -\ln(1 - x)$; then $f'(x) = 1/(1 - x)$ and $f''(x) = 1/(1 - x)^2 > 0$ for all $x < 1$.)

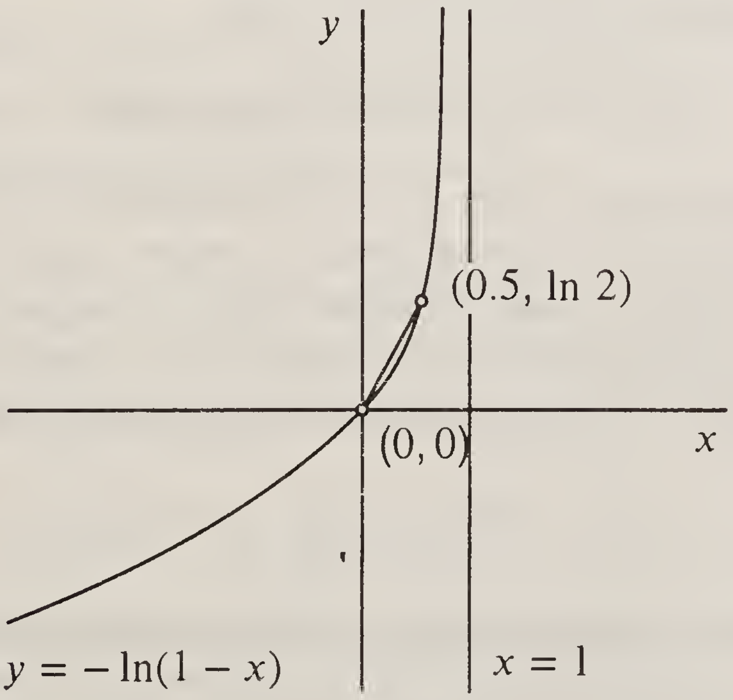


Figure 5.2 The graph shows that for $0 \leq x \leq 1/2$, we have $(2 \ln 2)\, x \geq \ln(1/(1 - x))$ for $0 \leq x \leq 1/2$.

The convexity of Γ implies that the chord joining the points A$(0,0)$ and B$(1/2, \ln 2)$ lies completely *above* the curve. The equation of AB is $y = (2\ln 2)\,x$, so that over the range $0 \le x \le 1/2$ we have the inequality:

$$(2 \ln 2)\,x \ge \ln\left(\frac{1}{1-x}\right).$$

Since $\ln 2 \approx 0.69315 < 0.7 = 7/10$, (9) follows.

Inequality (9) implies that

$$x > \frac{5}{7}\ln\left(\frac{1}{1-x}\right)$$

for $x = 1/2, x = 1/3, x = 1/5, \ldots$. Therefore, by addition,

$$\sum_{p\in P_n} \frac{1}{p} > \frac{5}{7}\left(\sum_{p\in P_n} \ln\frac{1}{1-1/p}\right). \tag{10}$$

Combining (8) and (10), we deduce that

$$\sum_{p\in P_n} \frac{1}{p} > \frac{5}{7}\ln\ln n.$$

As $n \to \infty$, the right side diverges to infinity, therefore so does the left side; so we reach our desired objective, that of showing the divergence of $\sum_i 1/p_i$.

An Alternative Proof

Here is an alternative proof of the claim that $\sum_i 1/p_i$ diverges. The proof has been written in an 'old-fashioned' style and purists will protest. Nevertheless, we shall present the proof and let the readers judge for themselves. Let S denote the sum $\sum_i 1/p_i$. We shall make use of the following result:

$$e^x \ge 1 + x \text{ for all real values of } x,$$

with equality holding precisely when $x = 0$. The graphs of e^x and $1 + x$ show why this is true; the former graph is convex over its entire extent (examine the second derivative of e^x to see why), while the latter, a line, is tangent to the former at the point (0, 1), and lies entirely below it everywhere else. Substituting the values $x = 1/2, x = 1/3, x = 1/5, \ldots$, successively into this inequality, we find that

$$e^{1/2} > 1 + \frac{1}{2}, \quad e^{1/3} > 1 + \frac{1}{3}, \quad e^{1/5} > 1 + \frac{1}{5}, \cdots.$$

Multiplying together the corresponding sides of these inequalities, we obtain:

$$e^S > \left(1+\frac{1}{2}\right)\left(1+\frac{1}{3}\right)\left(1+\frac{1}{5}\right)\cdots.$$

The infinite product on the right side yields the following series:

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{10}+\frac{1}{11}+\frac{1}{13}+\frac{1}{14}+\frac{1}{15}+\cdots.$$

This series is the sum of the reciprocals of all the positive integers whose prime factors are all distinct; equivalently, the positive integers that have no squared factors. These numbers are sometimes referred to as the *quadratfrei* or *square-free* numbers. Let Q denote this sum. We shall show that this series itself diverges, in other words, that $Q = \infty$. This will immediately imply that $S = \infty$ (for $e^S > Q$), and Euler's result will then follow.

We consider the product

$$Q \times \left(1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots\right).$$

This product, when expanded, gives the following series:

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots,$$

that is, we obtain the harmonic series. To see why, note that every positive integer n can be *uniquely* written as a product of a square-free number and a square; for example, $1000 = 10 \times 10^2, 2000 = 5 \times 20^2, 1728 = 3 \times 24^2$, and so on. Now when we multiply

$$\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{10}+\frac{1}{11}+\frac{1}{13}+\frac{1}{14}+\frac{1}{15}+\cdots\right)$$

with

$$\left(1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots\right)$$

we find, by virtue of the remark just made, that the reciprocal of each positive integer n occurs *precisely once* in the expanded product. This explains why the product is just the harmonic series. Now recall that the sum

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\cdots$$

is finite (indeed, we have shown that it is less than 2). It follows that

$$Q \times \text{ (some finite number) } = \infty.$$

Therefore $Q = \infty$, and Euler's result ($\sum_i 1/p_i = \infty$) follows. QED!

Readers who are unhappy with this style of presentation, in which ∞ is treated as an ordinary real number, will find it an interesting (but routine) exercise to rewrite the proof to accord with more exacting standards of rigour and precision.

XVI. *Summa seriei infinita harmonicè progressionalium,* $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}$ *& c.est infinita.*

Id primus deprehendit Frater: inventa namque per præced. Summa seriei $\frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}$, & c. visurus porrò, quid emergeret ex ista serie, $\frac{1}{2}+\frac{2}{6}+\frac{3}{12}+\frac{4}{20}+\frac{5}{30}$, & c. Si resolveretur methodo Prop. XIV. collegit p opositionis veritatem ex absurditate manifesta, quæ sequeretur, si summa. Seriei harmonicæ finita statueretur. Animadvertit enim,

Seriem A, $\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}$, & c.$\infty$. (fractionibus singulis in alias, quarum numeratores sunt 1, 2, 3, 4, & c. transmutatis)

Seriei B, $\underline{\frac{1}{2}+\frac{2}{6}+\frac{3}{12}+\frac{4}{20}+\frac{5}{30}+\frac{6}{42}, \text{ \& c.} \infty C + D + E + F, \text{\& c.}}$

$$\left.\begin{array}{llll}
C \cdot \frac{1}{2}+\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\frac{1}{42}, & \text{\& c.} \infty & \text{per præc.} & \frac{1}{1} \\
D \cdots +\frac{1}{6}+\frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\frac{1}{42}, & \text{\& c.} \infty & C-\frac{1}{2} & \infty\frac{1}{2} \\
E \cdots\cdots + \frac{1}{12}+\frac{1}{20}+\frac{1}{30}+\frac{1}{42}, & \text{\& c.} \infty & D-\frac{1}{6} & \infty\frac{1}{3} \\
F \cdots\cdots\cdots\cdots +\frac{1}{20}+\frac{1}{30}+\frac{1}{42}, & \text{\& c.} \infty & E-\frac{1}{1} & \infty\frac{1}{4} \\
 & \text{\& c.} \infty & & \text{\&c.}
\end{array}\right\} \infty G; \text{ unde}$$

sequitur, seriem G ∞ A, totum parti, si summa finita effet. Ego

Johann's divergence proof, from Jakob's *Tractatus de Seriebus Infinitis*, republished in 1713. (From page 197 of *Journey through Genius* by William Dunham.)

Conclusion

A much deeper—but also more difficult—analysis shows that the sum $1/p_1 + 1/p_2 + 1/p_3 + \cdots + 1/p_n$ is approximately equal to ln ln n. This is usually stated in the following form: as n tends to ∞, the fraction

$$\frac{1/p_1 + 1/p_2 + 1/p_3 + \cdots + 1/p_n}{\ln \ln n}$$

tends to 1. This is indeed a striking result, reminiscent of the earlier result that $1/1 + 1/2 + 1/3 + \cdots + 1/n$ is approximately equal to ln n. It shows the staggeringly slow rate of divergence of the sum of the reciprocals of the primes. The harmonic series $\sum_i 1/i$, diverges slowly enough—to achieve a sum of over 100, for instance, we would need to add more than 10^{43} terms, so it is certainly not a job that one can leave to finish off over a weekend. (Do you see where the number 10^{43} comes from?) On the other hand, to achieve a sum of over 100 with the series $\sum_i 1/p_i$, we need to add something like $10^{10^{43}}$ terms!! This number is so stupendously large that it is a hopeless task to make any visual image of it. Certainly there is no magnitude even remotely comparable to it in the whole of the known universe.

Suggested Reading

[1] G H Hardy, E M Wright. *An Introduction to the Theory of Numbers.* 4th ed. Oxford. Clarendon Press, 1960.

[2] Ivan Niven, Herbert S Zuckermann. *An Introduction to the Theory of Numbers.* Wiley Eastern Ltd., 1989.

[3] Tom Apostol. *An Introduction to Analytic Number Theory.* Narosa Publishing House, 1979.

SHAILESH A SHIRALI
Rishi Valley School
Rishi Valley 517 352
Andhra Pradesh

6

On Fermat's Two Squares Theorem

Shailesh A Shirali

Introduction

The purpose of this chapter is to present a proof of the two squares theorem: *every prime of the form* 1 (mod 4) *can be written as a sum of two squares.* The theorem was first stated by Fermat (as usual, with no proof!) and later proved by Euler. The proof given here is an elaboration of the one presented by Don Zagier in a crisp note that appeared in *The American Mathematical Monthly*, Vol. 97, # 2 (Feb 1990). As Zagier himself remarks in his paper, his proof is not constructive. In the final section we make an interesting conjecture which, if correct, will provide a constructive version of Zagier's proof.

Throughout, p refers to a fixed prime of the form 1 (mod 4), while **N** refers to the set of positive integers. For a finite set X, $|X|$ denotes the cardinality of X.

Proof of the Two Squares Theorem

The proof hinges on a study of the solutions in positive integers of the equation $x^2 + 4yz = p$. Let S_p denote the solution set:

$$S_p = \{(x, y, z) \in \mathbf{N}^3 : x^2 + 4yz = p\}. \tag{1}$$

It is easy to verify that S_p is non-empty (for $(1, 1, \frac{p-1}{4}) \in S_p$) and finite. We shall show that $|S_p|$ is odd.

Consider the following relations:

$$x^2 + 4yz = (x + 2z)^2 + 4z(y - x - z) = (2y - x)^2 + 4y(x - y + z). \tag{2}$$

From this we see that α, β, γ as defined by

$$\alpha(x, y, z) = (x + 2z, z, y - x - z), \tag{3}$$

$$\beta(x, y, z) = (2y - x, y, x - y + z), \tag{4}$$
$$\gamma(x, y, z) = (x - 2y, x - y + z, y), \tag{5}$$

are maps of the solution set in real numbers of $x^2 + 4yz = p$ into itself; still better, they are *unimodular* maps — they permute the integer solutions amongst themselves. (This can be checked by observing that the matrices corresponding to the three maps are all unimodular, that is, they have determinant ± 1.) Since our interest lies chiefly in the positive integral solutions, we define subsets A_p, B_p and C_p of S_p as follows:

$$A_p = \{(x, y, z) \in S_p, x < y - z\}, \tag{6}$$
$$B_p = \{(x, y, z) \in S_p, y - z < x < 2y\}, \tag{7}$$
$$C_p = \{(x, y, z) \in S_p, 2y < x\}. \tag{8}$$

We now make the following observations which are easy to verify.

- $S_p = A_p \cup B_p \cup C_p$, that is, A_p, B_p and C_p constitute a *partition* of S_p. Equality cannot hold in any of the defining inequalities because p is prime. Moreover, $(1, 1, \frac{p-1}{4}) \in B_p$.
- α maps A_p into C_p and γ maps C_p into A_p; moreover, α and γ are inverses of one another. Since A_p and C_p are finite sets, it follows that $|A_p| = |C_p|$.
- β maps B_p into itself, and β is its own inverse (it is an *involution*), so it pairs up elements of B_p with one another, except possibly for the fixed points—the triples (x, y, z) which get mapped to themselves; these have no mates and stand alone.
- β has just one fixed point. For, if (x, y, z) is a fixed point, then $(2y - x, y, x - y + z) = (x, y, z)$, so $x = y$. This gives $p = x(x + 4z)$, implying that $x = 1$ and $x + 4z = p$ since p is prime. It follows that $(1, 1, \frac{p-1}{4})$ is the sole fixed point of β.
- B_p is odd, for β is an involution on B_p with just one fixed point. In turn this implies that $|S_p|$ is odd (because $|A_p| = |C_p|$).

Observe that for each element $(x, y, z) \in S_p$, its 'mate' (x, z, y) also lies in S_p. Since S_p has an odd number of elements, it follows that S_p must contain an 'odd man out' which is its own 'mate'. If (r, s, s) is such an element of S_p, then $p = r^2 + (2s)^2$, and we are through.

Towards a Constructive Proof

Note that the proof presented is not constructive—it provides no clue as to how the desired (r, s) can be computed for a given p. (Curiously, this is true for most known proofs of the theorem.) However, the argument used does suggest the possibility of an algorithmic proof. I have empirically found that the following algorithm 'works', in the sense that it always seems to terminate. However, I have not been able to devise a proof of termination; if found, then a constructive proof of the two squares theorem is at hand.[1] Perhaps some reader would like to take up the challenge and settle the matter.

Consider the set I_p of integer triples (x, y, z) for which $x^2 + 4yz = p$. The set is non–empty, for $(1,.1, \frac{p-1}{4}) \in I_p$. Our objective is to find a triple in I_p of the form (r, s, s); this would immediately provide the desired representation of p as a sum of two squares ($p = r^2 + (2s)^2$). Towards this end we define a function $f\colon I_p \to I_p$ as follows:

$$f(x, y, z) = \begin{cases} (x + 2z, y - z - x, z) & \text{if } z + x < y, \\ (2y - x, z + x - y, y) & \text{if } z + x > y. \end{cases}$$

EXAMPLE. Let $p = 17$; then $f(1, 1, 4) = (1, 4, 1)$ and $f(1, 4, 1) = (3, 2, 1)$.

We now compute the orbit of the triple $(1, 1, \frac{p-1}{4})$ under action by f. If at some stage we reach a triple of the form (r, s, s) we terminate the computation. The curious thing is that we always seem to reach such a triple. Listed below are the initial segments of the orbits for a few p's. In each case we stop when the desired triple is reached.

- $\underline{p = 17}$

  $(1, 1, 4)$ $(1, 4, 1)$, $(3, 2, 1)$, $(1, 2, 2)$; result: $17 = 1^2 + 4^2$.
- $\underline{p = 29}$

  $(1, 1, 7)$, $(1, 7, 1)$, $(3, 5, 1)$, $(5, 1, 1)$; result: $29 = 5^2 + 2^2$.
- $\underline{p = 41}$

  $(1, 1,10)$, $(1, 10, 1)$, $(3, 8, 1)$, $(5, 4, 1)$, $(3, 2, 4)$, $(1, 5,2)$, $(5, 2, 2)$; result: $41 = 5^2 + 4^2$.
- $\underline{p = 53}$

  $(1, 1, 13)$, $(1, 13, 1)$, $(3, 11, 1)$, $(5, 7, 1)$, $(7, 1, 1)$; result: $53 = 7^2 + 2^2$.
- $\underline{p = 109}$

  $(1, 1, 27)$, $(1, 27, 1)$, $(3, 25, 1)$, $(5, 21, 1)$, $(7, 15, 1)$, $(9, 7, 1)$, $(5, 3, 7)$, $(1, 9, 3)$, $(7, 5, 3)$, $(3, 5, 5)$; result: $109 = 3^2 + 10^2$.

Any takers?

Further Remarks

- Weil writes, in his book (see Suggested Reading) that "all known proofs begin … by showing that -1 is a quadratic residue of $p = 4n + 1$". This being so, Zagier's proof is rather atypical.

[1] This conjecture was subsequently settled in the affirmative by B Bagchi; see Chapter 7.

- The theorem was stated by Fermat in 1640; he never published any proof but in all likelihood did possess one, probably based on the principle of infinite descent (which itself is one of Fermat's inventions). The first published proof, by Euler, appeared in the 1740's; it too uses the principle of infinite descent.

Suggested Reading

[1] Andre Weil. *Number Theory: An Approach Through History*, 1984.

SHAILESH A SHIRALI
Rishi Valley School
Rishi Valley 517 352
Andhra Pradesh

7

Fermat's Two Squares Theorem Revisited

B Bagchi

The Two Squares Theorem

Throughout this article, p is a prime such that $p \equiv 1 \pmod 4$. $\mathbb{N}$ and Z will denote, as usual, the set of all natural numbers (excluding zero) and the set of all integers (positive, negative or zero), respectively. Recall that the celebrated two squares theorem (first stated by Fermat and proved by Euler) says that p can be written as a sum of two perfect squares. Clearly one of these two squares must be even (and the other one is odd). Therefore, this theorem may be formulated by saying that there exists $(x, y) \in \mathbb{N} \times \mathbb{N}$ such that $x^2 + 4y^2 = p$. Any such pair (x, y) will be referred to as a representation of p. (Actually, as is well known, the representation is unique. For proof, see for instance Niven and Zuckerman in Suggested Reading.)

Permutations

G H Hardy writes that the two squares theorem 'is ranked, very justly, as one of the finest in arithmetic'. So it comes as a surprise to learn that its finest proof was found only in 1990. In that year, D Zagier modified a proof of the two squares theorem due to Heathbrown to create a remarkably short and elegant proof. Although Zagier's proof was presented in detail by Shirali in *Resonance* (see Suggested Reading), we shall begin with a brief account of this proof. To do so, we need to recall some facts about permutations.

If X is a finite set, then by a permutation of X we mean a function from X into itself under which each element of X has a unique pre-image. If π and σ are any two permutations of X, then we can form their 'product' $\pi\sigma$ by composition: $\pi\sigma(x) := \pi(\sigma(x))$, x in X. If X is of size n, there are only $n!$ permutations of X and they form a group with this product rule. (Though, strictly speaking, we need no group theory for this article, familiarity with the elements of this theory will still be useful.) Since we have defined the product of any two permutations, in particular we can form the powers $\pi = \pi^1, \pi^2, \ldots,$ of any given permutation π. Since there are only finitely many distinct permutations of X, some two of the powers of π must actually be equal. By cancellation, it follows that there must exist a natural number m such that π^m is the identity permutation id fixing all elements of X. The smallest such number is called the order of π. A permutation of order two is called an *involution*.

Any permutation π of X breaks up ('partitions') X into one or more parts such that two elements x and y of X are in the same part if and only if some power of π takes x to y. These parts are called the *orbits* of π. The singleton orbits are just the fixed points of π. A permutation of X is said to be transitive on X if it has only one orbit (namely, the whole of X).

It is easy to convince oneself that the size of any orbit of a permutation divides the order of the permutation. In particular, if the permutation π has prime order q, then (as 1 and q are the only divisors of q) each orbit is either a fixed point or has size q. It follows that, in this case, the number of fixed points of π is congruent modulo q to the size n of X. Hence π has a fixed point if n is not a multiple of q. As a special case ($q = 2$) of this observation, we see that an involution of X has a fixed point in X if X is an odd set (i.e., the number of elements of X is odd). This is the key fact which makes Zagier's proof (and its constructive versions presented here) work.

Zagier's Proof

Now we come to Zagier's proof. Let S denote the subset of $I\!N \times I\!N \times I\!N$ defined by

$$S = \{(x, y, z) \in I\!N \times I\!N \times I\!N : x^2 + 4yz = p\}.$$

Clearly S is a finite set. Zagier defines two involutions α and β of S by

$$\alpha(x, y, z) = \begin{cases} (x + 2z,\ z,\ y - x - z) & \text{if } x < y - z, \\ (2y - x,\ y,\ x + z - y) & \text{if } y - z < x < 2y, \\ (x - 2y,\ x + z - y,\ y) & \text{if } x > 2y. \end{cases}$$

$$\beta(x, y, z) = (x, z, y).$$

The involution α of the finite set S has a unique fixed point (namely $(1, 1, \frac{p-1}{4})$). It follows that S is an odd set. Therefore, the involution β of the odd set S must have an odd number (hence at least one) of fixed points in S. But $(x, y) \mapsto (x, y, y)$ is a bijection of the set of representations of p onto the set of fixed points of β. Hence p has at least one representation (as a sum of two squares). This completes Zagier's proof of the two squares theorem.

Shirali's Conjecture

Zagier notes in his paper that his proof 'is not constructive: it does not give a method to actually find the representation of p as a sum of two squares'. Perhaps provoked by this statement, S A Shirali gave a conjectural way to 'constructivize' this proof.

Shirali's conjecture may be phrased as follows. Define a finite subset $\widehat{S}$ of $Z \times I\!N \times I\!N$ by

$$\widehat{S} = \{(x, y, z) \in Z \times I\!N \times I\!N : x + y > z \text{ and } x^2 + 4yz = p\}.$$

Define a function $\hat{\gamma} : \widehat{S} \to \widehat{S}$ by

$$\hat{\gamma}(x, y, z) = \begin{cases} (x + 2z,\ y - x - z,\ z) & \text{if } x + z < y, \\ (2y - x,\ x + z - y,\ y) & \text{if } x + z > y. \end{cases}$$

Then, Shirali conjectures that the orbit of the point $(1, \frac{p-1}{4}, 1)$ under $\hat{\gamma}$ contains a point of the form (x, y, y). That is, to obtain a point $(x, y, y) \in \widehat{S}$ (and hence a square plus square representation of p), begin with the point $(1, \frac{p-1}{4}, 1)$ and look at the successive iterates (powers) of $\hat{\gamma}$ on this point until a point (x, y, y) is obtained.

(Actually, Shirali defines his function on the (infinite) set of all points (x, y, z) in $Z \times Z \times Z$ satisfying $x^2 + 4yz = p$, and proposes to begin with the α-fixed point $(1, 1, \frac{p-1}{4})$. However, we observed that this function fixes the finite subset $\widehat{S}$ introduced above and on this subset restricts it to $\hat{\gamma}$ as defined. Though the α-fixed point itself does not belong to this subset, its image under Shirali's original function is $(1, \frac{p-1}{4}, 1)$, which does belong. Therefore, our formulation of the conjecture is entirely equivalent to Shirali's original formulation.)

A Constructive Version of Zagier's Proof

Notice that the function $\hat{\gamma}$ is a 'perturbation' of the permutation $\gamma := \alpha\beta$ of S obtained by composing Zagier's involutions α and β. So it is natural to ask if Shirali's conjecture is valid with $\hat{\gamma}$ replaced by γ. In the following theorem, we show that this modified conjecture is indeed correct. Note that we now stay within the set S, and this is closer to Zagier's original proof.

Theorem. Let k denote the size of the orbit T under $\gamma := \alpha\beta$ which contains the α-fixed point a. Then k is odd; T contains a unique β-fixed point b and is given by the formula $b = \gamma^{(k-1)/2}(a)$. In fact, the orbit T satisfies the symmetry relation $\gamma^{k-1-n}(a) = \beta(\gamma^n(a))$ for $0 \leq n \leq k - 1$.

