ISBN 0-9639903-0-6

Fermat's Last Theorem

A Supplement to the Video

Edited by

Robert Osserman

for the

MATHEMATICAL SCIENCES RESEARCH INSTITUTE

Berkeley, California

Correspondence concerning the video "Fermat's Last Theorem"
and orders should be sent to the:

MATHEMATICAL SCIENCES RESEARCH INSTITUTE
1000 Centennial Drive
Berkeley, California 94720 USA
Internet: video@msri.org

ISBN 0–9639903–0–6

Text prepared by QUANTEXT-intl., and typeset by AMS-TEX.
Slides and figures prepared by MSRI.
Printed in the United States of America.

Fermat's Last Theorem
A Supplement to the Video

Robert Osserman, Editor

MATHEMATICAL SCIENCES RESEARCH INSTITUTE
Berkeley, California

Contents

Introduction

This booklet is designed to accompany the video tape: **Fermat's Last Theorem, The Theorem and its Proof: An Exploration of Issues and Ideas**. The videotape is based on a presentation at the Palace of Fine Arts Theater in San Francisco that took place on July 28, 1993, exactly five weeks after Andrew Wiles announced that he had a proof of Fermat's Last Theorem. The presentation was sponsored jointly by the Mathematical Sciences Research Institute of Berkeley, California, and the Exploratorium of San Francisco. The videotape also includes excerpts from an interview with Andrew Wiles made in Oxford, England, the day after he announced his proof in Cambridge.

Both the booklet and the videotape are designed for a general audience. They deal not only with Fermat's Last Theorem itself, but with the excitement of mathematical discovery in general and the way mathematical ideas can illuminate other subjects. In fact, as of this writing (November 30, 1993), the proof presented by Wiles is still being checked by experts prior to publication. It is likely to be months, and possibly even years, before we can be totally confident that the proof is complete and correct. The issues and ideas presented here are of permanent value, independent of future developments regarding the proof.

Note on the Wolfskehl Prize

Fermat's Last Theorem gained particular notoriety when in 1907 a large amount of money was offered as a prize for its solution by P. Wolfskehl in Germany. The offer expires in 2007 if the prize has not been awarded by then. The prize is administered by the Academy of Sciences in Göttingen, who are instructed to wait two years after a proof appears in print before making a decision. The value of the prize money was reduced to nothing by the runaway inflation in Germany in the 1920s, but in recent years it has once again built up to a current value of approximately $40,000.

2

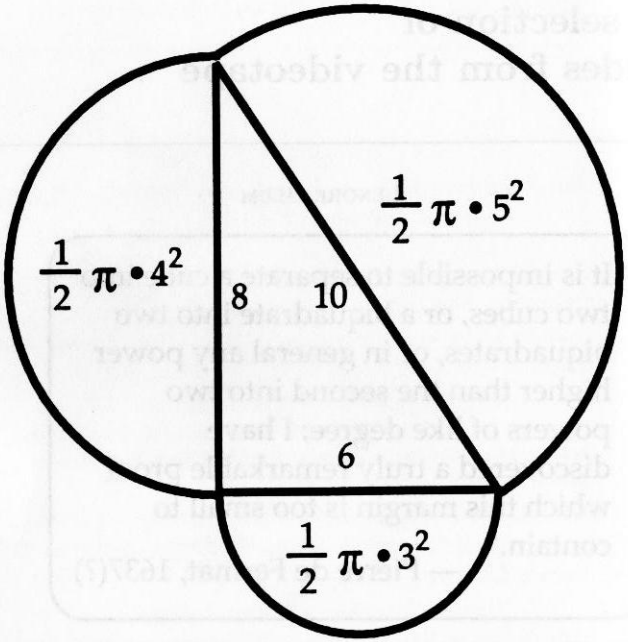
A selection of slides from the videotape

LENORE BLUM

"It is impossible to separate a cube into two cubes, or a biquadrate into two biquadrates, or in general any power higher than the second into two powers of like degree; I have discovered a truly remarkable proof which this margin is too small to contain."