Thus, to obtain a β-fixed point (x, y, y) (and hence the representation $p = x^2 + (2y)^2$), begin with the α-fixed point and iterate $\alpha\beta$ on it; in a finite number of steps you will reach a β-fixed point. This theorem shows that exactly half the orbit has to be traversed before this point is reached; the remaining half may be found (in reverse order) simply by applying β to the first half.

Proof. Since α and β are involutions, α 'normalises' γ: $\alpha\gamma\alpha^{-1} = \beta\alpha = \gamma^{-1}$. Therefore, α maps the orbits of γ to orbits of γ. (To see this, let s_1 and s_2 be two points from a common γ-orbit. By definition, this means that there is an integer ℓ such that $\gamma^{\ell}(s_1) = s_2$. Then $\alpha(s_2) = \alpha\gamma^{\ell}(s_1) = \alpha\gamma^{\ell}\alpha^{-1}(\alpha(s_1)) = \gamma^{-\ell}(\alpha(s_1))$.) Thus, whenever s_1 and s_2 in S are from a common γ-orbit, $\alpha(s_1)$ and $\alpha(s_2)$ are also in a common γ-orbit. So the image under α of any γ-orbit is again a γ-orbit.) In particular, if T is the orbit under γ which contains the fixed point a of α, then $\alpha(T)$ is an orbit which meets the orbit T in this fixed point, hence we must have $\alpha(T) = T$. Since the restriction of α to T is an involution of T with a unique fixed point, it follows as before that T is an odd set. Since both α and γ fix T, so does $\beta = \alpha\gamma$. Thus the restriction to T of β is an involution of the odd set T, and hence β must have a fixed point b in T. So there is an ℓ, $0 \leq \ell \leq k-1$, such that $b = \gamma^{\ell}(a)$ is fixed by β. To prove the uniqueness of this fixed point, it suffices to show that $k = 2\ell + 1$ is forced on us.

For $m \in Z$, we have $\beta(\gamma^m(b)) = \beta\gamma^m\beta^{-1}(\beta(b)) = \gamma^{-m}(b)$. Substituting $\gamma^{\ell}(a)$ for b, we find that the orbit T has a two-fold symmetry around its ℓth term:

$$\gamma^{\ell+m}(a) = \beta(\gamma^{\ell-m}(a)) \ \forall m \in Z.$$

In particular, taking $m = \ell + 1$ in this identity, we get $\gamma^{2\ell+1}(a) = \beta\gamma^{-1}(a) = \beta^2\alpha(a) = \alpha(a) = a$. From the definition of k, one sees that an integer h satisfies $\gamma^h(a) = a$ iff h is an integral multiple of k. Since $h = 2\ell + 1$ satisfies this condition, $2\ell + 1$ is a multiple of k. Since $1 \leq 2\ell + 1 < 2k$, this forces $2\ell + 1 = k$. Finally, substituting $\ell = \frac{k-1}{2}$, $m = \frac{k-1}{2} - n$ in the displayed identity, we get the last assertion of the theorem.

Shirali's Conjecture Vindicated

Define the involutions $\hat{\alpha}$ and $\hat{\beta}$ of the finite set $\widehat{S}$ as follows:

$$\hat{\alpha}(x, y, z) = (2z - x, x + y - z, z),$$
$$\hat{\beta}(x, y, z) = \begin{cases} (-x, y, z) & \text{if } x + z < y, \\ (x, z, y) & \text{if } x + z > y. \end{cases}$$

One readily verifies that (i) these are indeed involutions of $\widehat{S}$, (ii) $\hat{\alpha}$ has a unique fixed point, namely $\hat{a} := (1, \frac{p-1}{4}, 1)$, and $(x, y) \mapsto (x, y, y)$ is a bijection from the representations of p onto the fixed points of $\hat{\beta}$. Thus, in Zagier's proof, one may replace α, β and S by $\hat{\alpha}, \hat{\beta}$ and $\widehat{S}$, respectively. Finally, Shirali's function $\hat{\gamma}$ is related to these involutions by $\hat{\gamma} = \hat{\alpha}\hat{\beta}$. Therefore, the indicated substitutions in the proof of the above theorem yields a 'hatted' version of the theorem. In particular, this proves Shirali's conjecture.

Uniqueness of the Square Plus Square Representation of *p*

Aside from being non-constructive, Zagier's proof has another shortcoming. As already mentioned, the prime p has a unique representation as a sum of two squares. Or, what amounts to the same thing, β also has a unique fixed point in S. But this does not emerge from Zagier's proof (or from its constructive variations given above). We are unable to remedy this defect. Notice, however, that in view of the uniqueness assertion in the above theorem, it would suffice to show that γ acts transitively on S. (For, this would mean that $T = S$, and we know that β has a unique fixed point in T.) Computations by hand show that this is indeed correct for primes below hundred. One might therefore be tempted to conjecture that, generally, γ acts transitively on S. If correct, this would provide a neat explanation for the uniqueness of the β-fixed point. Unfortunately, this conjecture is incorrect. Its validity for small primes turns out to be yet another instance of the 'strong law of small numbers'. (If you have never heard of this law then you are urged to take a look at the beautiful article by Guy[1].)

We see this as follows.

For each fixed x, the number of points in S with the given first coordinate equals $d(\frac{p-x^2}{4})$. Therefore we have the formula

$$\#(S) = \sum_x d\left(\frac{p-x^2}{4}\right),$$

where the sum is over all odd numbers x in the range $1 \leq x < \sqrt{p}$. (Here $d(\cdot)$ is the usual divisor function: for $n \in I\!N, d(n)$ is the number of divisors of n including 1 and n.)

Let p be of the form $k^2 + 4$ (for an odd number k). Then, in the iterates under γ of the point $a = (1, 1, \frac{p-1}{4})$, the first coordinate increases in steps of two until the point $b = (k, 1, 1)$ is reached, then it decreases in steps of two until we reach the end point $(1, \frac{p-1}{4}, 1)$ of the orbit. This shows that in this case, the size k of the orbit T is related to the prime p by $p = k^2 + 4$. Also, the sum in the formula for $\#(S)$ given above has $(k+1)/2$ terms of which one term equals 1 while the remaining $(k-1)/2$ terms are ≥ 2. Since $d(n) = 2$ iff n is a prime, it follows that *for a prime of the form $p = k^2 + 4$, γ is transitive on S (i.e., $k = \#(S)$) iff $(p - x^2)/4$ is a prime for all odd numbers x in the range $1 \leq x < k$.* This shows, for instance, that we do not have transitivity for $p = 229$.

Inefficiency of the Algorithm

Clearly, the α-β algorithm needs at most $\frac{1}{2}\#(S)$ steps. Since $d(n) = O(n^\epsilon)$ and the formula for $\#(S)$ has $O(p^{\frac{1}{2}})$ terms in it, the number of necessary iterations is $O(p^{\frac{1}{2}+\epsilon})$. The example of primes of the form square plus four (presumably there are infinitely many such primes) shows that this estimate is close to the best possible. Wagon describes known algorithms whose complexity is polynomial in $\log p$, and the α-β algorithm compares very unfavourably (see Suggested Reading). But it may be that we have looked at the worst case, and for some large class of primes its performance is much better. Moreover, it may be possible to significantly improve on the performance of the algorithm as follows. The set S can be partitioned into three parts on each of which γ is linear (the permutation $\hat{\gamma}$ is even better in this respect; we have a partition of $\widehat{S}$ into two parts on each of which $\hat{\gamma}$ is linear). The runs of iteration during which the iterates stay in the same piece of S may easily be combined into a single step.

A Combinatorial Lemma

The perceptive reader may have suspected by now that the theorem presented above does not have much to do with primes or their representations by squares. This is indeed correct, and the theorem is a manifestation of a combinatorial phenomenon. We have:

LEMMA. *For any two involutions α and β of a finite set S, there are only three possibilities for any $\alpha\beta$-orbit: (i) neither involution has a fixed point in the orbit, or (ii) each of them has a unique fixed point in the orbit, or (iii) one of them has two fixed points in the orbit while the other has none.*

At first glance, this statement may look very strange. (For readers with a reasonable amount of familiarity with groups and group actions, here is a hint for a group-theoretic proof of this lemma: think of the group of isometries of a regular polygon.) But here is an elementary ('graph-theoretic') proof.

Let $\gamma = \alpha\beta$. Fix a γ-orbit T. If neither α nor β has a fixed point in T, then there is nothing to prove: we are in case (i) of the lemma. So assume that one of these two involutions has at least one fixed point. Then, arguing as in the proof of the above theorem, one sees that T is fixed by both α and β. Thus T is a union of α-orbits as well as of β-orbits. If T is a singleton, then we are in case (ii) and again there is nothing to prove. So we may assume that T has at least two elements. Hence no element of T is fixed by γ.

Now consider the graph G defined as follows. The vertices of G are the elements of T. Two distinct elements x, y of T are joined by an edge in G if (and only if) $y = \alpha(x)$ or $y = \beta(x)$ (i.e., if $\{x, y\}$ is an orbit of one of the involutions). Clearly, this is an undirected graph. Note that, for each x in $T, \alpha(x)$ and $\beta(x)$ are distinct elements of T—or else x would be fixed by γ, contrary to our assumption. It follows that each vertex x is of degree 1 or 2 in G (i.e., x is joined to one or two vertices), according to whether x is or is not fixed by one (and only one) of the two involutions. Since we have assumed that at least one of them has a fixed point in T, it follows that G has at least one vertex of degree one. Also, since $\gamma = \alpha\beta$ is transitive on T (T is a γ-orbit!), it follows that G is connected. Now, here is the punch line: the only connected graphs with all vertices of degree ≤ 2 and at least one vertex of degree 1 are the paths. Hence G is a path. So G has exactly two vertices of degree 1 (the two ends of the path) and hence we are in case (ii) or (iii). This proves the lemma.

EXERCISE: Continue this argument to see that if the elements of T are arranged on a circle according to the action of γ, then the two ends of G are placed opposite to each other. This explains the symmetry observed in the theorem.

A Prime Testing Algorithm?

If $n \equiv 1 \pmod 4$ is a number (not necessarily a prime) which is not a perfect square, then S, α, β may be defined as before with n replacing p. What happens if one runs the α-β algorithm in this case? Our combinatorial lemma shows that if we look inside the orbit T containing the fixed point $(1, 1, \frac{n-1}{4})$ of α, either we may find a fixed point of β and hence a representation of n as a sum of two squares, or we find a second fixed point (x, x, z) of α and hence a nontrivial factorisation $n = x(x + 4z)$ of n. The second case is bound to occur if the square free part of n has a 3 (mod 4) factor (since in this case n has no representation as a sum of two squares). In the former case, of course, we are unable to decide whether n is a prime or not (for instance, this case occurs if n is a number of the form $k^2 + 4$, even when n is composite). If, however, we happen to know a two squares representation of n and the algorithm is lucky enough to produce a second representation, then we can still conclude that n is composite (because a prime has at most one such representation). Perhaps it will be interesting to characterise those numbers n for which the first case occurs.

Suggested Reading

[1] R K Guy. The strong law of small numbers. *Amer. Math. Monthly*. Vol. 95, No. 8, pp 697–711, 1988.

[2] I Niven and H S Zuckerman. *An Introduction to the Theory of Numbers*. Third edition. Wiley, 1972.

[3] S A Shirali. On Fermat's two squares theorem. *Resonance*. Vol. 2, No. 3, pp 69–73, 1997.

[4] S Wagon. The Euclidean algorithm strikes again. *Amer. Math. Monthly*. Vol. 97, No. 2, pp 125–126, 1990.

[5] D Zagier. A one-sentence proof that every prime $p \equiv 1 \pmod 4$ is a sum of two squares. *Amer. Math. Monthly*. Vol. 97, No. 2, p 144, 1990.

B Bagchi
Statistics and Mathematics Unit
Indian Statistical Institute
Bangalore 560 059

8

Factoring Fermat Numbers

A Unique Computational Experiment for Factoring F_9

C E Veni Madhavan

Fermat observed that the numbers $F_k = 2^{2^k} + 1$, $k = 0, 1, 2, 3, 4$ are prime, and wondered whether this was true for all k. Euler found that the very next Fermat number is composite: $F_5 = 2^{32} + 1 = 641 \times 6700417$. So far it has been verified that $F_k, 5 \leq k \leq 22$ are all *composite*. No one knows whether any other F_k is prime. The numbers F_k grow rapidly with k—each is almost a square of the previous number—and it is a very difficult task to decide their primality. We give below an outline of the relevant computational challenges.

First note that, if k is odd, 3 divides $2^k + 1$ and in general, $2^a + 1$ divides $2^{ak} + 1$. Thus, if k is not a power of two, $2^k + 1$ is not prime. Fermat hazarded a guess that the converse was also true. In 1877, François Pépin published a necessary and sufficient condition which states that $F_k, k > 1$ is prime if and only if F_k divides $5^{(F_k-1)/2} + 1$. This condition is the basis for determining whether F_k is prime for any given k. Failure of this condition means that F_k is composite. It does not reveal any information about the factors.

Today, sophisticated number theoretic methods and powerful computing platforms are used for testing primality and factoring of large integers. These find applications in many practical problems such as cryptography. The recent records in Fermat number factoring have been achieved by means of two techniques called *number field sieve* (NFS) and *elliptic curve method* (ECM).

The complete factoring of F_9, which has about 150 decimal digits was carried out in 1992 by a unique computational experiment. Hundreds of computers in different parts of the world, working independently and in their spare time generated certain seed numbers. These computers sent their seeds by electronic mail to a host computer in USA. The host carried out the combination of the seeds and the factoring. The NFS method, requiring the generation of an enormous number of such seeds, was thus eminently suitable for this exercise. However, this method is quite difficult to implement.

Last year the number F_{22} was determined to be composite, using Pépin's criterion and extremely fast arithmetical algorithms implemented on supercomputers. This number of about 1.3 million decimal digits (about 500 times as long as this chapter) required about 10^{16} arithmetical operations and about seven months of real time. Complete factorization of Fermat numbers is known only for $k \leq 9$ and $k = 11$. No prime factors of F_{14} and F_{20} are known.

C E Veni Madhavan
Department of Computer
Science and Automation
Indian Institute of Science
Bangalore 560 012

9

The Class Number Problem

Binary Quadratic Forms

Rajat Tandon

Introducing the reader to the notion of 'class numbers', this chapter defines class numbers the way they arose in the study of 'binary quadratic equations'.

Remember the formula $\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ that we all learnt in school. Indians have a long history of work on quadratics. The high point seems to have been when Brahmagupta in the early seventh century gave a method by which, knowing one solution (x, y) in integers of the equation $cX^2 + 1 = Y^2$ (Pell's Equation!), where c is a constant integer, he could generate an infinite family of solutions. But I am interested here in the above formula. I will always assume that a, b, c are integers. The quantity $b^2 - 4ac$ under the square root sign gives us information about the quadratic $aX^2 + bX + c$. For instance, it tells us whether the quadratic has any real roots—it must be positive for this to be so. It tells us whether the quadratic has any rational roots—it must be a perfect square for this to be so. We call it the discriminant of the quadratic $aX^2 + bX + c$. The class number problem is concerned with the following questions:

1) Given an integer Δ, are there any quadratics $F(X) = aX^2 + bX + c, (a, b, c$ integers) whose discriminant $b^2 - 4ac$ equals Δ?

2) If so, how many such quadratics exist? Can we classify them in any way?

It is obvious that if $\Delta = b^2 - 4ac$, then 4 divides Δ or 4 divides $\Delta - 1$, i.e., $\Delta \equiv 0$ or $1 (\text{mod } 4)$. This is a necessary condition for there to be an integral quadratic with discriminant Δ. It is a simple exercise to show that it is also sufficient. So we have a complete answer to the first question. The second question is considerably more complex.

Before proceeding further let me give a quick recap of the notion of an *equivalence relation*. A relation '$\sim$' on a set S is called an equivalence relation if it is reflexive $(x \sim x)$, symmetric $(x \sim y \Rightarrow y \sim x)$ and transitive $(x \sim y$ and $y \sim z \Rightarrow x \sim z)$. Let $[x]$ denote the subset of S consisting of elements equivalent to x. It is called an equivalence class; note that any two equivalence classes are either identical or disjoint. Then S is the disjoint union of distinct equivalence classes. We denote the set of equivalence classes by $S/\sim$.

Suppose we replace X by $X+1$ in the quadratic aX^2+bX+c. We have $a(X+1)^2 + b(X+1)+c = aX^2+(b+2a)X+(a+b+c)$. The discriminant of this is $(b+2a)^2 - 4a(a+b+c) = b^2-4ac$. So the discriminant does not change if we replace $F(X) = aX^2+bX+c$ by $F(X+1)$ and hence by $F(X+2), F(X+3), \ldots$. Similarly, for $F(X-1), F(X-2)$ etc. Notice also that $\Delta = b^2-4ac$ is symmetric in a and c, i.e., if we replace aX^2+bX+c by cX^2+bX+a then the discriminant does not change. This indicates that it might be better to replace $F(X) = aX^2+bX+c$ by the corresponding homogeneous polynomial in 2 variables $F(X,Y) = aX^2+bXY+cY^2$. Then instead of the transformation $X \to X+1$ we take the transformation $X \to X+Y, Y \to Y$.

Let $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $W = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. T and W are members of $SL(2,\mathbf{Z})$, the group of 2×2 matrices with integer coefficients and determinant 1. If $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2,\mathbf{Z})$ and F is a homogeneous quadratic polynomial in two variables, we denote by $A \cdot F$ the polynomial obtained by replacing X by $\alpha X + \beta Y$ and Y by $\gamma X + \delta Y$. Observe that if $A, B \in SL(2,\mathbf{Z})$ then $A \cdot (B \cdot F) = AB \cdot F$. It is easy to check that if F has discriminant Δ, then so does $A \cdot F$ for any $A \in SL(2,\mathbf{Z})$. Denote by $S(\Delta)$ the set of all integral homogeneous quadratic polynomials in two variables of discriminant Δ.

We define an equivalence relation on $S(\Delta)$ by $F \sim G$ if either $F = G$ or there exists a chain $F_1, F_2, \ldots, F_n$ in $S(\Delta)$ such that $F = F_1$, $G = F_n$ and each F_{i+1} is either $T \cdot F_i$ or $T^{-1} \cdot F_i$ or $W \cdot F_i$; such a chain is called a chain from F to G. It is easy to see that this gives an equivalence relation on $S(\Delta)$. Hence $S(\Delta)$ can be partitioned into equivalence classes. We remark that it can be shown that $SL(2,\mathbf{Z})$ is generated by T and W and hence two forms F and G are equivalent if and only if there exists an $A \in SL(2,\mathbf{Z})$ such that $A \cdot F = G$.

Assume from now on that $\Delta < 0$. This is not because the case $\Delta > 0$ is uninteresting but because it is more difficult and less is known in this case. If the discriminant of $F(X,Y) = aX^2+bXY+cY^2$ is Δ then it is also so for $-F$. Note that $\Delta < 0$ implies that a and c have the same sign. We define $S_1(\Delta)$ to be the subset of $S(\Delta)$ consisting of those forms F for which a and c are positive, and $S_2(\Delta)$ its complement. Then $F \mapsto -F$ is a bijection from $S_1(\Delta)$ to $S_2(\Delta)$. It is also easy to see that no member of $S_1(\Delta)$ can be equivalent to any member of $S_2(\Delta)$. We restrict ourselves to $S_1(\Delta)$.

Definition. The form $F(X,Y) = aX^2+bXY+cY^2$ of $S_1(\Delta)$ is said to be *almost reduced* if $|b| \leq a \leq c$.

Theorem. Each equivalence class in $S_1(\Delta)$ has at least one almost reduced form.

Proof. Consider an equivalence class with an element $F(X,Y) = aX^2+bXY+cY^2$ in it. If $a > c$ replace F by $W \cdot F = F_1$ (say). Then $F_1(X,Y) = a_1X^2 + b_1XY + c_1Y^2$ with $a_1 = c$ and $c_1 = a$, and so $a_1 \leq c_1$. Notice $a > a_1$. If now

$|b_1| \leq a_1$, F_1 is reduced. If not, find an integer n such that $|b_1 + 2a_1n| \leq a_1$. Replace F_1 by $F_2 = T^n \cdot F_1$. Then

$$\begin{aligned} F_2(X,Y) &= a_1(X+nY)^2 + b_1(X+nY)Y + c_1Y^2 \\ &= a_1X^2 + (b_1 + 2a_1n)XY + (a_1n^2 + b_1n + c_1)Y^2 \\ &= a_2X^2 + b_2XY + c_2Y^2 \end{aligned}$$

(say), with $|b_2| \leq a_2$ and $a_2 = a_1$. But now a_2 may not be less than or equal to c_2. If so, again apply W and continue as before. After a finite number of steps we get an almost reduced form (finite because $a \geq a_1 \geq a_2 \geq \cdots > 0$).

Corollary. The number of equivalence classes in $S_1(\Delta)$ is finite.

Proof. It suffices to show that the number of almost reduced forms is finite. If $aX^2 + bXY + cY^2$ is almost reduced of discriminant Δ then

$$a \leq c \Rightarrow 4a^2 \leq 4ac = b^2 - \Delta \leq a^2 - \Delta.$$

Hence $3a^2 \leq |\Delta|$. Since a is a positive integer, there are only finitely many possible values of a and hence of b. Once a and b are given, c is uniquely determined.

A natural question to ask is: is there precisely one almost reduced form in each equivalence class? The answer is—almost but not quite.

We know that $X^2 + \frac{b}{a}X + \frac{c}{a}$ has two non-real roots (because $\Delta < 0$), say τ and $\bar{\tau}$. One of these (say τ) will lie in the upper half plane $\{x + iy : x, y \in \mathrm{S}, y > 0\}$. Hence $F(X,Y) = aX^2 + bXY + cY^2 = a(X + \tau Y)(X + \bar{\tau}Y)$ with $b = a(\tau + \bar{\tau})$ and $c = a\tau\bar{\tau}$. Hence to say that F is almost reduced is equivalent to saying that $|\tau + \bar{\tau}| \leq 1$ and $\tau\bar{\tau} \geq 1$, i.e., $\tau \in \mathrm{S}$ where S is the region shown in Figure 9.1 (including the boundary).

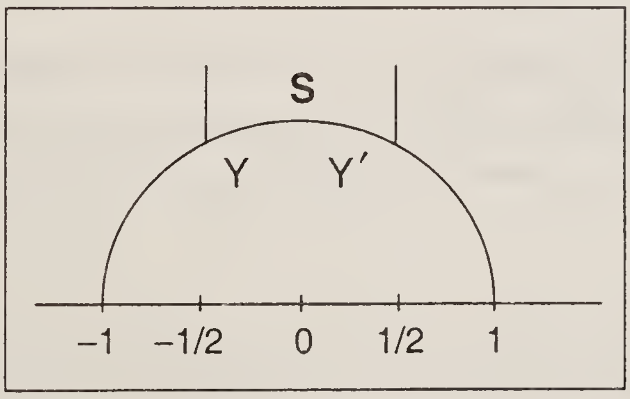


Figure 9.1

Notice that if τ is on the left vertical boundary of S then $\tau + 1$ is on the right vertical boundary which is also in S. Similarly, if τ is on the curve at Y then $\frac{-1}{\tau}$ is at Y'. In view of this we make the following definition:

Definition. We say that $F(X,Y) = aX^2 + bXY + cY^2$ is *reduced* if the corresponding $\tau \in \mathrm{S}$ but $\tau \notin$ the left boundary of S, i.e., the left vertical boundary and curve Y. This is equivalent to saying that $|b| \leq a \leq c$, and in case $a = |b|$ then $b > 0$, and in case $a = c$ then $b \geq 0$. We now have the expected theorem.

Δ	possible a	possible b, c	reduced forms	$\underline{h}(\Delta)$
-3	$a = 1$	$b = 1, c = 1$	$X^2 + XY + Y^2$	1
-4	$a = 1$	$b = 0, c = 1$	$X^2 + Y^2$	1
-7	$a = 1$	$b = 1, c = 2$	$X^2 + XY + 2Y^2$	1
-8	$a = 1$	$b = 0, c = 2$	$X^2 + 2Y^2$	1
-11	$a = 1$	$b = 1, c = 3$	$X^2 + XY + 3Y^2$	1
-12	$a = 1$ or 2	$b = 0, c = 3$ if $a = 1$	$X^2 + 3Y^2$	2
		$b = 2, c = 2$ if $a = 2$	$2(X^2 + XY + Y^2)$	
-15	$a = 1$ or 2	$b = 1, c = 4$ if $a = 1$	$X^2 + XY + 4Y^2$	2
		$b = 1, c = 2$ if $a = 2$	$2X^2 + XY + 2Y^2$	
-16	$a = 1$ or 2	$b = 0, c = 4$ if $a = 1$	$X^2 + 4Y^2$	2
		$b = 0, c = 2$ if $a = 2$	$2(X^2 + Y^2)$	
-19	$a = 1$ or 2	$b = 1, c = 5$ if $a = 1$	$X^2 + XY + 5Y^2$	1
-20	$a = 1$ or 2	$b = 0, c = 5$ if $a = 1$	$X^2 + 5Y^2$	2
		$b = 2, c = 3$ if $a = 2$	$2X^2 + 2XY + 3Y^2$	
-23	$a = 1$ or 2	$b = 1, c = 6$ if $a = 1$	$X^2 + XY + 6Y^2$	3
		$b = 1, c = 3$ if $a = 2$	$2X^2 + XY + 3Y^2$	
		$b = -1, c = 3$ if $a = 2$	$2X^2 - XY + 3Y^2$	

THEOREM. In each equivalence class of $S_1(\Delta)$ there is precisely one reduced form. The chart gives a list of reduced forms for low values of $|\Delta|$ is shown; $\underline{h}(\Delta)$ is the number of reduced forms.

We notice that some forms in the list are constant multiples of forms which came earlier in the list.

DEFINITION. A form $aX^2 + bXY + cY^2$ is said to be *primitive* if $(a, b, c) = 1$.

We let $h(\Delta)$ be the number of primitive reduced forms of discriminant Δ in $S_1(\Delta)$. $h(\Delta)$ is known as the *class number* of the forms with discriminant Δ. Notice that $h(\Delta)$ is 1 for $\Delta = -3, -4, -7, -8, -11, -12, -16, -19$ in the list.

DEFINITION. An integer $\Delta \equiv 0$ or $1 \pmod 4$ is said to be a *fundamental discriminant* if it is not of the form $\Delta_0 n^2$ where Δ_0 is a discriminant and n an integer greater than 1.

For instance, -12 and -16 are not fundamental discriminants. Notice that if Δ is fundamental, then a form of discriminant Δ is always primitive. Notice also that if Δ is fundamental, then it cannot have an odd square factor. We will see later that if Δ is fundamental then it has another interpretation.

In 1934, Heilbronn showed that $h(\Delta) \to \infty$ as $\Delta \to -\infty$ from which it follows (how?) that given any natural number N there are only a finite number of negative fundamental discriminants Δ for which the class number, $h(\Delta) = N$. One of the questions that suggests itself from the above is: what are the negative fundamental Δ for which $h(\Delta)$ is 1? Above we have given six such Δ's. Here are three more: $\Delta = -43, -67, -163$. In 1800, Gauss conjectured that there were no more.

In 1936, Siegel showed that for every $\epsilon > 0$ there exists a positive constant C_ϵ such that $h(\Delta) \geq C_\epsilon |\Delta|^{\frac{1}{2}-\epsilon}$. However, the result showed the existence of C_ϵ but not how to compute it. His proof showed that there cannot be two 'large' values of $|\Delta|$'s for which $h(\Delta)$ is small. From this it was proved that there is possibly just one other Δ (call it Δ_{10}) for which $h(\Delta) = 1$ and this Δ must be very large indeed. In 1966, Harold Stark, in his thesis, showed that Δ_{10} does not exist[1]. The same methods were applied to the negative Δ for which $h(\Delta) = 2$ and it was found that there are 18 such Δ's, the largest value of $|\Delta|$ being 427 (Baker, Stark, Montgomery etc). In 1986, using powerful methods in algebraic geometry, D Goldfeld, B H Gross and D Zagier solved the problem of fundamental negative Δ with $h(\Delta) = 3$.

REMARK. The Δ for which $h(\Delta) = 1$ have remarkable properties. For instance, if p is a positive prime number which is congruent to 3(mod 4) and $h(-p) = 1$ then $x^2 + x + \frac{p+1}{4}$ is a prime number for all x such that $0 \leq x \leq \frac{p-7}{4}$.

Suggested Reading

[1] H M Stark. The complete determination of the complex quadratic fields of Class number one. *Michigan Math F.* 14. 1–27, 1967.

[2] J P Serre. *A course in Arithmetic.* Narosa Publishing House. New Delhi, 1979.

[3] D Flath. *Introduction to Number Theory.* John Wiley and Sons. New York, 1989.

RAJAT TANDON
Department of Mathematics
University of Hyderabad
Central University P.O.
Hyderabad 500 046

[1] In 1954, an amateur mathematician Heegner, in Germany, had proved the same result but his proof had some gaps which were responsible for mathematicians expressing reservations about the proof. But later it was shown by Stark that the arguments of Heegner can be made rigorous and he managed to make Heegner's proof work. In fact, Heegner's ideas, in particular his construction of what are now called *Heegner points*, have proved to be very fruitful in later work on elliptic curves.

10

The Class Number Problem

An Introduction to Algebraic Number Theory

Rajat Tandon

This chapter gives an introduction to 'algebraic number theory', defines class numbers for finite extensions of the field of rational numbers and proves that in the context of quadratic fields, this definition coincides with the definition of class numbers via binary quadratic forms given in the previous chapter.

We have seen in the previous chapter that some seemingly innocuous questions starting with the formula $\frac{-b \pm \sqrt{b^2-4ac}}{2a}$ lead to fairly deep mathematics. This is typical of the subject. It is so important to ask the right question—" ask an impertinent question and you get a pertinent answer ".

The roots of $aX^2 + bX + c = 0$ are given by $\frac{-b \pm \sqrt{\Delta}}{2a}$ where $\Delta = (b^2 - 4ac)$, i.e., they are of the form $x + y\sqrt{\Delta}$ with x and y rational. The set $\mathbf{Q}(\sqrt{\Delta})$ of elements of the form $x + y\sqrt{\Delta}$ with x and y rational, forms a subfield of the field of complex numbers, $\mathbf{C} \cdot \mathbf{Q}(\sqrt{\Delta})$ is also a vector space over the rationals if we define scalar multiplication by $\lambda(x + y\sqrt{\Delta}) = \lambda x + \lambda y\sqrt{\Delta}$. $\{1, \sqrt{\Delta}\}$ is a basis of $\mathbf{Q}(\sqrt{\Delta})$ over $\mathbf{Q}$, and $\mathbf{Q}$ is a subfield of $\mathbf{Q}(\sqrt{\Delta})$. This process can easily be generalised. For instance, let p be a prime and $\zeta = e^{2\pi i/p}$. Let $\mathbf{Q}(\zeta)$ be the set of complex numbers of the form $x_0 + x_1\zeta + x_2\zeta^2 + \cdots + x_{p-2}\zeta^{p-2}$ with x_i rational. Note that $1 + \zeta + \zeta^2 + \zeta^3 + \cdots + \zeta^{p-1} = 0$ so ζ^{p-1} can be written in terms of $1, \zeta, \zeta^2, \zeta^3, \ldots, \zeta^{p-2}$. Check that $\mathbf{Q}(\zeta)$ is a subfield of $\mathbf{C}$ containing $\mathbf{Q}$ and that $1, \zeta, \zeta^2, \ldots, \zeta^{p-2}$ is a basis of $\mathbf{Q}(\zeta)$ over $\mathbf{Q}$ with scalar multiplication being defined in the obvious way. These are examples of fields containing $\mathbf{Q}$ which are finite dimensional as vector spaces over $\mathbf{Q}$. Such fields are known as *algebraic number fields* and were the object of detailed study by Dedekind, Kronecker and Kummer in the 19th century. Amongst the several motivations for studying such fields were three problems suggested by Greek geometers:

(i) To trisect any given angle.

(ii) To construct a cube whose volume is twice that of a given cube.

(iii) To construct a square equal in area to a given circle.

These constructions were to be done by 'ruler and compass only' in the manner that we are taught at school. The second problem boils down to being able to construct by ruler and compass the real root of $X^3 - 2$. Galois and Abel looked at such problems and their work gave a huge impetus to the systematisation of algebra and algebraic number theory.

The examples given above, $\mathbf{Q}(\sqrt{\Delta})$ and $\mathbf{Q}(\zeta)$, have been generated by single elements ($\sqrt{\Delta}$ and ζ) which satisfy some polynomial with rational (in fact, integral) coefficients ($X^2 - \Delta, X^p - 1$ respectively). Indeed, it can be shown that any subfield of $\mathbf{C}$ containing $\mathbf{Q}$ which is n-dimensional as a vector space over $\mathbf{Q}$ consists of elements of the form $x_0 + x_1\alpha + x_2\alpha^2 + \cdots + x_{n-1}\alpha^{n-1}$ where the x_i are rationals and α is a complex number which satisfies a polynomial equation of degree n with rational coefficients.

The first thing we would want to know about such fields is whether they have a subring in them in much the same way that $\mathbf{Q}$ contains $\mathbf{Z}$ and every element of $\mathbf{Q}$ is a ratio of two (one non-zero) elements of $\mathbf{Z}$. One 'natural' possibility in $\mathbf{Q}(\sqrt{\Delta})$ could be $\mathbf{Z} + \mathbf{Z}\sqrt{\Delta}$, i.e., elements of the form $a + b\sqrt{\Delta}$ with a and b integers or in other words $\mathbf{Z}$-linear combinations of the basis $1, \sqrt{\Delta}$. Similarly one could consider $\mathbf{Z} + \mathbf{Z}\zeta + \mathbf{Z}\zeta^2 + \mathbf{Z}\zeta^3 + \cdots + \mathbf{Z}\zeta^{p-2}$ in $\mathbf{Q}(\zeta)$. But immediately one would recognise a difficulty in basing a definition which depends on the choice of a basis. For instance, $\mathbf{Q}(\sqrt{\Delta}) = \mathbf{Q}(\sqrt{4\Delta})$ but $\mathbf{Z} + \mathbf{Z}\sqrt{\Delta} \neq \mathbf{Z} + \mathbf{Z}\sqrt{4\Delta}$ or observe that if $p = 3$ then $\zeta = e^{2\pi i/3} = \frac{-1+\sqrt{-3}}{2}$, so $\mathbf{Q}(\zeta) = \mathbf{Q}(\sqrt{-3})$ but $\mathbf{Z} + \mathbf{Z}\zeta \neq \mathbf{Z} + \mathbf{Z}\sqrt{-3}$. To get around the problem of square factors of Δ, we will henceforth assume that Δ is a fundamental discriminant. See previous chapter. Hence the only square factor Δ can have is 4.

We have already seen that the fields above are generated by elements which satisfy a monic (leading coefficient 1) polynomial with rational coefficients. In fact, every element $a + b\sqrt{\Delta}$ in $\mathbf{Q}(\sqrt{\Delta})$ satisfies the polynomial $X^2 - 2aX + (a^2 - b^2\Delta) = 0$. This suggests an alternative. Why not consider those elements of $\mathbf{Q}(\sqrt{\Delta})$ (or $\mathbf{Q}(\zeta)$) which satisfy a monic polynomial with coefficients in $\mathbf{Z}$? Such elements are called *algebraic integers* (in the given field). Do such elements form a subring I, i.e., are they closed under addition and multiplication? The answer is 'yes'. Observe that $a + b\sqrt{\Delta}$ will be an element of the given type provided $2a \in \mathbf{Z}$ and $a^2 - b^2\Delta \in \mathbf{Z}$. Suppose then that $a + b\sqrt{\Delta}$ and $c + d\sqrt{\Delta}$ are such that $2a, 2c \in \mathbf{Z}$ and $a^2 - b^2\Delta, c^2 - d^2\Delta \in \mathbf{Z}$. Observe that $2(a + c) \in \mathbf{Z}$ and $(a + c)^2 - (b + d)^2\Delta = (a^2 - b^2\Delta) + (c^2 - d^2\Delta) + 2ac - 2bd\Delta$. We say that a rational number is a half integer if it is of the form $l/2$, where l is odd. We make the following observations which can easily be proved by the reader: for $a, b \in \mathbf{Q}$, $2a$ and $a^2 - b^2\Delta$ are integers implies

(i) $2b \in \mathbf{Z}$ since Δ has no square free factor other than possibly 4;

(ii) if Δ is even, then a must be an integer and b either an integer or half integer;

(iii) if Δ is odd, a and b must be either both integers or both half integers.

In all cases it can then be seen that if $2a, 2c \in \mathbf{Z}$ and $a^2 - b^2\Delta, c^2 - d^2\Delta \in \mathbf{Z}$ then $2ac - 2bd\Delta \in \mathbf{Z}$ and therefore that $(a + c)^2 - (b + d)^2\Delta \in \mathbf{Z}$. On the other hand,

$$(a + b\sqrt{\Delta}) \cdot (c + d\sqrt{\Delta}) = ac + bd\Delta + (ad + bc)\sqrt{\Delta} \text{ and,}$$
$$(ac + bd\Delta)^2 - (ad + bc)^2\Delta = (a^2 - b^2\Delta) \cdot (c^2 - d^2\Delta)$$

are both in $\mathbf{Z}$. Hence I is indeed closed under addition and multiplication.

EXERCISE. Show that

(i) in $\mathbf{Q}(\sqrt{-1})$ we have $I = \{a + b\sqrt{-1} \mid a, b \in \mathbf{Z}\}$

(ii) in $\mathbf{Q}(\sqrt{-3})$, $I = \{\frac{a+b\sqrt{-3}}{2} \mid a, b \in \mathbf{Z}, a \equiv b \pmod 2\} = \mathbf{Z} + \mathbf{Z}\zeta$, where $\zeta = \frac{-1+\sqrt{-3}}{2}$ is a cube root of unity.

Would every element of $\mathbf{Q}(\sqrt{\Delta})$ be a ratio of two elements of I? We note that $\mathbf{Z} + \mathbf{Z}\sqrt{\Delta} \subseteq I$ and $\frac{a}{b} + \frac{c}{d}\sqrt{\Delta} = \frac{ad+bc\sqrt{\Delta}}{bd}$, so this is trivially true. What other properties of $\mathbf{Z}$ would we like I to have? The best would be unique factorisation. In $\mathbf{Z}$ we have the notion of a prime number and we know that every number can be written upto sign uniquely as a product of distinct prime powers, viz,

$$n = \pm p_1^{e_1} p_2^{e_2} \ldots p_r^{e_r}$$

where the p_i are distinct primes and, moreover, if n is also equal to $\pm q_1^{f_1} q_2^{f_2} \ldots q_s^{f_s}$, then after changing the order of the q_i's, if necessary, we have $r = s$, $p_i = q_i$ and $e_i = f_i$ for all i.

Imagine the usefulness of having such a property in I. For instance, consider $\mathbf{Q}(\zeta)$ as above and the ring of integers I in $\mathbf{Q}(\zeta)$, i.e., the set of all elements in $\mathbf{Q}(\zeta)$ which satisfy a monic polynomial in $\mathbf{Z}[X]$, the ring of polynomials in one variable with integer coefficients. Suppose there exist non-zero integers x, y, z such that $x^p + y^p = z^p$. Then,

$$x^p = z^p - y^p = (z - y)(z - \zeta y)(z - \zeta^2 y) \ldots (z - \zeta^{p-1} y). \tag{1}$$

It is easy to see that $x \in I$ and $z - \zeta^i y \in I$. If we have unique factorisation in I, there is just a chance that (1) may give us a contradiction to unique factorisation (or allow us to use the method of descent) and we may prove Fermat's[1] last theorem! It is just possible that Fermat had some such proof in mind when he wrote in the margin ....

We would first need the notion of a prime element in I. This is accomplished more or less as in $\mathbf{Z}$—negatives allowed. So we consider $-2, -3, -5, \ldots$ also as primes.

DEFINITION 1. An integer n is a prime if whenever n is written as a product ab of two integers then either a or b must be ± 1. Note that ± 1 are the only units in $\mathbf{Z}$, i.e., elements in $\mathbf{Z}$ with a multiplicative inverse.

There is another way of defining a prime number.

DEFINITION 1′. An integer $p \neq \pm 1$ is a prime if and only if whenever p divides a product of integers ab then p must divide either a or b.

[1] Incidentally, this is what Gauss had to say about FLT. "I confess that Fermat's theorem as an isolated proposition has very little interest for me because I could easily lay down a multitude of such propositions which one could neither prove nor dispose off." Gauss said that FLT had induced him to recall some of his earlier ideas in higher arithmetic but that he was not in a position to go back to that work because of his circumstances. "Still I am convinced that if I am as lucky as I dare hope and if I succeed in taking some of the principal steps in that theory, then Fermat's theorem will appear as only one of the least interesting corollaries."

Recall that if n is an integer then $n\mathbf{Z}$, the set of multiples of n, forms an ideal in $\mathbf{Z}$ (an ideal J in a commutative ring R is an additive subgroup of R which has the property: $x \in J$, $r \in R$ implies $rx \in J$). If an ideal I in a ring satisfies the property: $ab \in I$ implies either $a \in I$ or $b \in I$ it is called a prime ideal. So saying that the integer p is a prime number is the same as saying that $p\mathbf{Z}$ is a prime ideal in $\mathbf{Z}$. It is easy to see that the two definitions we have given are equivalent in $\mathbf{Z}$.

Based on the above, we could define in an arbitrary commutative ring with unity R (all our rings will be so) an element π to be prime either by requiring that whenever $\pi = ab$, either a or b must be a unit in R, or by requiring that the ideal πR, consisting of all multiples of π, is a prime ideal. Unfortunately, in an arbitrary ring the two definitions are not equivalent. An element π which satisfies the first property is said to be *irreducible* whereas if πR is a prime ideal we call π a *prime*. In integral domains (commutative rings with no zero divisors) all primes are irreducible but not vice-versa. (Exercise: Prove this.)

A domain in which every non-zero non-unit can be written as a product of irreducibles in an essentially unique way, that is upto order and multiplication by units ($6 = 2 \cdot 3 = 3 \cdot 2 = (-2) \cdot (-3) = (-3) \cdot (-2)$) is called a *unique factorisation domain* (UFD). Clearly, $\mathbf{Z}$ is a UFD and it is easy to check that $J = \mathbf{Z} + \mathbf{Z}i$ is also a UFD.

$\mathbf{Z}$ has another property which is somewhat stronger—every ideal in $\mathbf{Z}$ is of the form $n\mathbf{Z}$ where n is an integer. A domain D which has the property that every ideal in it is of the form xD for some x in D is called a *principal ideal domain* (PID) and every PID is a UFD. If we could show that the ring of integers I in an algebraic number field is always a PID then we could use the argument given above for FLT. Unfortunately, I is not always a PID. For instance, consider $\mathbf{Q}(\sqrt{-20})$; then $I = \mathbf{Z} + \mathbf{Z}\sqrt{-5}$ and we have $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$. It is easy to check that 2, 3, $1 \pm \sqrt{-5}$ are all irreducible elements in I. We remark that the ring of integers of an algebraic number field is a UFD if and only if it is a PID.

Recall that if $\mathcal{A}$ and $\mathcal{B}$ are two ideals in a ring R then we define their product as $\mathcal{A} \cdot \mathcal{B} = \left\{ \sum_{i=1}^{i=n} a_i b_i | a_i \in \mathcal{A}, b_i \in \mathcal{B}, \text{ for some } n \right\}$. This is also an ideal. Though I is not always a PID it is true that every ideal in I can be written uniquely, except for order, as a product of prime ideals. This gives us the first hint that the concept of an ideal may be at least as important as the notion of an element. Note that in a PID the two notions are almost the same as every ideal is generated by a single element which is uniquely determined upto units.

So if I is not always a PID then how 'bad' is it? The set $\mathcal{I}$ of ideals in I under the product defined above form a semigroup (I itself is the identity). We define an equivalence relation on this set $\mathcal{I}$ as follows: $\mathcal{A} \sim \mathcal{B}$ if there exist $\alpha, \beta \in I$ such that $\alpha I \cdot \mathcal{A} = \beta I \cdot \mathcal{B}$. It is easy to check that this gives us an equivalence relation on $\mathcal{I}$ and the product on $\mathcal{I}$ induces a product on the set of equivalence classes $\mathcal{I}/\sim$: $[\mathcal{A}] \cdot [\mathcal{B}] = [\mathcal{A} \cdot \mathcal{B}]$. The crucial point here is to check that '$\cdot$' as defined above is well defined, i.e., if $\mathcal{A} \sim \mathcal{A}'$ and $\mathcal{B} \sim \mathcal{B}'$ then $\mathcal{A} \cdot \mathcal{B} \sim \mathcal{A}' \cdot \mathcal{B}'$. The set of equivalence classes $\mathcal{I}/\sim$ with this product is actually a group. It is one of the fundamental theorems of algebraic number theory that this group is finite—not just for quadratic extensions of $\mathbf{Q}$ but for any finite extension of $\mathbf{Q}$. The order of this group is called the *class number* of the extension. The class number of $\mathbf{Q}(\sqrt{\Delta})$ will be denoted by $h'(\Delta)$. Note that the class number is one if and only if I is a PID.

Now let Δ be a negative fundamental discriminant, i.e., a negative integer Δ which is congruent to 0 or 1 modulus 4 and which cannot be written in the form $\Delta_0 n^2$ where Δ_0 is another discriminant and n is an integer greater than 1. Hence 4 is the only possible square factor of Δ. Recall that we have defined $h(\Delta)$ to be the number of equivalence classes of primitive binary integral quadratic forms. Remarkably:

Theorem. $h(\Delta) = h'(\Delta)$. In order to prove this we first observe that if $\alpha = a + b\sqrt{\Delta}$ is in the ring of integers I of $\mathbf{Q}(\sqrt{\Delta})$ then so also is $\overline{\alpha} = a - b\sqrt{\Delta}$. Hence so also is $\alpha\overline{\alpha}$ which is an integer. Hence if $\mathcal{A}$ is any non-zero ideal of I then $\mathcal{A} \cap \mathbf{Z} \neq (0)$. Clearly $\mathcal{A} \cap \mathbf{Z}$ is an ideal in $\mathbf{Z}$ so $\mathcal{A} \cap \mathbf{Z} = a\mathbf{Z}$ for some integer $a > 0$. Observe also that any non-zero ideal of I cannot be contained in $\mathbf{Z}$.