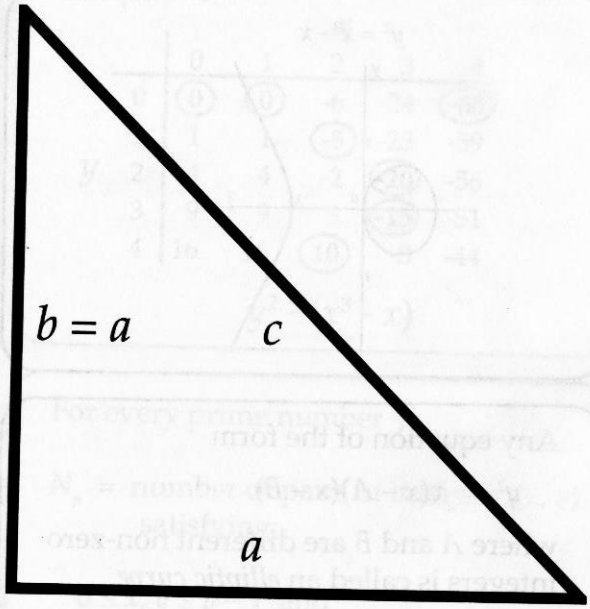
— Pierre de Fermat, 1637(?)

(?)1640: $n = 3(?), 4$	Fermat
1753: $n = 3$	Euler
1825: $n = 5$	Legendre, Dirichlet
1839: $n = 7$	Lamé
1847: Regular n	Kummer
1857: $n < 100$	Kummer
1930: $n < 600$	Vandner
1951: $n < 4{,}000$	Lehmer
1977: $n < 125{,}000$	Wagstaff
1992: $n < 4{,}000{,}000$	B, C, E, M, S
1993: all n!	Wiles



$$\frac{1}{2}\pi \cdot 9 + \frac{1}{2}\pi \cdot 16 = \frac{1}{2}\pi \cdot 25$$

$$9 + 16 = 25$$

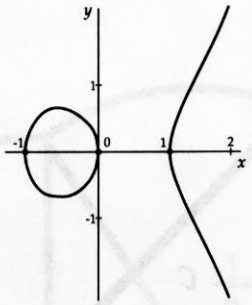


$$c^2 = a^2 + b^2 = 2a^2$$

$$\frac{c^2}{a^2} = 2$$

$$\frac{c}{a} = \sqrt{2}$$

$y^2 = x^3 - x$



Any equation of the form

$$y^2 = x(x - A)(x - B)$$

where A and B are different non-zero integers is called an *elliptic curve*.

(In the previous example we have

$$x^3 - x = x(x - 1)(x + 1)$$

so we can take $A = 1$, $B = -1$.)

An elliptic curve can be *modular*.

Instead of asking how often

$$y^2 = x^3 - x \quad \text{(i.e., } y^2 - (x^3 - x) = 0\text{)},$$

we will ask, for every prime number p: how often is $y^2 - (x^3 - x)$ a multiple of p?

Example: $p = 5$

y \ x	0	1	2	3	4
0	**0**	**0**	-6	-24	**-60**
1	1	1	**-5**	-23	-59
2	4	4	-2	**-20**	-56
3	9	9	3	**-15**	-51
4	16	16	**10**	-8	-44

$y^2 - (x^3 - x)$

For every prime number p,

N_p = number of pairs of integers (x, y) satisfying:

$0 \le x, y \le p - 1$ and
$y^2 - (x^3 - x)$ is a multiple of p

so $N_5 = 7$.

In 1814, Gauss found a recipe for calculating N_p for this elliptic curve:

- $N_2 = 2$.
- If p is 1 less than a multiple of 4, then $N_p = p$.
- If p is 1 more than a multiple of 4, then there is a more complicated formula.