In order to make life a bit easier, we will assume in what follows that Δ is odd and hence that $I = \mathbf{Z} + \mathbf{Z}[(1 + \sqrt{\Delta})/2]$ (proof?). Let $\mathcal{A}$ be an ideal in I. Define

$$J = \left\{ r \in \mathbf{Z} | r \cdot \frac{1 + \sqrt{\Delta}}{2} + s \in \mathcal{A} \right\}$$

for some $s \in \mathbf{Z}$. Then J is an ideal in $\mathbf{Z}$ and since $\mathcal{A} \not\subseteq \mathbf{Z}$, J is non-zero. Let $J = t\mathbf{Z}$, $t > 0$. Then there exists an $s \in \mathbf{Z}$ such that $t[(1 + \sqrt{\Delta})/2] + s \in \mathcal{A}$. We claim that $\mathcal{A} = a\mathbf{Z} + [(t + 2s + t\sqrt{\Delta})/2]\mathbf{Z}$. Clearly, the right-hand side is contained in $\mathcal{A}$. Let $\alpha = u + v[(1 + \sqrt{\Delta})/2] \in \mathcal{A}$. Then $v \in J$ so $v = tv'$ for some $v' \in \mathbf{Z}$. Therefore

$$\begin{aligned} \alpha - v'\frac{(t + 2s) + t\sqrt{\Delta}}{2} &= u + tv'\frac{1 + \sqrt{\Delta}}{2} - v'\frac{(t + 2s) + t\sqrt{\Delta}}{2} \\ &= u - sv' \in \mathcal{A} \cap \mathbf{Z} = a\mathbf{Z}. \end{aligned}$$

Therefore, $\alpha \in a\mathbf{Z} + [(t + 2s + t\sqrt{\Delta})/2]\mathbf{Z}$. Hence, every ideal $\mathcal{A}$ in I is of the form $a\mathbf{Z} + [(b + c\sqrt{\Delta})/2]\mathbf{Z}$, $a > 0$, $c > 0$. For this to be an ideal, it must be closed under multiplication by $(1 + \sqrt{\Delta})/2$. Hence $a[(1 + \sqrt{\Delta})/2] \in a\mathbf{Z} + [(b + c\sqrt{\Delta})/2]\mathbf{Z}$, i.e., there exist integers m, n such that $a[(1 + \sqrt{\Delta})/2] = ma + n[(b + c\sqrt{\Delta})/2] \Longrightarrow a = nc$ and $1 = 2m + \frac{b}{c}$ i.e., c divides a, c divides b and $\frac{b}{c}$ is odd. Let $a = tc$, $b = uc$, u odd. Then $a\mathbf{Z} + [(b + c\sqrt{\Delta})/2]\mathbf{Z} = tc\mathbf{Z} + [(uc + c\sqrt{\Delta})/2]\mathbf{Z} = c[t\mathbf{Z} + [(u + \sqrt{\Delta})/2]\mathbf{Z}]$. Hence, every ideal $\mathcal{A}$ in I is of the form $c[t\mathbf{Z} + [(u + \sqrt{\Delta})/2]\mathbf{Z}]$, with $c > 0$, $t > 0$ and u odd. Further, since $\mathcal{A}$ is closed under multiplication by $(1 + \sqrt{\Delta})/2$, $c[u + \sqrt{\Delta})/2][(1 + \sqrt{\Delta})/2] \in \mathcal{A}$. Hence there exist integers h, k such that $[(u + \Delta) + (1 + u)\sqrt{\Delta}]/4 = ht + k[(u + \sqrt{\Delta})/2]$. Therefore, $k = \frac{1+u}{2}$ and $\frac{(u+\Delta)}{4} = ht + \frac{ku}{2} = ht + [u(1 + u)/4]$. Hence $\Delta = u^2 + 4ht$. We have proved:

Proposition. Every ideal in I is of the form $t[a\mathbf{Z} + \{(b + \sqrt{\Delta})/2\}\mathbf{Z}]$ for some integers a, b, t with $t > 0$, $a > 0$ and such that there exists an integer c with $\Delta = b^2 - 4ac$.

Proof of the Theorem. We denote by $[aX^2 + bXY + cY^2]$ the equivalence class of the form $aX^2 + bXY + cY^2$ in $S_1(\Delta)$. We denote by $[\mathcal{A}]$ the equivalence class of the ideal $\mathcal{A}$ in $\mathcal{I}$. Define

$$\epsilon : S_1(\Delta)/\sim \longrightarrow \mathcal{I}/\sim$$

$$[aX^2+bXY+cY^2] \longmapsto \left[a\mathbf{Z}+\frac{b+\sqrt{\Delta}}{2}\mathbf{Z}\right].$$

Then the proposition we have proved above shows that ϵ is subjective. We need, of course, to show that ϵ is well defined. For this we must show that if

$$A\cdot(aX^2+bXY+cY^2)=a'X^2+b'XY+c'Y^2$$

where A is either $\begin{pmatrix}1&1\\0&1\end{pmatrix}$ or $\begin{pmatrix}0&1\\-1&0\end{pmatrix}$ then $a\mathbf{Z}+\left[\frac{b+\sqrt{\Delta}}{2}\right]\mathbf{Z}\sim a'\mathbf{Z}+\left[\frac{b'+\sqrt{\Delta}}{2}\right]\mathbf{Z}$.

If $A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$ then $a'=a$ and $b'=b+2a$ which implies that $a\mathbf{Z}+\frac{b+\sqrt{\Delta}}{2}\mathbf{Z}$ $=a'\mathbf{Z}+\left[(b'+\sqrt{\Delta})/2\right]\mathbf{Z}$.

If $A=\begin{pmatrix}0&1\\-1&0\end{pmatrix}$ then $a'=c$ and $b'=-b$ so

$$a'\mathbf{Z}+\frac{b'+\sqrt{\Delta}}{2}\mathbf{Z}=c\mathbf{Z}+\frac{-b+\sqrt{\Delta}}{2}\mathbf{Z}=\frac{b^2-\Delta}{4a}\mathbf{Z}+\frac{-b+\sqrt{\Delta}}{2}\mathbf{Z}.$$

Therefore, $a\left(a'\mathbf{Z}+\frac{b'+\sqrt{\Delta}}{2}\mathbf{Z}\right)=\frac{(-b+\sqrt{\Delta})}{2}.\left(a\mathbf{Z}+\frac{b+\sqrt{\Delta}}{2}\mathbf{Z}\right)$ and we have proved what was required.

In order to prove our theorem we must show that ϵ is a bijection. Only the injectivity of ϵ is left. Before proving injectivity we make two remarks:

(a) If $\mathcal{A}$ and $\mathcal{B}$ are two ideals in I then they are equivalent if there exists $\alpha,\beta\in I$ such that $\alpha.\mathcal{A}=\beta.\mathcal{B}$. But this is equivalent to $\alpha\bar{\alpha}\mathcal{A}=\bar{\alpha}\beta\mathcal{B}$ and $\alpha\bar{\alpha}$ is a positive integer. Hence $\mathcal{A}\sim\mathcal{B}$ if and only if there exists an integer $t>0$ and $\beta\in I$ such that $t\cdot\mathcal{A}=\beta\cdot\mathcal{B}$.

(b) If $\alpha,\beta\in I$ and $\alpha\mathbf{Z}+\beta\mathbf{Z}=\gamma\mathbf{Z}+\delta\mathbf{Z}$ then there exists an integral 2×2 matrix A of determinant ±1 such that $A\cdot\begin{pmatrix}\alpha\\\beta\end{pmatrix}=\begin{pmatrix}\gamma\\\delta\end{pmatrix}$.

Now suppose that

$$\epsilon\left([aX^2+bXY+cY^2]\right)=\epsilon\left([a'X^2+b'XY+c'Y^2]\right),$$

i.e.,

$$a\mathbf{Z}+\frac{b+\sqrt{\Delta}}{2}\mathbf{Z}\sim a'\mathbf{Z}+\frac{b'+\sqrt{\Delta}}{2}\mathbf{Z}.$$

Hence there exists an integer $t'>0$ and $\alpha=\frac{p+q\sqrt{\Delta}}{2}$ in I such that $\alpha\cdot(a\mathbf{Z}+\frac{b+\sqrt{\Delta}}{2}\mathbf{Z})=t'\cdot(a'\mathbf{Z}+\frac{b'+\sqrt{\Delta}}{2}\mathbf{Z})=\mathcal{A}$(say). We must show that

$$a'X^2+b'XY+c'Y^2=A\cdot(aX^2+bXY+cY^2)$$

for some A in SL(2, $\mathbf{Z}$).

Case 1: Let $q = 0$ and $t = p/2$. Then $at\mathbf{Z} = a't'\mathbf{Z} = \mathcal{A}\cap\mathbf{Z}$. We may without loss of generality assume that $at = a't'$ and hence $t > 0$. There exist integers m, n such that $t[(b+\sqrt{\Delta})/2] = ma't' + nt'[(b'+\sqrt{\Delta})/2]$ which implies that $t = nt'$ and hence $a' = na$. There also exist integers k, l such that $t'[(b' + \sqrt{\Delta})/2] = kta + lt[(b + \sqrt{\Delta})/2]$. Hence $ln = 1$ or $n = 1, t = t'$, $a = a'$ and $b' = b + 2ak$. It is now easy to see that

$$a'X^2 + b'XY + c'Y^2 = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \cdot (aX^2 + bXY + cY^2).$$

Case 2: $(q \neq 0)$. In view of case 1 we may assume that $(p, q) = 1$. By the proposition above and remark (b) there exists an integral matrix $A = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ of determinant ± 1 such that

$$A \cdot \begin{pmatrix} a\frac{p+q\sqrt{\Delta}}{2} \\ \left(\frac{b+\sqrt{\Delta}}{2}\right) \cdot \left(\frac{p+q\sqrt{\Delta}}{2}\right) \end{pmatrix} = \begin{pmatrix} t'a' \\ t' \cdot \frac{b'+\sqrt{\Delta}}{2} \end{pmatrix},$$

or, in fact, by multiplying by the matrix $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, if necessary, we can assume that A is in SL(2, $\mathbf{Z}$) and

$$A \cdot \begin{pmatrix} a\frac{p+q\sqrt{\Delta}}{2} \\ \left(\frac{b+\sqrt{\Delta}}{2}\right) \cdot \left(\frac{p+q\sqrt{\Delta}}{2}\right) \end{pmatrix} = \begin{pmatrix} \pm t'a' \\ t' \cdot \frac{b'+\sqrt{\Delta}}{2} \end{pmatrix}. \tag{2}$$

Therefore, $xa[(p + q\sqrt{\Delta})/2] + y[\{(bp + q\Delta) + (p + bq)\sqrt{\Delta}\}/4] = \pm t'a'$ which implies that $xa(p/2) + y[(bp + q\Delta)/4] = \pm t'a'$ and $xa(q/2) + y[(p + bq)/4] = 0$. Hence, $2xaq = -y(p + bq)$. Let e be the positive g.c.d. of $2a$ and $p + bq$. Then $x[(2aq)/e] = -y[(p + bq)/e]$, so $\frac{2aq}{e}$ divides y and $y = \frac{2aq}{e} . r$ for some integer r. Then $x = -r[(p+bq)/e]$. Since $(x, y) = 1$ we get $r = \pm 1$. A simple calculation now shows that $xa(p/2)+y[(bp+q\Delta)/4] = \pm t'a' = -\frac{2ar}{e}\alpha\bar{\alpha}$. Hence, keeping in view the various signs, we get $t'a' = (2a/e)\alpha\bar{\alpha}$. Furthermore, since $xw - yz = 1$, substituting the values of x and y given above, we get $w(p + bq) + 2aqz = -re$. We further get from (2) that $2a[(p + q\sqrt{\Delta})/2] + w[\{(b + \sqrt{\Delta})/2\}\{(p + q\sqrt{\Delta})/2\}] = t'[(b' + \sqrt{\Delta})/2]$ which implies that $2zap + w(bp + q\Delta) = 2t'b'$ and $2zaq + w(p + bq) = 2t'$, i.e., $-re = 2t'$. Hence $2zap + w(bp + q\Delta) = -reb'$. It is now easy to check that $A^t.(aX^2 + bXY + cY^2) = a'X^2 + b'XY + c'Y^2$. For instance, the coefficient of X^2, if we replace X by $xX + zY$ and Y by $yX + wY$ in the expression $aX^2 + bXY + cY^2$, is $ax^2 + bxy + cy^2$. Substituting $x = -r[(p + bq)/e]$ and $y = \frac{2aq}{e} \cdot r$ and using the fact that $t'a' = (2a/e)\alpha\bar{\alpha}$ and $2t' = -re$, we get $ax^2 + bxy + cy^2 = a'$. Similarly, the coefficient of XY on the required transformation is $2axz + bxw + byz + 2cyw$ which on substitution is just b'. Therefore $A^t \cdot (aX^2 + bXY + cY^2) = a'X^2 + b'XY + c'Y^2$ and ϵ is injective.

This is a beautiful example in mathematics where two apparently unrelated objects turn out to be equal. Maybe the reader can discover some more.

Suggested Reading

[1] D Flath. *Introduction to Number Theory*. John Wiley and Sons. New York, 1989.

[2] J P Serre. *A Course in Arithmetic*. Narosa Publishing House. New Delhi, 1979.

[3] H M Stark. The complete determination of the complex quadratic fields of class number one. *Michigan Math F*. 14. 1–27, 1967.

[4] Algebraic Number Theory. *Mathematical Pamphlets 4*. Tata Institute of Fundamental Research. Mumbai, 1964.

Rajat Tandon
Department of Mathematics
University of Hyderabad
Central University P.O.
Hyderabad 500 046

11

Roots are Not Contained in Cyclotomic Fields

Rajat Tandon

The square root of any integer is contained in a cyclotomic field, i.e., an extension field $\mathbb{Q}(\zeta_n)$ of $\mathbb{Q}$ generated by $\zeta_n = e^{2\pi i/n}$. There is a famous theorem of Kronecker and Weber (see the remarks at the end) which vastly generalises this fact. In what follows, if $\alpha_1, \alpha_2, \ldots, \alpha_n$ are complex numbers, we denote by $\mathbb{Q}(\alpha_1, \alpha_2, \ldots, \alpha_n)$ the smallest subfield of $\mathbb{C}$ containing the $\alpha_i's$. As in the case of Fermat's last theorem (FLT, where $x^n + y^n = z^n$ has integer solutions only in the case $n = 2$), the surprising fact is that other nth roots (other than square roots) are never contained in a cyclotomic extension. Of course, one must exercise a little care. For instance $\sqrt[4]{4} = \sqrt{2}$ is a square root and hence contained in a cyclotomic extension. The point here is that $\sqrt[4]{4}$ is not a genuine fourth root; it is, in fact, a square root.

Definition 1. If a is an integer greater than 1 then the real number $\sqrt[n]{a}$ is said to be a genuine nth root if it cannot be written in the form $\sqrt[m]{b}$ for some integer b and some $m < n$.

In particular, a genuine nth root for $n > 1$ is irrational; for if it is rational, then it is of the form $\sqrt[m]{b}$ for some integer b with $m = 1$. We have the following theorem:

Theorem 2. Let a be any integer. Then, $\sqrt{a}$ is contained in a cyclotomic field. If $\sqrt[n]{a}$ is a genuine nth root where a is an integer greater than 1 and n an integer greater than 2, then $\sqrt[n]{a}$ is not contained in any cyclotomic field.

The first assertion is very well-known and is easy to establish. While proving it, one actually proves a stronger statement viz.,

Proposition 3. If p is a prime, then $\sqrt{(-1)^{(p-1)/2}p} \in \mathbb{Q}(\zeta_p)$.

Observe that if $\sqrt[n]{a}$ is genuine and $a = p_1^{e_1} p_2^{e_2} \ldots p_r^{e_r}$ is the factorisation of a into distinct prime powers, then g.c.d.$(e_1, e_2, \ldots, e_r, n) = 1$. For, if $t = (e_1, e_2, \ldots, e_r, n)$

and $b = p_1^{e_1/t} p_2^{e_2/t} \dots p_r^{e_r/t}$, then $\sqrt[n]{a} = \sqrt[\frac{n}{t}]{b}$. Thus, if n has an odd prime factor p, in order to show that $\sqrt[n]{a}$ is not contained in any cyclotomic extension it suffices to show that $(\sqrt[n]{a})^{n/p} = \sqrt[p]{a}$ is not contained in a cyclotomic extension. On the other hand, if $n = 2^r$, $r \geq 2$, it suffices to show that $(\sqrt[n]{a})^{n/4} = \sqrt[4]{a}$ is not contained in any cyclotomic extension, i.e., it suffices (as in the case of FLT) to prove our theorem for $n = 4$ or p where p is any odd prime.

The proof follows from the following propositions which can be found in any standard text on Galois theory (see, for instance [1]). We will also refer to the article [2] on Galois theory by B Sury which appeared in *Resonance*. In what follows, K and F will always denote subfields of $\mathbb{C}$, and if K is any such field we denote by $G(K)$ the group of automorphisms of K. $[K : F]$ denotes the dimension of K as a vector space over F.

PROPOSITION 4. If $F \subseteq K \subseteq L$ then $[L : F] = [L : K][K : F]$.

It is easy to see that if α_i's form a basis of K over F and β_j's form a basis of L over K, then the $\alpha_i\beta_j$'s form a basis of L over F.

PROPOSITION 5. $[F(\alpha) : F]$ is equal to the degree of the unique monic polynomial f_α of minimal degree in $F[X]$ satisfied by α, and this is the same as the degree of any irreducible polynomial in $F[X]$ satisfied by α.

(See lemma in [2].) It can easily be seen by using the Euclidean algorithm for polynomials that f_α divides any polynomial in $F[X]$ that has α as a root and hence divides any irreducible polynomial g satisfied by α. Irreducibility of g implies that $g = cf_\alpha$ for some constant c in F.

PROPOSITION 6. The group $G(\mathbb{Q}(\zeta_m))$ of automorphisms of the field $\mathbb{Q}(\zeta_m)$ for any $m > 2$ is abelian; in fact, it is isomorphic to the group of units in the ring $\mathbb{Z}/m\mathbb{Z}$.

It is clear that if σ is an automorphism of $\mathbb{Q}(\zeta_m)$ then since $\zeta_m^m = 1$ we get $\sigma(\zeta_m)^m = 1$ so $\sigma(\zeta_m)$ is another mth root of 1. Since an automorphism of a group preserves order and σ is an automorphism of the multiplicative group $(\mathbb{Q}(\zeta_m) - \{0\})$, $\sigma(\zeta_m)$ has order m so $\sigma(\zeta_m) = \zeta_m^i$ for some i coprime to m. We thus have a map $\sigma \rightarrow i$ from $G(\mathbb{Q}(\zeta_m))$ to the group of units in $\mathbb{Z}/m\mathbb{Z}$. That the map is a homomorphism is a simple exercise. It is clear that σ is completely determined by its action on ζ_m since ζ_m generates $\mathbb{Q}(\zeta_m)$. Hence the map is injective. That the map is surjective follows from Proposition 8.

PROPOSITION 7. If $\mathbb{Q} \subseteq F \subseteq K$ where F and K are each generated over $\mathbb{Q}$ by the roots of some polynomials in $\mathbb{Q}[X]$, i.e., F and K are splitting fields of polynomials in $\mathbb{Q}[X]$ (see [2]), then $G(F)$ is isomorphic to $G(K)/G(K/F)$ where $G(K/F)$ denotes the subgroup of $G(K)$ consisting of those automorphisms of K which fix the elements of F. Hence if $G(K)$ is abelian, so is $G(F)$.

This follows easily if we consider the restriction map from $G(K)$ to $G(F)$. The fact that if $\sigma \in G(K)$, then $\sigma(F) = F$ follows from the fact that F is normal over $\mathbb{Q}$

(refer [2], Box 14). For, suppose $F = \mathbb{Q}(\alpha_1, \alpha_2, \dots, \alpha_n)$, where the α_i's are roots of some polynomial $f(x)$ in $\mathbb{Q}[X]$. Since $f(\alpha_i) = 0$ we have $\sigma(f(\alpha_i)) = 0$. But $\sigma(f(\alpha_i)) = f(\sigma(\alpha_i))$, so $\sigma(\alpha_i)$ must be another root of f, i.e., $\sigma(\alpha_i) = \alpha_j$ for some j; σ permutes the roots of f and so $\sigma(F) = F$.

PROPOSITION 8. If K is generated over F by the roots of some polynomial in $F[X]$ and α, α' are two roots in K of an irreducible polynomial in $F[X]$, then there exists an automorphism σ in $G(K/F)$ such that $\sigma(\alpha) = \alpha'$.

We have an isomorphism (just the substitution map) from $\frac{F[X]}{(f)}$ to $F(\alpha)$ which maps $X + (f)$ to α and similarly an isomorphism from $\frac{F[X]}{(f)}$ to $F(\alpha')$ which maps $X + (f)$ to α'. Hence we have an isomorphism from $F(\alpha)$ to $F(\alpha')$ which maps α to α'. This map extends to an automorphism of K (see Proposition 5.2 in [1]).

PROPOSITION 9. If p is an odd prime or 4 and if $\sqrt[p]{a}$ is genuine with $a > 1$ then $G(\sqrt[p]{a}, \zeta_p)$ is not abelian.

If p is an odd prime, $\mathbb{Q}(\zeta_p)$ is the field generated over $\mathbb{Q}$ by the roots of the polynomial $1 + X + X^2 + \cdots + X^{p-1}$. If p is an odd prime or 4 and $a > 1$, then $\mathbb{Q}(\sqrt[p]{a}, \zeta_p)$ is the field generated over $\mathbb{Q}$ by the roots of $X^p - a$. Both these polynomials are irreducible over $\mathbb{Q}$. Hence

$$[\mathbb{Q}(\zeta_p) : \mathbb{Q}] = \begin{cases} p - 1 & \text{if } p \text{ is odd} \\ 2 & \text{if } p = 4. \end{cases}$$

and $[\mathbb{Q}(\sqrt[p]{a}) : \mathbb{Q}] = p$. Hence by Proposition 4

$$[\mathbb{Q}(\sqrt[p]{a}, \zeta_p) : \mathbb{Q}] = \begin{cases} p(p - 1) & \text{if } p \text{ is odd} \\ 8 & \text{if } p = 4. \end{cases}$$

It follows again by Proposition 4 that $[\mathbb{Q}(\sqrt[p]{a}, \zeta_p) : \mathbb{Q}(\zeta_p)] = p$ and $[\mathbb{Q}(\sqrt[p]{a}, \zeta_p) : \mathbb{Q}(\sqrt[p]{a})] = p - 1$ or 2 according as p is odd or 4, respectively. Hence by Proposition 5, $X^p - a$ is irreducible over $\mathbb{Q}(\zeta_p)$ and $1 + X + X^2 + \cdots + X^{p-1}$ is irreducible over $\mathbb{Q}(\sqrt[p]{a})$ if p is odd whilst $X^2 + 1$ is irreducible over $\mathbb{Q}(\sqrt[4]{a})$. Observe that $\sqrt[p]{a}$ and $\sqrt[p]{a}\zeta_p$ are roots of $X^p - a$ and ζ_p and ζ_p^2 are roots of $1 + X + \cdots + X^{p-1}$ if p is odd whereas ζ_p and ζ_p^3 are roots of $X^2 + 1$ if $p = 4$. Hence by Proposition 8 there exists an automorphism $\sigma \in G(\mathbb{Q}(\sqrt[p]{a}, \zeta_p) : \mathbb{Q}(\zeta_p))$ (i.e. σ fixes ζ_p) such that $\sigma(\sqrt[p]{a}) = \sqrt[p]{a}\zeta_p$ and there exists a $\tau \in G(\mathbb{Q}(\sqrt[p]{a}, \zeta_p) : \mathbb{Q}(\sqrt[p]{a}))$ (i.e. τ fixes $\sqrt[p]{a}$) such that $\tau(\zeta_p) = \zeta_p^2$ if p is odd and $\tau(\zeta_p) = \zeta_p^3$ if $p = 4$. Hence,

$$\sigma\tau(\sqrt[p]{a}) = \sigma(\sqrt[p]{a}) = \sqrt[p]{a}\zeta_p,$$

whereas

$$\tau\sigma(\sqrt[p]{a}) = \tau(\sqrt[p]{a}\zeta_p) = \tau(\sqrt[p]{a})\tau(\zeta_p) = \begin{cases} \sqrt[p]{a}\zeta_p^2 & \text{if } p \text{ is odd} \\ \sqrt[p]{a}\zeta_p^3 & \text{if } p = 4. \end{cases}$$

In either case $\sigma\tau \neq \tau\sigma$ and $G(\mathbb{Q}(\sqrt[p]{a}, \zeta_p))$ is not abelian.

Observe that if $\mathbb{Q}(\sqrt[n]{a}) \subseteq \mathbb{Q}(\zeta_m)$, then $\mathbb{Q}(\sqrt[n]{a}, \zeta_n) \subseteq \mathbb{Q}(\zeta_m, \zeta_n) = \mathbb{Q}(\zeta_{[m,n]})$ where $[m, n]$ is the l.c.m. of m and n. If $\mathbb{Q}(\sqrt[p]{a}, \zeta_p)$ was contained in the cyclotomic extension $\mathbb{Q}(\zeta_m)$, its group of automorphisms would, by Proposition 7, be the quotient of the abelian group $G(\mathbb{Q}(\zeta_m))$, and hence abelian.

Remarks

We have apparently proved the stronger result that for a genuine pth root $\sqrt[p]{a}$ (with p an odd prime and $a > 1$), the Galois extension field generated by it is not an abelian extension of $\mathbb{Q}$. However, this is not really a stronger statement. The deep theorem of Kronecker and Weber referred to in the introduction says that any abelian extension of $\mathbb{Q}$ is contained in a cyclotomic extension. The interesting question is whether one can similarly obtain the abelian extensions of any algebraic number field by adjoining special values of transcendental functions. For imaginary quadratic fields $\mathbb{Q}(\sqrt{-d})$, this has been solved using the so-called theory of complex multiplication. Roughly, the role of the function $e^{2\pi ix}$ is taken by the elliptic modular j-function and the values are considered at points of finite order on the elliptic curves (in place of the circle as was in the case of the Kronecker–Weber theorem). The general question is known as Kronecker's 'jugendtraum' (the german word means 'dream of youth') and is still open. It is one of the famous 'Hilbert problems' (the 12th problem). Hilbert writes in his 1900 address at the International Congress of Mathematicians that the extension of Kronecker's theorem to any algebraic number field seems to him to be of the greatest importance and that he regards this as one of the most profound and far-reaching problems in the theory of numbers.

Suggested Reading

[1] M Artin. *Algebra*. Prentice-Hall of India. New Delhi. 1994.

[2] B Sury. The theory of equations and the birth of modern group theory. *Resonance*, Vol 4, No. 10, 1999.

Rajat Tandon
Department of Mathematics
and Statistics
University of Hyderabad
Hyderabad 500 046

12

Die ganzen zahlen hat Gott gemacht

Polynomials with Integer Values

B Sury

A quote attributed to the famous mathematician L Kronecker is '*Die ganzen zahlen hat Gott gemacht, alles andere ist menschenwerk.*' A translation might be '*God gave us integers and all else is man's work.*' All of us are familiar already from middle school with the similarities between the set of integers and the set of all polynomials in one variable. A paradigm of this is the Euclidean (division) algorithm. However, it requires an astute observer to notice that one has to deal with polynomials with real or rational coefficients rather than just integer coefficients for a strict analogy. There are also some apparent dissimilarities—for instance, there is no notion among integers corresponding to the derivative of a polynomial. In this discussion, we shall consider polynomials with integer coefficients. Of course a complete study of this encompasses the whole subject of algebraic number theory, one might say. For the most of this paper (in fact, with the exception of Lemma 5, Lemma 7 and Exercise 3), we adhere to fairly elementary methods and address a number of rather natural questions. To give a prelude, one such question might be "if an integral polynomial takes only values which are perfect squares, then must it be the square of a polynomial?" Note that for a natural number n, the polynomial $\binom{X}{n} = \frac{X(X-1)\cdots(X-n+1)}{n(n-1)\cdots 1}$ takes integer values at all integers although it does not have integer coefficients. By Z, we shall denote the set of integers.

Prime Values and Irreducibility

The first observation about polynomials taking integral values is

LEMMA 1. A polynomial P takes Z to Z if, and only if $P(X) = a_0 + a_1\binom{X}{1} + \cdots + a_n\binom{X}{n}$ with $a_i \in Z$.

PROOF. The sufficiency is evident. For the converse, we first note that any polynomial whatsoever can be written in this form for some n and some (possibly nonintegral) a_i's. Writing P in this form and assuming that $P(Z) \subset Z$, we have

$$P(0) = a_0 \in Z$$
$$P(1) = a_0 + a_1 \in Z$$
$$P(2) = a_0 + a_1 \binom{2}{1} + a_2 \in Z$$

and so on. Inductively, since $P(m) \in Z \; \forall m$, we get $a_i \in Z \; \forall i$.

COROLLARY 1. If a polynomial P takes Z to Z and has degree n, then $n!P(X) \in Z[X]$.

LEMMA 2. A nonconstant integral polynomial $P(X)$ cannot take only prime values.

PROOF. If all values are composite, then there is nothing to prove. So assume that $P(a) = p$ for some integer a and prime p. Now, as P is nonconstant,

$$\lim_{n \to \infty} |P(a + np)| = \infty.$$

So, for big enough n, $|P(a + np)| > p$. But $P(a + np) \equiv P(a) \equiv 0 \bmod p$, which shows $P(a + np)$ is composite.

REMARK 1. Infinitely many primes can occur as integral values of a polynomial. For example, if $(a, b) = 1$, then the well-known (but deep) Dirichlet's theorem on primes in progression shows that the polynomial $aX + b$ takes infinitely many prime values. In general, it may be very difficult to decide whether a given polynomial takes infinitely many prime values. For instance, it is not known if $X^2 + 1$ represents infinitely many primes. In fact, there is no polynomial of degree ≥ 2 which is known to take infinitely many prime values.

LEMMA 3. If P is a nonconstant, integral-valued polynomial, then the number of prime divisors of its values $\{P(m)\}_{m \in Z}$, is infinite, i.e., not all terms of the sequence $P(0), P(1), \ldots$ can be built from finitely many primes.

PROOF. It is clear from Corollary 1 above that it is enough to prove this for $P(X) \in Z[X]$, which we will henceforth assume. Now, $P(X) = \sum_{i=0}^{n} a_i X^i$, where $n \geq 1$. If $a_0 = 0$, then clearly $P(p) \equiv 0 \bmod p$ for any prime p. If $a_0 \neq 0$, let us consider for any integer t the polynomial

$$P(a_0 t X) = \sum_{i=0}^{n} a_i (a_0 t X)^i = a_0 \left\{ 1 + \sum_{i=1}^{n} a_i a_0^{i-1} t^i X^i \right\} = a_0 Q(X).$$

There exists some prime number p such that $Q(m) \equiv 0 \bmod p$ for some m and some prime p, because Q can take the values 0, 1, -1 only at finitely many points. Since $Q(m) \equiv 1 \bmod t$, we have $(p, t) = 1$. Then $P(a_0 t m) \equiv 0 \bmod p$. Since t was arbitrary, the set of p arising in this manner is infinite.

REMARK 2.

(a) Note that it may be possible to construct infinitely many terms of the sequence $\{P(m)\}_{m \in Z}$ using only a finite number of primes. For example, take $(a, d) = 1, a \geq d \geq 1$. Since, by Euler's theorem, $a^{\varphi(d)} \equiv 1 \bmod d$, the numbers $\frac{a(a^{\varphi(d)n} - 1)}{d} \in Z \ \forall \ n$. For the polynomial $P(X) = dX + a$, the infinitely many values $P(\frac{a}{d}(a^{\varphi(d)n} - 1)) = a^{\varphi(d)n+1}$ have only prime factors coming from primes dividing a.

(b) In order that the values of an integral polynomial $P(X)$ be prime for infinitely many integers, $P(X)$ must be irreducible over Z and of content 1. By content, we mean the greatest common divisor of the coefficients.

Box 1. Eisenstein's Criterion and More

Perhaps the only general criterion known to check whether an integral polynomial of a special kind is irreducible is due to G Eisenstein, a student of Gauss and an outstanding mathematician, whom Gauss is said to have rated above himself. Eisenstein died when he was 27.

Let $f(X) = a_0 + a_1X + \cdots + a_nX^n$ be an integral polynomial satisfying the following property with respect to some prime p. The prime p divides $a_0, a_1, \ldots, a_{n-1}$ but does not divide a_n. Also, assume that p^2 does not divide a_0. Then, f is irreducible.

The proof is indeed very simple high school algebra. Suppose, if possible, that $f(X) = g(X)h(X) = (b_0 + b_1X + \cdots + b_rX^r)(c_0 + c_1X + \cdots + c_sX^s)$ with $r, s \geq 1$. Comparing coefficients, one has

$$a_0 = b_0c_0, \quad a_1 = a_0b_1 + b_0a_1, \ldots, a_n = b_rc_s, \quad r + s = n.$$

Since $a_0 = b_0c_0 \equiv 0 \bmod p$, either $b_0 \equiv 0 \bmod p$ or $c_0 \equiv 0 \bmod p$.

To fix notations, we may assume that $b_0 \equiv 0 \bmod p$. Since $a_0 \not\equiv 0 \bmod p^2$, we must have $c_0 \not\equiv 0 \bmod p$. Now $a_1 = b_0c_1 + b_1c_0 \equiv b_1c_0 \bmod p$; so $b_1 \equiv 0 \bmod p$. Proceeding inductively in this manner, it is clear that all the b_i's are multiples of p. This is a manifest contradiction of the fact that $a_n = b_rc_s$ is not a multiple of p. This finishes the proof.

It may be noted that one may reverse the roles of a_0 and a_n and obtain another version of the criterion:

Let $f(X) = a_0 + a_1X + \cdots + a_nX^n$ be an integral polynomial satisfying the following property with respect to some prime p. The prime p divides $a_1, a_2, \ldots, a_n$ but does not divide a_0. Also, assume that p^2 does not divide a_n. Then, f is irreducible.

The following generalisation is similar to prove and is left as an exercise.

Let $f(X) = a_0 + a_1X + \cdots + a_nX^n$ be an integral polynomial satisfying the following property with respect to some prime p. Let t be such that the prime p divides $a_0, a_1, \ldots, a_{n-t}$ but does not divide a_n. Also, assume that p^2 does not divide a_0. Then, f is either irreducible or has a nonconstant factor of degree less that t.

In general, it is difficult to decide whether a given integral polynomial is irreducible or not. We note that the irreducibility of $P(X)$ and the condition that it have content 1 are not sufficient to ensure that $P(X)$ takes infinitely many prime values. For instance, the polynomial $X^n + 105X + 12$ is irreducible, by Eisenstein's criterion (see Box 1). But, it cannot take any prime value because it takes only even values, and it does not take either of the values ± 2 since both $X^n + 105X + 10$ and $X^n + 105X + 14$ are irreducible, again by Eisenstein's criterion.

LEMMA 4. Let $a_1, \ldots, a_n$ be distinct integers. Then $P(X) = (X-a_1)\cdots(X-a_n)-1$ is irreducible.

PROOF. Suppose, if possible, $P(X) = f(X)g(X)$ with deg. f, deg. $g < n$. Evidently, as $f(a_i)g(a_i) = -1$, $f(a_i) = -g(a_i) = \pm 1 \; \forall \; 1 \leq i \leq n$. Now, $f(X) + g(X)$ being a polynomial of degree $<n$ which vanishes at the n distinct integers, $a_1, \ldots, a_n$ must be identically zero. This gives $P(X) = -f(X)^2$, but this is impossible as can be seen by comparing the coefficients of X^n.

EXERCISE 1. Let n be odd and $a_1, \ldots, a_n$ be distinct integers. Prove that $(X-a_1)\cdots(X-a_n)+1$ is irreducible.

Let us consider the following situation. Suppose $p = a_n \ldots a_0$ is a prime number expressed in the usual decimal system, i.e., $p = a_0 + 10a_1 + 100a_2 + \cdots + 10^n a_n$, $0 \leq a_i \leq 9$. Then, is the polynomial $a_0 + a_1X + \cdots + a_nX^n$ irreducible? This is, in fact, true and, more generally,

LEMMA 5. Let $P(X) \in Z[X]$ and assume that there exists an integer n such that

(i) the zeros of P lie in the half plane Re $(z) < n - \frac{1}{2}$,

(ii) $P(n-1) \neq 0$,

(iii) $P(n)$ is a prime number.

Then $P(X)$ is irreducible.

PROOF. Suppose, if possible $P(X) = f(X)g(X)$ over Z. All the zeros of $f(X)$ also lie in $\mathrm{Re}(z) < n - \frac{1}{2}$. Therefore, $|f(n - \frac{1}{2} - t)| < |f(n - \frac{1}{2} + t)| \forall t > 0$. Since $f(n-1) \neq 0$ and $f(n-1)$ is integral, we have $|f(n-1)| \geq 1$. Thus $|f(n)| > |f(n-1)| \geq 1$. A similar thing holding for $g(X)$, we get that $P(n)$ has proper divisors $f(n)$, $g(n)$ which contradicts our hypothesis.

Irreducibility and Congruence Modulo *p*

For an integral polynomial to take the value zero at an integer or even to be reducible, it is clearly necessary that these properties hold modulo any integer m. Conversely, if $P(X)$ has a root modulo any integer, it must itself have a root in Z. In fact, if $P(X) \in Z[X]$ has a linear factor modulo all but finitely many prime numbers, the $P(X)$ itself has a linear factor. This fact can be proved only by deep methods, viz. using the so-called Ćebotarev density theorem. On the other hand, (see Lemma 7) it was first observed by Hilbert that the reducibility of a polynomial modulo every integer is not sufficient to guarantee its reducibility over Z. Regarding roots of a polynomial modulo a prime, there is following general result due to Lagrange:

LEMMA 6. Let p be a prime number and let $P(X) \in Z[X]$ be of degree n. Assume that not all coefficients of P are multiples of p. Then the number of solutions mod p to $P(X) \equiv 0 \bmod p$ is, at the most, n.

The proof is obvious using the division algorithm over Z/p. In fact, the general result of this kind (provable by the division algorithm again) is that a nonzero polynomial over any field has at the most its degree number of roots.

REMARK 3. Since $1, 2, \ldots, p-1$ are solutions to $X^{p-1} \equiv 1 \bmod p$, we have $X^{p-1} -1 \equiv (X-1)(X-2)\cdots(X-(p-1)) \bmod p$. For odd p, putting $X = 0$ gives Wilson's theorem that $(p-1)! \equiv -1 \bmod p$.

Note that we have observed earlier that any non-constant integral polynomial has a root modulo infinitely many primes. However, as first observed by Hilbert, the reducibility of a polynomial modulo every integer does not imply its reducibility over Z. For example, we have the following result:

LEMMA 7. Let p, q be odd prime numbers such that $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right) = 1$ and $p \equiv 1 \bmod 8$. Here $\left(\frac{p}{q}\right)$ denotes the Legendre symbol defined to be 1 or -1 according as p is a square or not modulo q. Then, the polynomial $P(X) = (X^2 - p - q)^2 - 4pq$ is irreducible, whereas it is reducible modulo any integer.

PROOF.

$$\begin{aligned} P(X) &= X^4 - 2(p+q)X^2 + (p-q)^2 \\ &= (X - \sqrt{p} - \sqrt{q})(X + \sqrt{p} + \sqrt{q})(X - \sqrt{p} + \sqrt{q})(X + \sqrt{p} - \sqrt{q}). \end{aligned}$$

Since $\sqrt{p}, \sqrt{q}, \sqrt{p} \pm \sqrt{q}, \sqrt{pq}$ are all irrational, none of the linear or quadratic factors of $P(X)$ are in $Z[X]$, i.e., $P(X)$ is irreducible. Note that it is enough to show that a factorisation of P exists modulo any prime power as we can use Chinese reminder theorem to get a factorisation modulo a general integer.

Now, $P(X)$ can be written in the following ways:

$$\begin{aligned} P(X) &= X^4 - 2(p+q)X^2 + (p-q)^2 \\ &= (X^2 + p - q)^2 - 4pX^2 \\ &= (X^2 - p + q)^2 - 4qX^2 \\ &= (X^2 - p - q)^2 - 4pq. \end{aligned}$$

The second and third equalities above show that $P(X)$ is reducible modulo any p^n and any q^n. Also since $p \equiv 1 \bmod 8$, p is a square modulo any 2^n and the second equality above again shows that $P(X)$ is the difference of two squares modulo 2^n, and hence reducible mod 2^n.

If ℓ is a prime $\neq 2, p, q$, let us show now that $P(X)$ is reducible modulo l^n for any n.

At least one of $\left(\frac{p}{\ell}\right)$, $\left(\frac{q}{\ell}\right)$ and $\left(\frac{pq}{\ell}\right)$ is 1 because, by the product formula for Legendre symbols, $\left(\frac{p}{\ell}\right) \cdot \left(\frac{q}{\ell}\right) \cdot \left(\frac{pq}{\ell}\right) = 1$. According as $\left(\frac{p}{\ell}\right)$, $\left(\frac{q}{\ell}\right)$ or $\left(\frac{pq}{\ell}\right) = 1$, the second, third or fourth equality shows that $P(X)$ is reducible mod ℓ^n for any n.

We end this section with a result of Schur whose proof is surprising and elegant as well. This is:

SCHUR'S THEOREM. For any n, the truncated exponential polynomial $E_n(X) = n!\left(1 + X + \frac{X^2}{2!} + \cdots + \frac{X^n}{n!}\right)$ is irreducible over $\mathbf{Z}$.

Just for this proof, we need some nontrivial number theoretic facts. A reader unfamiliar with these notions but who is prepared to accept at face value a couple of results can still appreciate the beauty of Schur's proof. Here is where we have to take recourse to some very basic facts about prime decomposition in algebraic number fields. Suppose, if possible, that $E_n(X) = f(X)g(X)$ for some nonconsant, irreducible integral polynomial f. Let us write $f(X) = a_0 + a_1X + \cdots + X^r$ (evidently, we may take the top coefficients of f to be 1). Start with any (complex) root α of f and look at the field $K = \mathbf{Q}(\alpha)$ of all those complex numbers which can be written as polynomials in α with coefficients from $\mathbf{Q}$. The basic fact that we will be using (without proof) is that any nonzero ideal in 'the ring of integers of K' (i.e., the subring O_K of K made up of those elements, which satisfy a monic integral polynomial) is uniquely a product of nonzero prime ideals and a prime ideal can occur at the most deg f times. This is a good replacement for K of the usual unique factorisation of natural numbers into prime numbers. The proof also uses a fact about prime numbers observed by Sylvester but is not trivial to prove.

SYLVESTER'S THEOREM. If $m \geq r$, then $(m+1)(m+2)\cdots(m+r)$ has a prime factor $p > r$.

The special case $m = r$ is known as Bertrand's postulate.

PROOF OF SCHUR'S THEOREM. Now, the proof uses the following fact which is interesting in its own right:

Any prime dividing the constant term a_0 of f is less than the degree r of f.

To see this, note first that $N(\alpha)$, the 'norm of α' (a name for the product of all the roots of the minimal polynomial f of α), is a_0 upto sign. So, there is a prime ideal P of O_K such that $(\alpha) = P^kI$, $(p) = P^lJ$, where I, J are indivisible by P and k, $l \geq 1$. Here, (α) and (p) denote, respectively, the ideal of O_K generated by α and p. Since $E_n(\alpha) = 0$, we have

$$0 = n! + n!\alpha + n!\alpha^2/2! + \cdots + \alpha^n.$$

We know that the exact power of p dividing $n!$ is

$$h_n = [n/p] + [n/p^2] + \cdots.$$

Thus, in O_K, the ideal $(n!)$ is divisible by P^{lh_n} and no higher power of P. Similarly, for $1 \leq i \leq n$, the ideal generated by $n!\alpha^i/i!$ is divisible by $P^{lh_n - lh_i + ki}$. Because of the equality

$$-n! = n!\alpha + n!\alpha^2/2! + \cdots + \alpha^n,$$

it follows that we cannot have each $lh_n - lh_i + ki$ strictly bigger than lh_n, which is the exact power of P dividing the left-hand side. Therefore, there is some i such that $-lh_i + ki \leq 0$. Thus,

$$i \leq ki \leq lh_i = l([i/p] + [i/p^2] + \cdots) < \frac{li}{p-1}.$$

Thus, $p - 1 < l \leq r$, i.e., $p \leq r$. This confirms the observation.

To continue with the proof, we may clearly assume that the degree r of f is at most $n/2$. Now, we use Sylvester's theorem to choose a prime $q > r$ dividing the product $n(n-1)\cdots(n-r+1)$. Note that we can use this theorem because the smallest term $n-r+1$ of this r-fold consecutive product is bigger than r as $r \leq n/2$. Note also that the observation tells us that q cannot divide a_0. Now, we shall write $E_n(X)$ modulo the prime q. By choice, q divides the coefficients of X^i for $0 \leq i \leq n-r$.

So, $f(X)g(X) \equiv X^n + n!\frac{X^{n-1}}{(n-1)!} + \cdots + n!\frac{X^{n-r+1}}{(n-r+1)!} \bmod q$.

Write $f(X) = a_0 + a_1 X + \cdots + X^r$ and $g(X) = b_0 + b_1 X + \cdots + X^{n-r}$.

The above congruence gives $a_0 b_0 \equiv 0$, $a_0 b_1 + a_1 b_0 \equiv 0$ etc. mod q until the coefficient of X^{n-r} of $f(X)g(X)$. As $a_0 \not\equiv 0 \bmod q$, we get recursively (this is just like the proof of Eisenstein's criterion – see Box 1) that

$$b_0 \equiv b_1 \equiv \cdots b_{n-r} \equiv 0 \bmod q.$$

This is impossible as $b_{n-r} = 1$. Thus, Schur's assertion follows.

Polynomials taking Square Values

If an integral polynomial takes only values which are squares, is it true that the polynomial itself is a square of a polynomial? In this section, we will show that this, and more, is indeed true.

LEMMA 8. Let $P(X)$ be a Z-valued polynomial which is irreducible. If P is not a constant, then there exist arbitarily large integers n such that $P(n) \equiv 0$ and $P(n)$ $\not\equiv 0 \bmod p^2$ for some prime p.

PROOF. First, suppose that $P(X) \in Z[X]$. Since P is irreducible, P and P' have no common factors. Write $f(X)P(X) + g(X)P'(X) = 1$ for some $f, g \in Z[X]$. By Lemma 3 there is a prime p such that $P(n) \equiv 0 \bmod p$, where n can be as large as we want. So, $P'(n) \not\equiv 0 \bmod p$ as $f(n)P(n) = g(n)P'(n) = 1$. Since $P(n+p) - P(n)$ $\equiv P'(n) \bmod p^2$, either $P(n+p)$ or $P(n)$ is $\not\equiv 0 \bmod p^2$. To prove the result for general P, one can replace P by $m!P$ where $m = \deg P$.

LEMMA 9. Let $P(X)$ be a Z-valued polynomial such that the zeros of smallest multiplicity have multiplicity m. Then, there exist arbitrarily large integers n such that $P(n) \equiv 0 \bmod p^m$, $P(n) \not\equiv 0 \bmod p^{m+1}$ for some prime p.

PROOF. Let $P_1(X), \ldots, P_r(X)$ be the distinct irreducible factors of $P(X)$. Write $P(X) = P_1(X)^{m_1} \cdots P_r(X)^{m_r}$ with $m = m_1 \leq \cdots m_r$. By the above Lemma, one can find arbitrarily large n such that for some prime p, $P_1(n) \equiv 0 \bmod p$, $P_1(n) \not\equiv 0 \bmod p^2$ and, $P_i(n) \not\equiv 0 \bmod p$ for $i > 1$. Then, $P(n) \equiv 0 \bmod p^m$ and $\not\equiv 0 \bmod p^{m+1}$

COROLLARY 2. If $P(X)$ takes at every integer, a value which is the kth power of an integer, then $P(X)$ itself is the kth power of a polynomial.