This formula shows that the sequence

$$\begin{array}{cccccccc} N_2, & N_3, & N_5, & N_7, & N_{11}, & N_{13}, & \cdots \\ 2, & 3, & 7, & 7, & 11, & 7, & \cdots \end{array}$$

has a very special structure. Because of this structure we say that the elliptic curve

$y^2 = x^3 - x$ is *modular*.

Starting with any elliptic curve we can define the sequence

$N_2, N_3, N_5, N_7, N_{11}, N_{13}, \ldots$

in the same way. The elliptic curve is called *modular* if there is some rule or formula for this sequence which gives a structure like the one given by Gauss' formula.

A sequence must be very special to have this *modular* property.

Conjecture (Taniyama 1955, Shimura 1962). *Every elliptic curve is modular.*

Gerhard Frey (1985)

Suppose that we have a counter-example to Fermat's Last Theorem:

$a^n + b^n = c^n$ with $n > 2$.

Consider the elliptic curve

$y^2 = x(x - a^n)(x + b^n)$

This curve seems to be non-modular.

- Suppose $a^n + b^n = c^n$.
- Write down $y^2 = x(x - a^n)(x + b^n)$.
- This is a semistable elliptic curve.
- Wiles' 1993 theorem says it's modular.
- Ribet's 1986 theorem says it can't be modular. (Uses that $a^n + b^n$ is c^n.)
- This contradiction shows that there are no a, b, and c with $a^n + b^n = c^n$.

History

- Y. Taniyama, 1955. Conjecture: elliptic curves are modular.
- G. Frey, 1985. Fermat counterexample leads to elliptic curve. This curve seems non-modular.
- J-P. Serre, 1985. To show Frey's curve is non-modular, it suffices to prove two facts about modular forms.
- K. Ribet, 1986. Frey's curve is non-modular.
- A. Wiles, 1993. Frey's curve is modular.

Progress on Fermat's Last Theorem

- 1640 Fermat: $x^4 + y^4 \neq z^2$
- 1753 Euler: $x^3 + y^3 \neq z^3$
- 1828 Dirichlet: $x^5 + y^5 \neq z^5$
- 1839 Lamé: $x^7 + y^7 \neq z^7$

When Fermat's *other* theorems were proved

- 1749 Euler: $p = x^2 + y^2$
- 1772 Lagrange: 4 Squares
- 1801 Gauss: 3 Triangles
- 1813 Cauchy: 5 Pentagons, *etc.*
- 1840's Jacobi: Other theorems

10 Years Ago

Faltings proved the Mordell Conjecture. It follows that for each $n > 2$, $x^n + y^n = z^n$ has only finitely many essentially distinct solutions.

3
A copy of the page in Diophantus' book

(*in the original Greek, and Latin translation*)

Fermat's statement of his theorem was written in the margin of Diophantus' *Arithmetica*, next to Problem VIII of Book 2, where Diophantus solves a particular case of writing a square as the sum of two squares. He wants to find two fractions whose squares add up to 16. He obtains the answer:

$$16 = 4^2 = \left(\frac{12}{5}\right)^2 + \left(\frac{16}{5}\right)^2$$

or

$$400 = 20^2 = 12^2 + 16^2.$$

Τετάχθω ἡ μὲν ὑπεροχὴ αὐτῶν $\overset{\circ}{M}\bar{\beta}$, ὁ δὲ ἐλάσσων $\varsigma\,\bar{\alpha}$· ὁ ἄρα μείζων ἔσται $\varsigma\,\bar{\alpha}\,\overset{\circ}{M}\bar{\beta}$· δεήσει ἄρα $\varsigma\,\bar{\delta}\,\overset{\circ}{M}\bar{\delta}$ $\gamma^{\pi\lambda.}$ εἶναι $\overset{\circ}{M}\bar{\beta}$ καὶ ἔτι ὑπερέχειν $\overset{\circ}{M}\bar{\iota}$. τρὶς ἄρα $\overset{\circ}{M}\bar{\beta}$ μετὰ $\overset{\circ}{M}\bar{\iota}$ ἴσαι εἰσὶν $\varsigma\,\bar{\delta}\,\overset{\circ}{M}\bar{\delta}$· ἀλλὰ τρὶς $\overset{\circ}{M}\bar{\beta}$ μετὰ $\overset{\circ}{M}\bar{\iota}$ γίνονται $\overset{\circ}{M}\overline{\iota\varsigma}$· ταῦτα ἴσα $\varsigma\,\bar{\delta}\,\overset{\circ}{M}\bar{\delta}$, καὶ γίνεται ὁ $\varsigma\,\overset{\circ}{M}\bar{\gamma}$.