PROOF. If $P(X)$ is not an exact kth power, then one can write $P(X) = f(X)^k g(X)$ for polynomials f, g so that $g(X)$ has a zero whose multiplicity is $<k$. Once again, we can choose n and a prime p such that $g(n) \equiv 0 \bmod p$, $\not\equiv 0 \bmod p^k$. This contradicts the fact that $P(n)$ is a kth power.

[2] is an excellent source of results of this nature.

Cyclotomic Polynomials

These were referred to already in an earlier article ([1]). It was also shown there that one could use these polynomials to prove the existence of infinitely many primes congruent to 1 modulo n for any n. For a natural number d, recall that the cyclotomic polynomial $\Phi_d(X)$ is the irreducible, monic polynomial whose roots are the primitive dth roots of unity, i.e., $\Phi_d(X) = \prod_{a \leq d:(a,\, d)=1}(X - e^{2\pi i a/d})$. Note that $\Phi_1(X) = X - 1$ and that for a prime p, $\Phi_p(X) = X^{p-1} + \cdots + X + 1$. Observe that for any $n \geq 1$, $X^n - 1 = \prod_{d/n} \Phi_d(X)$.

EXERCISE 2. Prove that for any d, $\Phi_d(X)$ has integral coefficients, and is irreducible over Z.

Factorising an integral polynomial into irreducible factors is far from easy. Even if we know the irreducible factors, it might be difficult to decide whether a given polynomial divides another given one.

EXERCISE 3.

(a) Given positive integers $a_1 < \cdots < a_n$, consider the polynomials $P(X) = \prod_{i>j}(X^{a_i - a_j} - 1)$ and $Q(X) = \prod_{i>j}(X^{i-j} - 1)$. By factorising into cyclotomic polynomials, prove that $Q(X)$ divides $P(X)$. Conclude that $\prod_{i>j} \frac{a_i - a_j}{i - j}$ is always an integer.

(b) Consider the $n \times n$ matrix A whose (i, j)th entry is the Gaussian polynomial
$$\begin{bmatrix} a_i \\ j-1 \end{bmatrix}.$$
Compute det A to obtain part (a) again.

Here, for $m \geq r$, the Gaussian polynomial is defined as

$$\begin{bmatrix} m \\ r \end{bmatrix} = \frac{(X^m - 1)(X^{m-1} - 1)\cdots(X^{m-r+1} - 1)}{(X^r - 1)(X^{r-1} - 1)\cdots(X - 1)}.$$

Note that
$$\begin{bmatrix} m \\ r \end{bmatrix} = \begin{bmatrix} m-1 \\ r-1 \end{bmatrix} + X^r \begin{bmatrix} m-1 \\ r \end{bmatrix}.$$

It seems from looking at $\Phi_p(X)$ for prime p as though the coefficients of the cyclotomic polynomials $\Phi_d(X)$ for any d are all 0, 1 or -1. However, the following rather amazing fact was discovered by Schur. His proof uses a consequence of a deep result about prime numbers known as the prime number theorem. The prime number theorem tells us that $\pi(x) \sim x/\log(x)$ as $x \to \infty$. Here $\pi(x)$ denotes the number of primes until x. The reader does not need to be familiar with the prime number theorem but is urged to take on faith the consequence of it that for any constant c, there is n such that $\pi(2^n) \geq cn$.

PROPOSITION 1. Every integer occurs as a coefficient of some cyclotomic polynomial.

PROOF. First, we claim that for any integer $t > 2$, there are primes $p_1 < p_2 < \cdots < p_t$ such that $p_1 + p_2 > p_t$. Suppose this is not true. Then, for some $t > 2$, every set of t primes $p_1 < \cdots < p_t$ satisfies $p_1 + p_2 \leq p_t$. So, $2p_1 < p_t$. Therefore, the number of primes between 2^k and 2^{k+1} for any k is less than t. So, $\pi(2^k) < kt$. This contradicts the prime number theorem as noted above. Hence, it is indeed true that for any integer $t > 2$, there are primes $p_1 < p_2 < \cdots < p_t$ such that $p_1 + p_2 > p_t$.

Now, let us fix any odd $t > 2$. We shall demonstrate that both $-t + 1$ and $-t + 2$ occur as coefficents. This will prove that all negative integers occur as coefficients. Then, using the fact that for an odd $m > 1$, $\Phi_{2m}(X) = \Phi_m(-X)$, we can conclude that all integers are coefficients.

Consider now primes $p_1 < p_2 < \cdots < p_t$ such that $p_1 + p_2 > p_t$. Write $p_t = p$ for simplicity. Let $n = p_1 \cdots p_t$ and let us write $\Phi_n(X)$ modulo X^{p+1}. Since $X^n - 1 = \prod_{d/n} \Phi_d(X)$, and since $p_1 + p_2 > p_t$, we have

$$\Phi_n(X) \equiv \prod_{i=1}^{t} \frac{1 - X^{p_i}}{1 - X} \equiv (1 + \cdots + X^p)(1 - X^{p_1})\cdots(1 - X^{p_t})$$
$$\equiv (1 + \cdots + X^p)(1 - X^{p_1} - \cdots - X^{p_t}) \bmod X^{p+1}.$$

Therefore, the coefficients of X^p and X^{p-2} are $1 - t$ and $2 - t$, respectively. This completes the proof. Note that in the proof we have used the fact that if $P(X) = (1 - X^r)Q(X)$ for a polynomial $Q(X)$, then $Q(X) = P(X)(1 + X^r + X^{2r} + \cdots + \cdots)$ modulo any X^k.

EXERCISE 4.

(a) Let $A = (a_{ij})$ be a matrix in $GL(n, Z)$, i.e., both A and A^{-1} have integer entries. Consider the polynomials $p_i(X) = \sum_{j=0}^{n} a_{ij} X^j$ for $0 \leq i \leq n$.

Prove that any integral polynomial of degree at most n is an integral linear combination of the $p_i(X)$. In particular, if $a_0, \ldots, a_n \in Q$ are distinct, show that any rational polynomial of degree at most n is of the form $\sum_{i=0}^{n} \lambda_i (X + a_i)^n$ for some $\lambda_i \in Q$.

(b) Prove that $1 + X + \cdots + X^n = \sum_{i=0}^{[\frac{n}{2}]} (-1)^i \binom{n-i}{i} X^i (1+X)^{n-2i}$. Conclude that $\sum_{i \geq 0} \binom{n-i}{i} = \frac{\alpha^{n+1} - \beta^{n+1}}{\sqrt{5}}$, where $\alpha = \frac{1+\sqrt{5}}{2}$, $\beta = \frac{1-\sqrt{5}}{2}$. This is known as Binet's formula. Further, compute $\sum_{i \geq 0} (-1)^i \binom{n-i}{i}$.

REMRAK 4. It is easily seen by induction that $\sum_{i \geq 0} \binom{n-i}{i}$ is just the $(n+1)$th Fibonacci number F_{n+1}.

As remarked earlier, even for a polynomial of degree 2 (like $X^2 + 1$) it is unknown whether it takes infinitely many prime values. A general conjecture (Bouniakowsky, Schinzel and Sierpinski) in this context is:

> *A nonconstant irreducible integral polynomial whose coefficients have no nontrivial common factor always takes on a prime value.*

We end with an open question which is typical of many number-theoretic questions—a statement which can be understood by the proverbial layman but an answer which proves elusive to this day to professional mathematicians. For any irreducible, monic, integral polynomial $P(X)$, define its *Mahler measure* to be $M(P) = \prod_i \text{Max}(|\alpha_i|, 1)$, where the product is over the roots of P. The following is an easy exercise.

EXERCISE 5. $M(P) = 1$ if and only if P is cyclotomic.

D H Lehmer posed the following question:

> *Does there exist $C > 0$ such that $M(P) > 1 + C$ for all noncyclotomic (irreducible) polynomials P?*

This is expected to have an affirmative answer and, indeed, Lehmer's calculations indicate that the smallest possible value of $M(P) \neq 1$ is $1.176280821\ldots$, which occurs for the polynomial

$$P(X) = X^{10} + X^9 - X^7 - X^6 - X^5 - X^4 - X^3 + X + 1.$$

Lehmer's question can be formulated in terms of discrete subgroups of Lie groups. One may not be able to predict when it can be answered but it is more or less certain that one will need tools involving deep mathematics.

Suggested Reading

[1] B Sury. Cyclotomy. *Resonance*, Vol. 4, No. 12. pp 41–53. 1999.
[2] Polya and Szego. *Problems in Analysis*. I & II, Springer-Verlag. 1945.

B Sury
Statistics and Mathematics Unit
Indian Statistical Institute
Bangalore 560 059

13

Prime Representing Quadratics

N V Tejaswi

This chapter gives a proof of the result contained in the remark by R Tandon in Chapter 9. In fact, the converse of that statement is also true. This result was proved by Rabinowitz and Frobenius around 1912. A much simpler proof was given by Ayoub and Chowla in the *Journal of Number Theory* 13, 443–445 (1981). We believe the proof given here is new and is simpler than any of the available ones[1].

We use the notation contained in Chapter 9 with the additional observation that equivalent forms represent the same set of numbers. Let p be a prime with $p \equiv 3 \pmod 4$ and $n = (p+1)/4$. We have:

THEOREM 1. The class number of forms with discriminant $-p$ is 1, i.e., $h(-p) = 1$, if and only if for each x, $0 \leq x < n-1$, $x^2 + x + n$ is a prime number.

PROOF. Suppose there exists an integer b, $0 \leq b < n - 1 = (p-3)/4$ such that $b^2 + b + n$ is not a prime. Then there is a prime q such that

$$b^2 + b + n = aq,$$

with $q^2 \leq b^2 + b + (p+1)/4$. We have

$$4q^2 \leq (2b+1)^2 + p < \left(\frac{2(p-3)}{4} + 1\right)^2 + p = \left(\frac{p+1}{2}\right)^2,$$

i.e.,

$$q < \frac{p+1}{4},$$

and

$$4aq = (2b+1)^2 + p.$$

[1] After this article had been submitted to *Resonance*, I came to know that Frobenius proof closely resembles this proof.

Consider the quadratic forms

$$f(x,y) = x^2 + xy + ny^2 \quad \text{and} \quad g(x,y) = ax^2 + (2b+1)xy + qy^2.$$

Both have discriminant equal to $-p$. Since the class number is 1, both these forms should be equivalent, and hence should represent the same set of integers. Clearly, q is representable by $g(x,y)$, (take $x = 0$, $y = 1$). But q is not representable by $f(x,y)$. This follows from $y \neq 0$, for, if $y = 0$ then q would be a square, and for $y \neq 0$ we have

$$f(x,y) = \frac{1}{4}((2x+y)^2 + py^2) > \frac{p}{4},$$

while $q < (p+1)/4$.

For the converse, suppose that $h(-p) \geq 2$; note that $p > 7$ since $h(-3) = h(-7) = 1$. Then there exists a reduced form

$$g(x,y) = ax^2 + bxy + cy^2$$

with discriminant $-p$ which is not equivalent to the (reduced) form

$$f(x,y) = x^2 + xy + ny^2.$$

From the definition of reduced quadratic forms we have that $a, c > 1$ and $|b| \leq a \leq \sqrt{p/3}$. Further note that b is odd and hence $b = 2b' + 1$ for some integer b'. Clearly, $b' < n - 1$ (as $p > 7$) and we have

$$b'^2 + b' + n = ac,$$

which shows that for $x = b'$, $x^2 + x + n$ is not a prime number, thereby proving the converse.

The following problem appeared in the 26th International Mathematical Olympiad in 1986.

PROBLEM. Let n be a natural number. If $k^2 + k + n$ is a prime number for $0 \leq k \leq [\sqrt{n/3}]$ show that $k^2 + k + n$ is a prime for $0 \leq k \leq n - 2$.

In view of this we can restate the above theorem as

THEOREM 1′. The class number of forms with discriminant $-p$ is 1, i.e., $h(-p) = 1$, if and only if for each x, $0 \leq x \leq [\sqrt{n/3}]$, $x^2 + x + n$ is a prime number.

Acknowledgement: I thank C S Yogananda for suggesting the problem and for pointing out some errors in the first draft of this chapter.

N V TEJASWI
National Undergraduate Programme in
Mathematics and Computer Science
Chennai Mathematical Institute
92, G N Chetty Road
T. Nagar
Chennai 600 017

14

The Congruent Number Problem

V Chandrasekar

In Mathematics, especially number theory, one often comes across problems which arise naturally and are easy to pose, but whose solutions require very sophisticated methods. What is known as 'The Congruent Number Problem' is one such. Its statement is very simple and the problem dates back to antiquity, but it was only recently that a breakthrough was made, thanks to current developments in the Arithmetic of elliptic curves, an area of intense research in number theory.

Introduction

A positive integer n is called a *congruent number* if there exists a right-angled triangle whose sides are rational numbers and whose area is the given number n.

If we represent the sides of such a triangle by X, Y, Z, with Z as the hypotenuse, then by our definition, a positive integer n is a congruent number if and only if the two equations

$$X^2 + Y^2 = Z^2, \quad \frac{XY}{2} = n$$

have a solution with X, Y, Z all rational numbers.

EXAMPLES.

1. Consider the right-angled triangle with sides $X = 3$, $Y = 4$ and $Z = 5$. Its area n is $XY/2 = 6$, so 6 is a congruent number. Here we are lucky to find a suitable triangle for the number 6 whose sides are actually integers. It will be seen that this is in general an exceptional circumstance.
2. Consider the triangle with sides 3/2, 20/3 and 41/6. This is a right-angled triangle (!) and its area is 5. Therefore 5 is a congruent number.

QUESTION. Does there exist a right-angled triangle with integral sides and area equal to 5?

QUESTION. Is 1 a congruent number? (There is a lot of history behind this which will be narrated below.)

One can generate congruent numbers at will by making use of the identity

$$(X^2 - Y^2)^2 + (2XY)^2 = (X^2 + Y^2)^2$$

which corresponds to the right-angled triangle with sides $X^2 - Y^2$, $2XY$ and hypotenuse $X^2 + Y^2$. We substitute our choice of integer values for X and Y and obtain the congruent number $n = XY(X^2 - Y^2)$. For example, $X = 3$, $Y = 2$ yields the triangle with sides 5, 12, 13 and area 30. So 30 is a congruent number. For more examples refer to Box 14.1.

Now any positive integer n can be written as $n = u^2v$, where v has no square factors (v is a 'squarefree integer'). It is clear that n is a congruent number if and only if v is so; the right-angled triangle for v can be obtained from the corresponding one for n, if it exists, by scaling it down by a factor of u. (Remember that we allow the side lengths to take fractional values!) So when deciding whether n is congruent or not, we may assume that n is a squarefree integer. This will be done in what follows.

Box 14.1 Generating Congruent Numbers

Here p, q are arbitrary positive integers of opposite parity (that is, $p + q$ is odd), the congruent number n is the squarefree part of $pq(p^2 - q^2)$, and the sides of the triangle are proportional to $p^2 - q^2$, $2pq$, $p^2 + q^2$.

Serial number	p	q	n	Sides of the triangle
1	3	2	30	5, 12, 13
2	4	3	21	7/2, 12, 25/2
3	5	4	5	3/2, 20/3, 41/6
4	9	4	65	65/6, 12, 97/6
5	25	16	41	40/3, 123/20, 881/6

Now we are ready to formulate:

THE CONGRUENT NUMBER PROBLEM. Given a positive integer n, is there a simple criterion which enables us to decide whether or not n is congruent?

A few remarks are in order. To start with, if we restrict the sides of the triangle to integer values only, the question can be settled, at least in theory, in a finite number of steps. To see why, recall the equations

$$X^2 + Y^2 = Z^2, \quad \frac{XY}{2} = n.$$

Since X and Y are now integers, X and Y both divide $2n$. So to see if a solution exists, we let X run through the set of divisors of $2n$, let $Y = 2n/X$ and check whether $X^2 + Y^2$ is a square integer. Thus the problem can be settled in a routine manner. For example, we can easily verify that there is no integral solution for the case $n = 5$. (Note, however, that we do know that 5 is a congruent number.)

But once we allow the sides to have rational values, the problem acquires an entirely different status. There is no obvious starting point, unlike the case of integer solutions discussed above. One could endlessly churn out congruent numbers following the method in Box 14.1 without being certain when a given number n (or $n \times m^2$, for some integer m) will appear on the list. Continuing in this way would exhaust one's computing resources, not to mention one's patience! Also, this procedure is of no avail if n is not a congruent number.

To appreciate this better, consider the following right-angled triangle with area 101 which was found by Bastien in 1914. This triangle has sides

$$X = \frac{711024064578955010000}{118171431852779451900},$$
$$Y = \frac{3967272806033495003922}{118171431852779451900},$$

and hypotenuse

$$Z = \frac{2 \times 2015242462949760001961}{118171431852779451900}.$$

This is known to be the smallest solution (in terms of the sizes of the numerator and denominator) corresponding to the congruent number 101! The serial number of this triangle in the list in Box 14.1 would exceed 10^{20}!

The above considerations force us to look for a more indirect approach in our search for a criterion for characterizing congruent numbers.

Here we have yet another instance of a problem in number theory which is simple to state, yet has hidden depths. There have been instances when the solutions of such problems have emerged only centuries after being posed. In such instances, a lot of deep and beautiful mathematics gets generated as a result. A striking example from recent times is the proof of Fermat's last theorem by Andrew Wiles in 1995, which uses a mind-boggling variety of techniques from several fields in current mathematical research.

We shall see how the congruent number problem falls into this category by giving a brief account of its history and the concepts and techniques that were used in the solution of this problem which is deceptively so simple to state.

Brief History

The congruent number problem makes its earliest appearance in an Arab manuscript traced to the tenth century (*c* 972 AD). In his classic *History of the Theory of Numbers*, Vol 2 (Diophantine Analysis), Dickson quotes Woepeck's view that there is no indication that the Arabs knew Diophantus prior to the translation by Aboul Wafi (998 AD), but they may well have come across the problem from the Hindus who were already acquainted with his work. The Arabs figured out that the following numbers are congruent: 5, 6, 14, 15, 21, 30, 34, 65, 70, 110, 154, 190 and so on. In fact, their list contains ten congruent numbers greater than 100, for example, 10374.

The scene later shifts to Pisa, where Leonardo Pisano (better known as Fibonacci), by virtue of his position as a mathematical expert in his native city, is presented to the Emperor Frederic II. The king's scholars challenge him to find three rational numbers whose squares form an arithmetic progression with common difference 5. This is equivalent to finding integers X, Y, Z, T, with $T \neq 0$, such that $Y^2 - X^2 = Z^2 - Y^2 = 5T^2$, and this in turn reduces to finding a right triangle with rational sides

$$\frac{Z+X}{T}, \quad \frac{Z-X}{T}, \quad \frac{2Y}{T},$$

and area 5; in other words, to the question of whether 5 is a congruent number or not. Leonardo addressed the general problem in his memoir *Liber Quadratorum* (1225), which was lost to the world till it was found and published by Prince Boncompaign in the year 1856. In addition to showing that 5 and 7 are congruent numbers (the triangles have sides $3/2, 20/3, 41/6$ and $35/12, 24/5, 337/60$ respectively), he also states without proof that no congruent number can be a square, or equivalently that 1 is not a congruent number.

The proof of this statement had to wait for four centuries. Eventually it led to Fermat's discovery of his method of infinite descent, which was to have a profound effect on subsequent developments in arithmetic, or number theory as we now call it.

Fermat had been in correspondence with many of his contemporaries regarding the existence of a right-angled triangle with rational sides and a square area. An explicit reference to the application of his technique to prove that this is impossible appears in his letter to Huygens in 1659, where he states: "*As ordinary methods, such as are found in the books, are inadequate to proving such difficult propositions, I discovered at last a most singular method ... which I call the infinite descent. At first I used it only to prove negative assertions such as ... "there is no right angled triangle in numbers whose area is a square". To apply it to affirmative questions is much harder, so that, when I had to prove that "Every prime of the form* $4n+1$ *is a sum of two squares", I found myself in a sorry plight. But at last such questions proved amenable to my method.*" (We infer that the technique of infinite descent had its first application in number theory to the problem of congruent numbers.) Continuing, Fermat gives a cryptic description of his method: "*If the area of such a triangle were a square, then there would also be a smaller one with the same property, and so on, which is impossible, ...*". He adds that to explain how his method works would make his discourse too long, as the whole mystery of his method lay there. To quote Weil, "*Fortunately, just for once, he (Fermat) had found room for this mystery in the margin of the very last proposition of Diophantus*".

Before reproducing Fermat's proof we prove the following:

PROPOSITION 1. Let X, Y, Z be the sides of an integer-sided right-angled triangle, with Z the hypotenuse, such that X, Y, Z have no common factors. Then there exist, relatively prime integers p, q such that $p+q$ is odd, $\{X, Y\} = \{p^2 - q^2, 2pq\}$ and $Z = p^2 + q^2$.

PROOF. Clearly X and Y cannot be both even, as they have no common factors. Both cannot be odd, for in this case both X^2 and Y^2 would be 1 modulo 4, implying that $Z^2 \equiv 2 \pmod 4$; but this is absurd as no square is of the form $2 \pmod 4$. Thus one of them, say X, is odd and the other, Y, is even. It follows that Z is odd and that

$Z + X, Z - X$ are both even. Therefore $(Z + X)/2$ and $(Z - X)/2$ are integers; indeed they are coprime, because X and Z are themselves coprime.

Since $Y^2 = Z^2 - X^2$, we obtain:

$$\left(\frac{Y}{2}\right)^2 = \frac{Z+X}{2} \cdot \frac{Z-X}{2}.$$

By the unique factorization property of the integers, each factor on the right-side must be a square. Thus $(Z + X)/2 = p^2$, $(Z - X)/2 = q^2$ with p and q coprime. Solving, we obtain

$$X = p^2 - q^2, \quad Y = 2pq, \quad Z = p^2 + q^2.$$

Since X is odd, $p + q$ is odd.

Fermat's Legacy

We now reproduce Fermat's proof by the method of descent in the following:

THEOREM. 1 is not a congruent number.

PROOF. Suppose, on the contrary, that 1 is a congruent number; i.e., there exists a right-angled triangle with integral sides whose area is a square integer. In view of Proposition 1, its sides must be of the form $2pq, p^2 - q^2, p^2 + q^2$ with $p > q > 0$, $p + q$ odd and $(p, q) = 1$.

Since the area ($= pq(p - q)(p + q)$) is a square integer and the numbers p, q, $p - q$, $p + q$ are mutually coprime, it follows that *each* of these numbers is a square integer. We write

$$p = x^2, \quad q = y^2, \quad p + q = u^2, \quad p - q = v^2.$$

Since u and v are odd and coprime, it follows that the gcd of $u + v$ and $u - v$ is 2. But now we have

$$2y^2 = 2q = u^2 - v^2 = (u + v)(u - v).$$

Arguing as in Proposition 1, we see that there exist integers r, s such that $(u + v, u-v) = (2r^2, 4s^2)$ or $(u+v, u-v) = (4r^2, 2s^2)$. The former case leads to $u = r^2+2s^2$, $v = r^2 - 2s^2$ and therefore to

$$x^2 = \frac{u^2 + v^2}{2} = r^4 + 4s^4.$$

Hence $r^2, 2s^2, x$ are the sides of a right-angled triangle with area $(rs)^2$ and hypotenuse $x = \sqrt{p} < p^2 + q^2$ (the hypotenuse of the triangle with which we started). The case $u + v = 4r^2, u - v = 2s^2$ is dealt with in a similar fashion.

So, starting from a right-angled triangle with integral sides whose area is a square integer, we have produced another triangle of the same type with a smaller hypotenuse than the original triangle. Clearly this process can be repeated. But this gives rise to an infinite decreasing sequence of positive integers—a clear absurdity. (This is the central principle behind infinite descent.) We are thus led to a contradiction and we conclude that 1 is not a congruent number.

The non-congruent nature of the number 1 is of special interest because it shows that there is no non-trivial solution to the equation $X^4 - Y^4 = Z^2$, which in turn implies Fermat's last theorem ('The equation $X^n + Y^n = Z^n$ has no non-trivial solutions in integers for $n > 2$') for the case $n = 4$!

In the following two propositions we prove the claims made above.

PROPOSITION 2. A number n is congruent if and only if there exists a rational number a such that $a^2 + n$ and $a^2 - n$ are both squares of rational numbers.

PROOF. Let n be a congruent number and let X, Y, Z be rational numbers satisfying

$$X^2 + Y^2 = Z^2, \quad \frac{XY}{2} = n.$$

Then $X^2 + Y^2 \pm 2XY = Z^2 \pm 4n$, so

$$\left(\frac{X \pm Y}{2}\right)^2 = \left(\frac{Z}{2}\right)^2 \pm n.$$

So if we take $a = Z/2$, then a is rational and $a^2 + n$ and $a^2 - n$ are both squares of rational numbers.

For the converse, let a be a rational number such that $a^2 + n$ and $a^2 - n$ are squares of rational numbers. Let

$$X = \sqrt{a^2 + n} + \sqrt{a^2 - n}, \quad Y = \sqrt{a^2 + n} + \sqrt{a^2 - n},$$

and

$$Z = \sqrt{X^2 + Y^2} = \sqrt{4a^2} = 2a.$$

Then X, Y, Z are the sides of a right-angled triangle with rational sides and area $XY/2 = ((a^2 + n) - (a^2 - n))/2 = n$.

PROPOSITION 3. If there are non-zero integers X, Y, Z such that $X^4 - Y^4 = Z^2$, then 1 is a congruent number.

PROOF. Write the equation in the form

$$X^4 = Y^4 + Z^2.$$

Using Proposition 1, we deduce that there exist integers p, q such that $X^2 = p^2 + q^2$ and $Y = p^2 - q^2$. But this leads to

$$\frac{p^2}{q^2} + 1 = \left(\frac{X}{q}\right)^2, \quad \frac{p^2}{q^2} - 1 = \left(\frac{Y}{q}\right)^2.$$

So p^2/q^2 is a rational number such that $p^2/q^2 + 1$ and $p^2/q^2 - 1$ are squares of rational numbers. In other words 1 is a congruent number.

Combining Propositions 2 and 3 with the fact that 1 is a non-congruent number, we deduce Fermat's last theorem for $n = 4$.

Before closing this section, it is fitting to quote Weil's lavish praise of Fermat and his justly-famous method: "*The true breakthrough came in 1922 with Mordell's celebrated paper; here, if Fermat's name does not occur, the use of the words "infinite descent" shows that Mordell was well aware of his indebtedness to his remote predecessor. Since then the theory of elliptic curves, and its generalizations to curves of higher genus and to abelian varieties, has been one of the main topics of modern number theory. Fermat's name, and his method of infinite descent, are indissolubly bound with it; they promise to remain so in the future*".

Congruent Numbers and Elliptic Curves

Congruent numbers continued to excite the curiosity of number theorists over the years. Their congruence properties have been investigated and tables of such numbers constructed. Some classes of numbers have also been identified as congruent numbers. To cite an example, a result due to Heegner and Birch shows that if n is a prime number of the form 5 (mod 8) or of the form 7 (mod 8) then n is a congruent number. (See Box 14.2)

But what is ultimately sought is a simple and complete characterization of all congruent numbers; in other words, an algorithm which will quickly determine whether a given natural number n is congruent or not.

Box 14.2 Some Classes of Congruent Numbers

This box displays some results given in the paper by K Feng [5]. It characterises some classes of congruent and non-congruent numbers in terms of their divisibility properties.

To illustrate, Gross's result states that if an integer n is squarefree and has at most two prime factors of the form 5, 6 or 7 (mod 8), then n is a congruent number.

If p and q are odd primes, then the Legendre symbol (p/q) is 1 if p is a quadratic residue modulo q (that is, if the equation $x^2 \equiv p \pmod{q}$ has a solution), else -1.

In the following account, n is taken to be a squarefree integer. The symbol '**CN**' means 'congruent number', while '**Non-CN**' means 'non-congruent number'. p, q, r denote distinct primes and p_i refers to an arbitrary prime congruent to i mod 8.

For **CN**

- $n = 2p_3$ (Heegner 1952, Birch 1968).
- $n = p_5, p_7$ (Stevens 1975).
- $n = p^u q^v \equiv 5, 6, 7 \pmod{8}$, $0 \le u, v \le 1$ (B Gross 1985).
- $n = 2p_3p_5, 2p_5p_7$.

Contd...

Contd...

- $n = 2p_1p_7$, with $(p_1/p_7) = -1$ (Monsky 1990).
- $n = 2p_1p_3$, with $(p_1/p_3) = -1$.

For **Non-CN**

- $n = p_3, p_3q_3, 2p_5, 2p_5q_5$ (Genocchi 1855).
- $n = 2p$, with $p \equiv 9 \pmod{16}$ (Bastien 1913).
- $n = p_1p_3$, with $(p_1/p_3) = -1$ (Lagrange 1974).
- $n = 2p_1p_5$, with $(p_1/p_5) = -1$.
- $n = p_1p_3q_1$, with $(p_1/p_3) = (p_3/q_1) = -1$.

As it happened, the search for such an algorithm was made possible by relating the congruent number problem to the arithmetic of elliptic curves.

This connection is established as follows. From Proposition 2 we know that a number n is congruent means there exists a rational square, say u^2 such that $u^2 + n$ and $u^2 - n$ are both rational squares. This implies that $u^4 - n^2$ is a rational square, say v^2; or equivalently that $u^6 - n^2u^2 = u^2v^2$. Setting $x = u^2$ and $y = uv$ we arrive at the equation $y^2 = x^3 - n^2x$. Thus if n is a congruent number, we obtain a rational point (x, y) on the curve represented by the equation $y^2 = x^3 - n^2x$.

Now the curves corresponding to the equation $y^2 = x^3 - n^2x$ are examples of what are known as elliptic curves. The arithmetic of these curves has been a central topic of research in Number Theory over the years. In view of the above connection, it was natural to expect that the results relating to elliptic curves would be able to settle the congruent number problem. This expectation was realized when J Tunnell succeeded in finding a simple algorithm for the problem. (See Box 14.3 for a brief outline of the logical steps involved in Tunnell's method.)

Let the reader be reassured that to apply the algorithm one does not need to know anything about elliptic curves, modular forms, liftings or L-functions which are (to name a few) some of the concepts and techniques which lie at the basis of Tunnell's work!

In what follows, $\#S$ denotes the number of elements of a set S.

Tunnell's Theorem (1983). Let n be a squarefree congruent number (that is, n is the area of a right-angled triangle with rational sides). Define A_n, B_n, C_n, D_n as follows:

$$\begin{aligned}
A_n &= \#\{(x, y, z) \in \mathbf{Z}^3 \mid n = 2x^2 + y^2 + 32z^2\},\\
B_n &= \#\{(x, y, z) \in \mathbf{Z}^3 \mid n = 2x^2 + y^2 + 8z^2\},\\
C_n &= \#\{(x, y, z) \in \mathbf{Z}^3 \mid n = 8x^2 + 2y^2 + 64z^2\},\\
D_n &= \#\{(x, y, z) \in \mathbf{Z}^3 \mid n = 8x^2 + 2y^2 + 16z^2\}.
\end{aligned}$$

Then:

(A) $A_n = B_n/2$ *if n is odd; and*

(B) $C_n = D_n/2$ *if n is even.*

If the Birch–Swinnerton Dyer conjecture is true, then, conversely, these equalities imply that n is a congruent number.

Box 14.3 Elliptic Curves and the Congruent Number Problem

For each natural number n, let E_n denote the elliptic curve represented by the equation $y^2 = x^3 - n^2x$. Then we have the following correspondence between the set of right-angled triangles with rational sides and area n and the set of rational points on E_n. Let the sides be A, B, C where A, B, C are rational and $A < B < C$, and let (x, y) be a rational point on E_n such that: (a) x is the square of a rational number, (b) the denominator of x is even, (c) the numerator of x has no common factor with n. The correspondence is given as follows:

$$(x, \pm y) \longrightarrow \left(\sqrt{x+n} - \sqrt{x-n},\ \sqrt{x+n} + \sqrt{x-n},\ 2\sqrt{x}\right),$$

$$(A, B, C) \longrightarrow \left(\frac{C^2}{4}, \pm\frac{(B^2 - A^2)C}{8}\right).$$

It can be shown by means of the above bijection that a number n is congruent if and only if there exist infinitely many rational solutions (x, y) on the elliptic curve E_n.

To each elliptic curve E_n, there is associated an important number $L(E_n)$, which we shall not attempt to define. It is known (this is the Coates–Wiles Theorem) that if E_n has infinitely many rational solutions, then $L(E_n) = 0$. Combining this with the remark in the previous paragraph we deduce the following: *If $L(E_n)$ is not zero, then n is a non-congruent number.*

The converse statement, namely that $L(E_n) = 0$ implies the existence of infinitely many rational points on E_n (in other words, that $L(E_n) = 0$ implies that n is congruent) would follow from a famous conjecture due to Birch and Swinnerton–Dyer. (This conjecture has been made for all elliptic curves, not just for the E_n defined above.)

Now Tunnell's work can be summarized in one line; he has found an expression for $L(E_n)$ which is of the form

$$L(E_n) = \begin{cases} C \times (A_n - B_n/2), & \text{if} \quad n \text{ is odd,} \\ C \times (C_n - D_n/2), & \text{if} \quad n \text{ is even.} \end{cases}$$

Here C is a non-zero number, and A_n, B_n, C_n, D_n are the quantities defined in the statement of Tunnell's theorem.

The justification of Tunnell's algorithm follows from the above mentioned facts.

Observe that Tunnell's algorithm helps one to establish whether a given number n is non-congruent.

EXAMPLES.

1. Let $n = 1$; then $A_n = B_n = 2$, so equation (A) is not valid. We conclude that 1 is not a congruent number.
2. We show similarly that 2 and 3 are not congruent numbers.
3. Let n be squarefree, odd and congruent to 5 or 7 modulo 8. Since $2x^2 + y^2$ can never be congruent to 5 or 7 modulo 8, both cardinalities in (A) are 0 and hence the condition is satisfied. If the Birch–Swinnerton Dyer conjecture were true, we would be able to conclude that any such n is a congruent number. (There is supportive argument for this statement from the tables and the vanishing of the so called L-value of the corresponding elliptic curve.)

In particular, 157 would be a congruent number. This is in fact true. A proof of this fact is furnished by the right-angled triangle whose sides x, y, z, displayed below, were computed by Don Zagier. Again, this is the smallest solution for the area 157! The sides are X, Y where

$$X = \frac{6803298487826435051217540}{411340519227716149383203},$$
$$Y = \frac{411340519227716149383203}{21666555693714761309610},$$

and the hypotenuse is Z where

$$Z = \frac{224403517704336969924557513090674863160948472041}{8912332268928859588025535178967165700164808 30}.$$

A natural question on the part of the reader would concern the appropriateness of the word 'congruent' in the definition of congruent number. As to that, one cannot do better than to quote Richard Guy: "*Congruent Numbers are perhaps confusingly named*".

But, after all, what's there in a name?

Suggested Reading

[1] R K Guy. *Unsolved Problems in Number Theory*. Springer-Verlag, 1981.

[2] N Koblitz. *Introduction to Elliptic Curves and Modular Forms*. Springer-Verlag, 1984.

[3] J Tunnell. A classical Diophantine problem and modular forms of weight 3/2. *Inventiones Math*. 72. 323–33, 1983.

[4] A Weil. *Number Theory: An Approach Through History*. Birkhäuser, 1984.

[5] K Feng. Non-congruent numbers, odd graphs and the B-S-D conjecture. *Acta Arithmetica*; LXXV 1, 1996.

V CHANDRASEKAR
C/o Mr Sripathy
Spic Mathematics Institute
92, East-Coast Chambers
T. Nagar
Chennai 600 017

15

Fermat's Last Theorem

A Theorem at Last!

C S Yogananda

After more than three centuries of effort by some of the best mathematicians, Gerhard Frey, J-P Serre, Ken Ribet and Andrew Wiles have finally succeeded in proving Fermat's assertion that the equation $X^n + Y^n = Z^n$ has no solutions in non-zero integers if $n \geq 3$. Each of the four mathematicians made a decisive contribution, with Wiles delivering the *coup de grace*. The proof, as it finally came to be, is in some sense a triumph for Fermat.

When Pierre de Fermat died in 1665, he had not published a single mathematical work (except for an anonymous appendix to a book written by a colleague). His mathematical discoveries were contained in his correspondence with other mathematicians of his time, notably, Pascal, Frénicle de Bessy and Father Mersenne. He also left behind a few unpublished manuscripts and marginal notes in the books he studied. We have to be grateful to his son Samuel for whatever we know of Fermat's work. Samuel de Fermat went through his father's papers and books in addition to soliciting letters written by his father from his correspondents in order to publish them. Among Fermat's possessions was a copy of the Latin translation, by Bachet, of Diophantus' *Arithmetic* in which Fermat had made a number of marginal notes.

The first work Samuel chose to publish, in 1670, was a new edition of Bachet's Diophantus with an appendix containing forty eight marginal notes made by Fermat. The second of these notes appears alongside problem 8 in Book II of *Arithmetic: "... given a number which is square, write it as a sum of two other squares"*. Fermat's note states, in Latin, that *"on the other hand, it is impossible for a cube to be written as a sum of two cubes or a fourth power to be written as sum of two fourth powers or, in general, for any number which is a power greater than the second to be written as a sum of two like powers. I have a truly marvellous demonstration of this proposition which this margin is too narrow to contain"*. Thus, it was in 1670 that the world learnt of what has come to be termed as Fermat's Last Theorem (FLT): The equation

$$X^n + Y^n = Z^n$$

has no solutions in non-zero integers if $n \geq 3$. Fermat himself had given a proof of this assertion for $n = 4$ using *infinite descent*, a method he invented, and Euler proved the case, $n = 3$. Thus, to prove FLT we need to show that $X^p + Y^p = Z^p$ has no solutions in non-zero integers whenever p is a prime greater than 3 (do you see why?).

After more than three centuries of effort by some of the best mathematicians, Gerhard Frey, J-P Serre, Ken Ribet and Andrew Wiles have finally succeeded in proving Fermat's assertion, each of them making a decisive contribution, with Wiles delivering the *coup de grace*. The proof, as it finally came to be, is in some sense a triumph for Fermat. *Elliptic curves and infinite descent* play significant roles; it was Fermat who pioneered the use of elliptic curves in solving diophantine equations, and it is to him that we owe the method of infinite descent.

Diophantine Equations

The chief work of Diophantus of Alexandria (c. 250 A.D) known to us is the *Arithmetic*, a treatise in thirteen books, or *Elements*, of which only the first six have survived. This work consists of about 150 problems, each of which asks for the solution of a given set of algebraic equations in positive rational numbers, and so equations for which we seek integer (or rational) solutions are referred to as diophantine equations. The most familiar example we know is $X^2 + Y^2 = Z^2$ whose solutions are *Pythagorean triples*; (3, 4, 5), (5, 12, 13) are examples of such triples. If, instead, we ask for solutions in integers of $X^2 + Y^2 = 3Z^2$, we get an example of a diophantine equation for which there are no solutions in non-zero integers. (To see this, first observe that we may assume X, Y, Z to be pairwise relatively prime, by cancelling common factors, if any; and that any square when divided by 3 leaves remainder 0 or 1.) In fact, it is an interesting exercise to characterize the set of natural numbers m for which the equation $X^2 + Y^2 = mZ^2$ has no solutions in non-zero integers.

To understand the role of *geometry* in solving diophantine equations, let us consider the equation $X^2 + Y^2 = Z^2$. How do we characterize all solutions (in integers) of this equation? We could assume again that X, Y, Z is a *primitive* solution, i.e., X, Y, Z are pairwise relatively prime. Dividing by Z^2 and putting $X/Z = x$ and $Y/Z = y$ we get $x^2 + y^2 = 1$, that is to say, we get a *rational point* (a point both of whose coordinates are rational numbers), (x, y), on the unit circle centered at the origin. Conversely, a rational point on the circle $x^2 + y^2 = 1$ will give us a (primitive) Pythagorean triple. So, our problem reduces to finding all rational points on the unit circle. We do this by drawing a line with rational slope passing through the point $(-1, 0)$. This line will meet the circle at one more point and we claim that this point is also rational. I shall leave it to you to figure out why it is so. (You need to use the fact that if one root of a quadratic equation with rational coefficients is rational then the other root is also rational.) This way we obtain all rational points on the circle. Put $t = \tan\theta/2$ in the familiar parametrisation of the circle, $(\cos\theta, \sin\theta)$. Then we get the well-known characterisation of the Pythagorean triples: if m and n, $m > n$, are integers of opposite parity then the numbers

$$m^2 - n^2, \quad 2mn, \quad m^2 + n^2$$

History of FLT

- 1640, Fermat himself proved the case $n = 4$
- 1770, Euler proved the case $n = 3$; (Gauss also gave a proof).
- 1823, Sophie Germain proved the *first case* of FLT — first case of FLT holds if there is no solution for $X^p + Y^p = Z^p$ for which p does not divide the product XYZ — for a class of primes, *Sophie Germain primes* — primes p such that $2p + 1$ is also a prime.
- 1825, Dirichlet, Legendre proved FLT for $n = 5$.
- 1832, Dirichlet treated successfully the case $n = 14$.
- 1839, Lamé proved the case $n = 7$.
- 1847, Kummer proved FLT in the case when the exponent is a *regular prime*. But it is not known, even today, whether there are infinitely many Sophie Germain primes or regular primes.
- 1983, Faltings gave a proof of Mordell's conjecture.
- 1986, Frey–Ribet–Serre: Shimura–Taniyama–Weil conjecture implies FLT.
- 1994, Andrew Wiles: proof of S–T–W conjecture for semistable elliptic curves.

form a primitive Pythagorean triple and every primitive Pythagorean triple arises this way.

This method can be used to find all rational points on a conic section whose defining equation has rational coefficients, once we are able to find one such point.

Elliptic Curves

Consider the following classical problems.

(i) Find all n such that the sum of the squares of the first n natural numbers is a square. That is, we have to find natural numbers n and m such that

$$m^2 = n(n+1)(2n+1)/6.$$

(ii) (Diophantus) Find three rational right triangles of equal area.
Let A denote the area of the right triangle with sides $a\,(= p^2 - q^2)$, $b\,(= 2pq)$ and $c\,(= p^2 + q^2)$; thus $A = pq\,(p^2 - q^2)$. Then if we put $x = p/q$ we get a rational point $(p/q,\, 1/q^2)$ on the curve

$$Ay^2 = x^3 - x.$$

Conversely, if $(a/b, c/d)$ is a rational point on this curve then the right triangle with $d(a^2 - b^2)/b^2c$ and $2ad/bc$ as legs also has area equal to A.

(iii) (From an Arab manuscript dated before the 9th century) Given a natural number n, find a rational number u, such that both $u^2 + n$ and $u^2 - n$ are squares (of rational numbers).

What is *elliptic* about elliptic curves?

Ellipses are not elliptic curves! Elliptic curves are so called because it was in connection with the problem of computing arc lengths of ellipses that they were first studied systematically. When we compute the arc of a circle, we have to integrate the function $1/\sqrt{(1-x^2)}$, which we do in terms of sine and cos functions. The trignometric functions are therefore called *circular functions*. Similarly, to compute the arc length of an ellipse, we have to integrate functions of the form

$$1/\sqrt{[(1-x^2)(1-k^2x^2)]}.$$

This integral cannot be computed using circular functions and mathematicians worked on this problem for many years before Abel and Jacobi, independently introduced *elliptic functions* to compute such integrals. Just as sin and cos satisfy $x^2 + y^2 = 1$, the elliptic functions satisfy an equation of the form $y^2 = f(x)$ where $f(x)$ is a cubic.

If such a u can be found then n is called a congruent number. A number n being congruent is equivalent to the existence of a right triangle with rational sides and area n (see Chapter 12).

Let n be a congruent number and let u be such that $u^2 + n = a^2$ and $u^2 - n = b^2$. Multiplying the two equations together we get

$$u^4 - n^2 = (ab)^2.$$

Multiplying by u^2 throughout to get

$$u^6 - n^2u^2 = (abu)^2.$$

Putting $u^2 = x$ and $abu = y$ we get a rational point on the curve, E, defined by the equation

$$y^2 = x^3 - n^2x.$$

EXERCISE. Conversely, if (x, y) is a rational point on E such that x is a rational square and has an even denominator, then n (whose square appears as the coefficient of x) is a congruent number.

In each of the above problems, we were led to consider equations of the form $y^2 = f(x)$, where $f(x)$ is a cubic polynomial in x with rational coefficients and distinct roots. Such equations define *elliptic curves*. We could think of elliptic curves as the set of all rational/real/complex solutions of such equations. The set of all complex solutions of an elliptic curve can be identified with the points on a *torus*. The figures below (Figure 15.1) show what the real and complex points on an elliptic curve look like.

Finding rational points on an elliptic curve turns out to be a difficult problem and though many deep results have been proved (one of them by Andrew Wiles

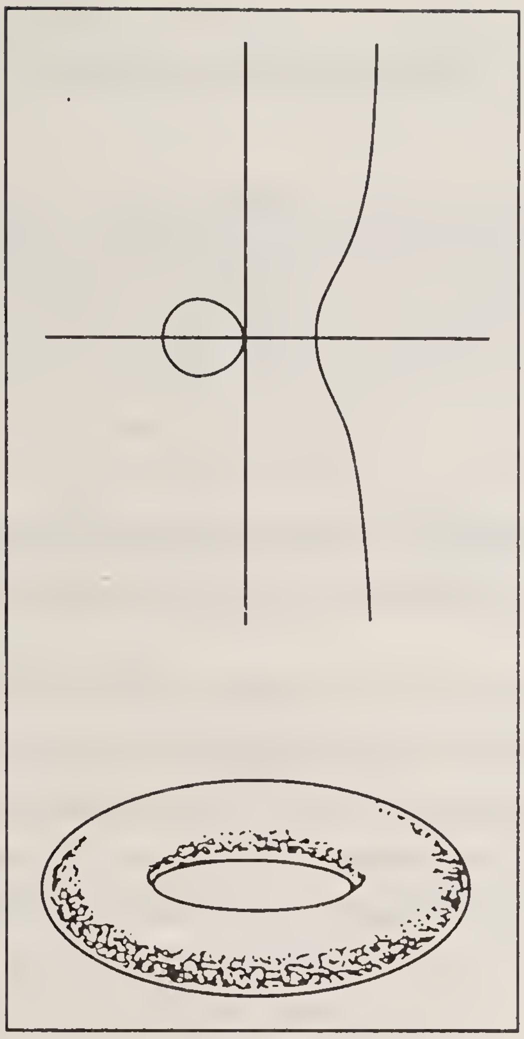


Figure 15.1 Typical illustration depicting how the real/complex points on an elliptic curve look like.

along with John Coates), a lot remains to be done in this area. The study of elliptic curves is currently a very active field of research involving many different areas of mathematics.

If we try to imitate the method we used for a conic to get more rational points from one such point we are stuck. This is because generally, a line meets a cubic curve at three points and we cannot conclude that the other points of intersection are rational. That is, if one root of a cubic equation with rational coefficients is rational, the other two roots could be irrational; they could be conjugate surds, for instance. What is true is that if you draw the line joining two rational points, then the third point where this line meets the cubic will also be a rational point. Thus, we can 'add' two rational points to get a third rational point. It turns out that we could take the 'point at infinity' as the identity or the 'zero' element and obtain a structure of a *group* (in fact, a commutative group) on the set of rational points of an elliptic curve by declaring the sum of three collinear points to be zero; the inverse or 'negative' of the point (x, y) is the point $(x, -y)$. Thus, to add two points P and Q join them by a straight line, find the third point of intersection of the line with the curve and reflect it in the x-axis to get a point, R, on the curve which will then be the 'sum' of P and Q.

EXERCISE. Consider the elliptic curve, E, defined by the equation $y^2 = ax^3 + bx^2 + cx + d$. Obtain an expression for the coordinates x_3, y_3 of the sum of the two points $\mathrm{P} = (x_1, y_1)$ and $\mathrm{Q} = (x_2, y_2)$ on E in terms of x_1, x_2, y_1, y_2.

Hint: If P is not equal to Q, $x_3 = -x_1 - x_2 - (b/a) + (y_2 - y_1)^2/a(x_2 - x_1)^2$ and if P = Q, $x_3 = -2x_1 - (b/a) + (f'(x_1))^2/a(2y_1)^2$ where $f(x)$ denotes the cubic.

The structure of a group on the set of rational points of an elliptic curve provides us with a powerful tool to study diophantine equations. For instance, in problem (ii), if we get one rational point then we could 'double' (i.e., draw a tangent at that point) it to get one more point and then add these two to get yet another point, and so on. In fact, this is what Fermat used to get more solutions to the problem (even Diophantus used this procedure but he gave only three rational points). In the congruent number problem, it turns out that the double of any rational point which is not of order 2 is such that the x-coordinate is a square number with an even denominator.

The method we used to show the non-existence of solutions of $X^2 + Y^2 = 3Z^2$ by showing that the equation has no solutions *modulo 3* is a standard method we use in studying diophantine equations. Assume that the equation has integer coefficients by clearing the denominators, if necessary. We *reduce* the equation modulo a prime p by replacing the coefficients of the equation by their remainders when divided by p and consider the set of solutions of the reduced equation in the *finite field* $\{0, 1, 2, \ldots, p - 1\}$. If, for example, we find a prime for which there are no solutions for the reduced equation, it follows immediately that the original equation has no rational roots.

Consider an elliptic curve E defined by $y^2 = f(x)$. Except for a finite set of primes depending on the cubic $f(x)$, the reduced equation will also define an elliptic curve. In fact, the *exceptional* set of primes is precisely the set of prime divisors of the discriminant of the cubic $f(x)$. For a prime p not dividing the discriminant, let N_p denote the number of points of E modulo p, i.e., the number of pairs (x, y), with x, y in $\{0, 1, 2, \ldots, p - 1\}$, satisfying the equation modulo p. Define integers a_p by

$$N_p = p + 1 - a_p.$$

These a_p's could be positive or negative and Hasse proved the following inequality in 1930:

$$|a_p| \leq 2\sqrt{p}.$$

These numbers contain a lot of information about the rational points of the elliptic curve and there are many conjectures concerning their properties among which the Birch–Swinnerton–Dyer conjecture and the Shimura–Taniyama–Weil conjecture are the most important.

The content of the Shimura–Taniyama–Weil (S–T–W) conjecture is that these a_p's are the *Fourier coefficients of a cusp form* (of weight 2 and a certain level N). The definition of cusp forms is beyond the scope of this chapter and we content ourselves by saying that they are certain functions on the upper half-plane (please see Suggested Reading at the end). Elliptic curves for which the a_p's satisfy the S–T–W conjecture are called *modular* elliptic curves.

Frey Elliptic Curve and Fermat's Last Theorem

The study of rational points on higher degree curves witnessed a breakthrough in 1983 when Gerd Faltings proved a conjecture of Mordell. As a corollary, it stated that the curve $X^n + Y^n = 1$ has only finitely many rational points if $n \geq 5$, which means that there would be at most finitely many solutions to the Fermat equation

$$X^n + Y^n = Z^n.$$

The general feeling among mathematicians following this was one of satisfaction since there was no reason or heuristic basis as to why FLT should be true; *at most finitely many solutions* was good enough.

But FLT bounced back soon after in 1985, when Gerhard Frey linked a counter example of FLT, if there is one, with an elliptic curve which did not seem to satisfy the S–T–W conjecture! Frey's was a simple but very ingenious idea: if, for some prime $p > 3$, there are non-zero integers u, v, w such that $u^p + v^p = w^p$, then consider the elliptic curve, now referred to as the *Frey curve*,

$$y^2 = x(x + u^p)(x - v^p).$$

Thus for the first time, FLT for *any exponent* was connected with a *cubic* curve instead of the higher degree curve which the equation itself defines.

Then things started happening fast and in the summer of 1986, building on the work of Frey and Serre, Ribet succeeded in proving that S–T–W implies FLT by showing that the Frey curve could not be modular. Now, FLT was not just a curiosity but was related to a deep conjecture; if it were not true and we had a counter example, the Frey curve would be sticking out like a sore thumb!

Soon after he heard of Ribet's result, Andrew Wiles went to work on the S–T–W conjecture in the late summer of 1986. After working hard on it for seven years, during which time even his closest friends did not get to know what he was up to, Wiles stunned the mathematical world by claiming that he had proved the FLT by proving a particular case of the S–T–W conjecture, the case of *semi-stable* elliptic curves. He made the announcement at the end of a series of lectures at the Isaac Newton Institute in Cambridge, England on the morning of Wednesday, June 23, 1993. But experts checking his proof found many gaps of which he could overcome all but one. It is to the credit of Wiles that he did not let this setback deter him. Rather, encouraged and mathematically supported by his students and close friends, notably Henri Darmon, Fred Diamond and Richard Taylor, he circumvented the gap in September 1994. His paper, along with another one of his jointly with Richard Taylor, occupies one whole issue of the leading journal *Annals of Mathematics*, **142** (1995). It should be remarked that the theorem Wiles proved has a very significant result with far-reaching consequences and FLT follows as a simple corollary.

Apparently, Fermat's favourite target for his problems and challenges were the English mathematicians; after all, he was French! Thus, it is fitting that his most famous challenge has been answered by Wiles, an Englishman, though it took a while (A Wiles!) coming!

Suggested Reading

[1] Paulo Ribenboim. *13 Lectures on Fermat's Last Theorem.* Springer-Verlag. 1979.

[2] H M Edwards. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory.* Springer-Verlag. 1977.
The above two books contain historical accounts of the various attempts to prove FLT and developments stemming from these attempts, especially the work of Kummer.

[3] J-P Serre. *A Course in Arithmetic.* Springer International Student Edition, 1979.
This extraordinary book covers in just hundred pages many important theorems in number theory (with proofs) and contains an introduction to modular forms.

[4] Neal Koblitz. *Introduction to Elliptic Curves and Modular Forms.* Springer-Verlag. 1984.
This contains a beautiful introduction to elliptic curves and modular forms via the *congruent number problem.*

C S Yogananda
Scientist, NBHM (DAE)
Department of Mathematics
Indian Institute of Science
Bangalore 560 012

16

Some Unsolved Problems in Number Theory

Progress Made in Recent Times

K Ramachandra

The beauty of the theory of numbers is that it poses so many simple-looking problems, most of which remain unsolved even today. Many of these problems have come down to us from ancient times, indicating the age-old fascination that human beings have felt for numbers. We list a few of these problems below, describing some known results and indicating the progress made in recent times.

The Infinitude of Primes

It is easy to show that the list $2, 3, 5, 7, 11, \ldots$ of primes does not terminate. The biggest prime known explicitly today has more than 10^5 digits! Now consider pairs of primes that differ by 2, for instance $(3,5), (5,7), (11,13), (17,19), \ldots$. These are the so-called *twin primes*. It is not known as of today whether the list of twin primes terminates or not. It *is* known that the sum $1/3 + 1/5 + 1/11 + 1/17 + \cdots = \sum 1/p$ taken over all primes p such that $p + 2$ is prime is finite (indeed, the sum, known as *Brun's constant*, can be computed to a fair degree of accuracy), but this does not prove that there are only finitely many such primes. (It is clearly possible for a sum of infinitely many positive numbers to be finite; for instance, this happens with the sets $\left\{\frac{1}{1}, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots\right\}$ and $\left\{\frac{1}{1}, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \frac{1}{25}, \ldots\right\}$. Ancient Greeks believed this was impossible. The well-known paradoxes of Zeno are related to this observation.) The best that we know today is that the list of pairs (p, q) of primes with

$$0 < p - q < c \ln p \qquad \left(c = \frac{1}{4}\right)$$

does not terminate. This is a very deep result due to H Maier of Germany. (Actually his constant c is slightly less than $1/4$.) We are very far from this result for, say, $c = 1/100$.

Another question deals with the number $\pi(x)$ of primes p below x. It was noticed by Legendre, Gauss, Riemann and others that $\pi(x)$ is roughly equal to $x/\ln x$; this is equivalent to saying that the nth prime is roughly equal to $n \ln n$. Chebychev showed that there exist constants a, b such that

$$a\frac{x}{\ln x} < \pi(x) < b\frac{x}{\ln x}$$

for all x. Using the methods of complex variables, Hadamard and de la Vallee Poussin proved independently in the 1890's that

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln(x)} = 1.$$

Instead of $\pi(x)$, it is nicer to deal with the function $Q(x)$ which counts the prime p with the weight $\ln p$; that is, $Q(x) = \sum_{p\le x} \ln p$. It was proved around the turn of the century that

$$|Q(x) - x| < x\left(e^{\sqrt{\ln x}}\right)^{-h}$$

for all $x > 10^{100}$ and a certain absolute positive constant h. The precise value of h is not important. One of the deepest results in prime number theory is the theorem that the term $e^{\sqrt{\ln x}}$ can be replaced by

$$e^{(\ln x)^{3/5}(\ln\ln x)^{-1/5}}$$

This result is due to the Soviet mathematician I M Vinogradov.

Additive Prime Number Theory

In 1742 Goldbach asked, in a letter to Euler, whether every even number from 6 onwards can be expressed as a sum of two odd primes. The answer to this question is unknown even today! The achievements in this problem have a very long history. Using the so-called 'circle method' pioneered by Ramanujan–Hardy, Hardy and Littlewood showed that if the hypothesis formulated below holds true, then every odd number from some point onwards can be expressed as a sum of 3 odd primes.

The hypothesis is stated in terms of the following function μ defined on the set of positive integers:

$$\mu(n) = \begin{cases} 1, & \text{for} \quad n = 1; \\ 0, & \text{if} \quad n \text{ is divisible by the square of a prime;} \\ (-1)^k, & \text{if} \quad n \text{ is the product of } k \text{ distinct primes.} \end{cases}$$

Let a, b be positive integers, and let $h > 3/4$ be a constant. The hypothesis then states that the following inequality holds for all $x > N(a, b, h)$, where N is some function that depends only on a, b, h:

$$\left|\sum \mu(an + b)\right| \le x^h.$$

This hypothesis is open as of today. It is considered very difficult to prove, even in the special case $a = b = 1$, $h = 1 - 10^{-100}$.

The Circle Method

(The 'circle method' was developed by Ramanujan and Hardy while they were working on the partition problem. The problem is to find an asymptotically accurate formula for $p(n)$, the number of partitions of n or the number of ways that n can be written as an unordered sum of positive integers ($p(1) = 1$, $p(2) = 2$, $p(3) = 3$, $p(4) = 5, \ldots$). It has been known from the time of Euler that

$$\prod_{j=1}^{\infty} \left(1 - z^j\right)^{-1} = \sum_{n=1}^{\infty} p(n) z^n.$$

Let $f(z)$ denote the infinite product on the left side. The singularities of $f(z)$ are the roots of unity and lie densely on the unit circle $|z| = 1$; thus $f(z)$ has the unit circle as its circle of convergence. Using Cauchy's residue theorem, we obtain

$$p(n) = \frac{1}{2\pi i} \oint_{|z|=r} \frac{f(z)}{z^{n+1}} dz,$$

for $0 < r < 1$. Thus the problem of estimating $p(n)$ has been converted into one of estimating an integral. The beautiful and amazingly productive idea pioneered by Ramanujan and Hardy was to estimate the integral by identifying the points where 'most' of the contribution comes from; these are clearly the points on $|z| = r$ that lie 'close' to the poles of $f(z)$. The practical details are formidable, but what is of significance is that the method, originally conceived to tackle the partition problem, has turned out to be applicable to a large class of related problems—for instance, Waring's problem.)

However, in 1937 Vinogradov proved the same result without having to use any unproved hypothesis. A recent result in the direction of Goldbach's conjecture is the one by O Ramare: *Every positive even number can be expressed as a sum of not more than 6 primes.* Another result by the author and his colleagues A Sankarayanan and K Srinivas is the following: let g_n denote the nth even number expressible as a sum of 2 odd primes ($g_1 = 6, g_2 = 8, g_3 = 10, \ldots$). We do not know whether the range of g exhausts the even numbers beyond 6, but the following is now known:

$$(g_{n+1} - g_n)^{37} < k g_n \qquad \text{for all} \quad n,$$

where k is a positive constant independent of n.

Waring's Problem

Let k be any natural number greater than 1. More than two centuries back, E Waring conjectured the following. *Let* $g(k) = 2^k + [1.5^k] - 2$ *and write* g *for* $g(k)$. *Then* *every positive integer n can be expressed as a sum of g or fewer positive k^{th} powers; that is, for all $n \in \mathbf{N}$ there exist non-negative integers $x_1, x_2, \ldots, x_g$ such that* $n = x_1^k + x_2^k + \cdots + x_g^k$. It is not too hard to check that the number $q = 2^k[1.5^k] - 1$ cannot be expressed as a sum of fewer than g positive k^{th} powers; that is, the equation

$$x_1^k + x_2^k + \cdots + x_{g-1}^k = q$$

has no solution in non-negative integers x_i. (*Example*: Let $k = 3$; then $g = 8 + 3 - 2 = 9$ and $q = (8 \times 3) - 1 = 23$. Since $23 < 3^3$, to express 23 as a sum of positive cubes we must use only the summands 1 and 8, and since $23 = (2 \times 8) + (7 \times 1)$, we require at least 9 such summands. Thus 23 cannot be expressed as a sum of fewer than 9 positive cubes.) Thus g is the most economical number of summands.

The current status of the problem is as follows: *There exists an absolute positive constant C such that Waring's conjecture is true for all $k > C$.* The proof derives from the ideas of Ramanujan, Hardy, Littlewood, Vinogradov, Dickson, Ridout and Mahler and is very complicated, running to hundreds of pages. It should be mentioned that the proof only establishes the existence of C and gives no clue as to its magnitude; no C, however large, can be calculated by the method of proof.

Problems on Irrationality

Consider the zeta function $\zeta(t)$ defined for real numbers $t > 1$ as follows: $\zeta(t) = \sum_{n \geq 1} 1/n^t$. One of the grand achievements of the century is the proof that

$$\zeta(3) = 1 + \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{4^3} + \cdots$$

is irrational. (An *irrational number* is one that is not expressible as a ratio of two non-zero integers. Related to the idea of irrationality is the notion of transcendence. A number is *algebraic* if it is the root of a polynomial with integral coefficients; else it is *transcendental*. Examples of algebraic irrationals are $\sqrt{2}$, $\sqrt[3]{2}$ and $\sqrt[3]{10} + \sqrt[5]{21}$, and examples of transcendental numbers are π, e and $\ln 2$ (here $e = 2.71828\ldots$ is Euler's number). The proof that a given number is transcendental can be extremely difficult.) The proof is due to R Apery. What happens when t is an odd positive integer greater than 3 is open. Strangely, a great deal is known when t is an even positive integer. Indeed, it is known that the value of $\zeta(t)$ is a rational multiple of π^t whenever t is an even positive integer. (This has been known since the time of Euler.) This immediately implies that $\zeta(t)$ is irrational, indeed transcendental, when t is an even positive integer. The paucity of positive conclusions for the case when t is an odd positive integer is extremely curious.

Much the same can be said for Euler's constant γ defined thus:

$$\gamma = \lim_{n \to \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) - \ln n.$$

Amazingly, it is not known whether γ is rational or not.

The transcendency of numbers such as $\pi + \ln 2$ was first proved by A Baker. These are deep results.

Concluding Remarks

It appears that there is no dearth of attractive problems. What is needed are solutions! What has been solved is very little and what remains to be solved is vast. In figurative terms, what has been solved can be likened to an egg-shell, and what remains to be solved to the infinite space surrounding it.

Addendum to "Some Unsolved Problems in Number Theory"

(***Resonance***, **May 1997**)

1. S S Pillai—The omission of the name S S Pillai (Siva Sankaranarayana Pillai) in connection with Waring's problem is very serious. In a series of papers, Pillai proved that if $k \geq 6$ and further if $(3^k + 1)/(2^k - 1) \leq [1.5^k] + 1$ then Waring's conjecture is correct for that k. Around the same time (but a little later) L E Dickson proved this with $k \geq 7$ and $(3^k + 1)/(2^k - 1) \leq [1.5^k] + 1$. The inequality $(3^k + 1)/(2^k - 1) \leq [1.5^k] + 1$ was proved for all integers exceeding a certain constant C (same C as in the paragraph on Waring's problem) by K Mahler. The history of this discovery is very well explained in *Introduction to the Theory of Numbers* by G H Hardy and E M Wright (see notes at the end of the chapter XXI). For another treasure house of information regarding priority of Pillai's work see K Chandrasekharan, S S Pillai (obituary), *J. Indian Math. Soc.*, Vol.15, pp 1–10, 1951. Regarding Pillai's achievements I mention the following: when I was in the Institute for Advanced Studies, Princeton, USA, during 1970–71, I noticed in the Institute Library a book by G H Hardy where he places Pillai as the greatest Indian mathematician after Srinivasa Ramanujan. Waring's conjecture was proved for $k = 5$ by Chen-Jing-Run (around 1970) and for $k = 4$ by R Balasubramanian, J-M Deshouillers and F Dress in 1989. Cases $k = 2$ and 3 were disposed off (by simpler methods) by Lagrange and Wieferich respectively. About Pillai I have the following comment: Once I was talking to a responsible Indian specialist dealing with History of Mathematics. I was very surprised when I came to know that he had not heard of Pillai at all. I can account for it as follows. Pillai was very unassuming; he was a member of the Indian Mathematical Society alright; but he was not a fellow of any of the academies and he had no publicity whatsoever amongst mathematicians who had not looked at the book by G H Hardy and E M Wright mentioned earlier.
2. The equation $\left|\sum \mu(an+b)\right| \leq x^h$ under the section 'Additive Prime Number Theory' should read $\left|\sum_{1\leq n\leq x} \mu(an+b)\right| \leq x^h$.

Cont...

Contd...

3. A comment on *The Circle Method* in the box on page 78:

The function $f(z)$ is analytic in $|z| < 1$ and it does not exist anywhere in $|z| \geq 1$. (So the terminology *poles of* $f(z)$ is not correct). We have to make r a suitable function of n but still less than 1. Then decompose this circle into small bits in a particular way and obtain asymptotics of each bit. The cumulative effect of adding all these asymptotics will give the Hardy–Ramanujan formula for partitions. Actually Ramanujan in his first letter (this letter was written from the Madras Port Trust) to Hardy mentions (see equation 1.14 of *Twelve Lectures*) that the integer $q(n)$ defined by

$$\left(\sum_{n=-\infty}^{\infty}(-x)^{n^2}\right)^{-1} = \sum_{n=0}^{\infty} q(n)x^n \quad \text{(note that LHS is the product } \prod_{h=1}^{\infty}\{(1-x^h)(1-x^{2h-1})\}^{-1}\text{)}$$

is the integer nearest to

$$\frac{1}{2}\frac{d}{dn}\left(\frac{\sinh(\pi\sqrt{n})}{\pi\sqrt{n}}\right).$$

When questioned about this, he wrote in a letter that it is "not the integer nearest to but this main term plus ...". (Compare this main term with the first term of the Hardy–Ramanujan–Rademacher formula for $p(n)$).

K Ramachandra

K Ramachandra
Honorary Visiting Professor
National Institute of Advanced Studies
Indian Institute of Science Campus
Bangalore 560 012

When and Where the Articles Appeared in Resonance

1. *To Find Four Distinct Positive Integers such that the Sum of Any Two of them is a Square*, Jun. 1997, 2(6), 89–90: S H Aravind.
2. *Bachet's Problem*, Nov. 1996, 1(11), 88–9: B Bagchi.
3. *On Fermat's Two Squares Theorem*, Mar. 1997, 2(3), 69–73: Shailesh A Shirali.
4. *Fermat's Two Squares Theorem Revisited*, July 1999, 4(7), 59–67: B Bagchi.
5. *Mathematical Induction—An Impresario of the Infinite*, Feb. 1998, 3(2), 69–7: B Sury.
6. *On Provability versus Consistency in Elementary Mathematics*, Sept. 1996, 1(9), 75–79: Shailesh A Shirali.
7. *Factoring Fermat Numbers—A Unique Computational Experiment for Factoring F_9*, Jan. 1996, 1(1), 108: C E Veni Madhavan.
8. *On the Infinitude of Prime Numbers—Euler's Proof*, Mar. 1996, 1(3), 78–95: Shailesh A Shirali.
9. *The Class Number Problem—Binary Quadratic Forms*, June 1998, 3(6), 25–30: Rajat Tandon.
10. *The Class Number Problem—An Introduction to Algebraic Number Theory*, July 1998, 3(7), 28–37: Rajat Tandon.
11. *Prime Representing Quadratics*, June 1999, 4(6), 91–92: N J Tejaswi.
12. *The Congruent Number Problem*, Aug. 1998, 3(8), 33–45: V Chandrashekar.
13. *Fermat's Last Theorem—A Theorem at Last*, Jan. 1996, 1(1), 71–79: C S Yogananda.
14. *Some Unsolved Problems in Number Theory—Progress Made in Recent Times*, May 1997, 2(5), 77–80: K Ramachandra.
15. *Roots are not Contained in Cyclotomic Fields*, Apr. 2001, 6(4), 78–83: Rajat Tandon.
16. *Die ganzen zahlen hat Gott gemacht—Polynomials with Integer Values*, Sept. 2001, 6(9), 46–60: B Sury.

Index

www.universitiespress.com

Number theory has fascinated mathematicians from the most ancient of times. A remarkable feature of number theory is the fact that there is something in it for everyone—from puzzle enthusiasts, problem solvers and amateur mathematicians to professional scientists and technologists. In this book we offer the reader some articles in number theory that appeared in *Resonance* over the years 1996–2001. The articles included within form a varied lot, beginning with a puzzle, "Find four positive integers such that the sum of any two is a square," to an expository article on one of the great mathematical achievements of the 20th century – the proof of "Fermat's Last Theorem".

Shailesh A Shirali is currently the Principal of Rishi Valley School and has taught Mathematics there for nearly two decades. He is closely involved with the Mathematical Olympiads, including the International Mathematical Olympiad. He was Chairman of the Problem Committee at the IMO held in 1996 in Mumbai, India, and leader of the Indian IMO teams that appeared for the IMO-1997 and IMO-1998. He is particularly interested in problem solving in the fields of geometry, number theory and combinatorics. He serves on the editorial boards of *Resonance* and *Samasya*.

C S Yogananda obtained his Ph.D. in mathematics in 1990 from the Institute of Mathematical Sciences, Madras. He has been involved in the Mathematical Olympiad programmes at various levels since 1989. His research interests lie in number theory; his other interests include classical music and mountaineering. He serves on the editorial boards of *Resonance* and *Samasya*.

Rs 125.00

ISBN 81 7371 454 1

9 788173 714542

Universities Press

Shirali and Yogananda: *Number Theory*