ἔσται ὁ μὲν ἐλάσσων ἀριθμὸς $\overset{\circ}{M}\bar{\gamma}$, ὁ δὲ μείζων $M\,\bar{\varepsilon}$, καὶ ποιοῦσι τὸ πρόβλημα.

η.

Τὸν ἐπιταχθέντα τετράγωνον διελεῖν εἰς δύο τετραγώνους.

Ἐπιτετάχθω δὴ τὸν $\overline{\iota\varsigma}$ διελεῖν εἰς δύο τετραγώνους.

Καὶ τετάχθω ὁ $\alpha^{\text{ος}}$ $\Delta^{\text{ν}}\,\bar{\alpha}$, ὁ ἄρα ἕτερος ἔσται $\overset{\circ}{M}\overline{\iota\varsigma}\,\wedge\,\Delta^{\text{ν}}\,\bar{\alpha}$· δεήσει ἄρα $\overset{\circ}{M}\overline{\iota\varsigma}\,\wedge\,\Delta^{\text{ν}}\,\bar{\alpha}$ ἴσας εἶναι $\square^{\text{ῳ}}$.

πλάσσω τὸν $\square^{\text{ον}}$ ἀπὸ $\varsigma^{\text{ῶν}}$ ὅσων δήποτε $\wedge$ τοσούτων M ὅσων ἐστὶν ἡ τῶν $\overline{\iota\varsigma}\,\overset{\circ}{M}$ πλευρά· ἔστω $\varsigma\,\bar{\beta}\,\wedge\,\overset{\circ}{M}\bar{\delta}$. αὐτὸς ἄρα ὁ $\square^{\text{ος}}$ ἔσται $\Delta^{\text{ν}}\,\bar{\delta}\,\overset{\circ}{M}\overline{\iota\varsigma}\,\wedge\,\varsigma\,\overline{\iota\varsigma}$· ταῦτα ἴσα $\overset{\circ}{M}\overline{\iota\varsigma}\,\wedge\,\Delta^{\text{ν}}\,\bar{\alpha}$. κοινὴ προσκείσθω ἡ λεῖψις καὶ ἀπὸ ὁμοίων ὅμοια.

$\Delta^{\text{ν}}$ ἄρα $\bar{\varepsilon}$ ἴσαι $\varsigma\,\overline{\iota\varsigma}$, καὶ γίνεται ὁ $\varsigma\,\overline{\iota\varsigma}$ πέμπτων.

ἔσται ὁ μὲν $\overset{\kappa\varepsilon}{\overline{\sigma\nu\varsigma}}$, ὁ δὲ $\overset{\kappa\varepsilon}{\overline{\rho\mu\delta}}$, καὶ οἱ δύο συντεθέντες ποιοῦσι $\overset{\kappa\varepsilon}{\overline{\upsilon}}$, ἤτοι $\overset{\circ}{M}\overline{\iota\varsigma}$, καὶ ἔστιν ἑκάτερος τετράγωνος.

---

4 ἀλλὰ . . . $\overset{\circ}{M}\bar{\delta}$ (5) om. *Ba*. 6 ἀριθμὸς om. *Ba*. 12 ὁ ἄρα . . . $\Delta^{\text{ν}}\,\bar{\alpha}$ (13) om. B. 14/15 τοσαύτας A. 15 $\overset{\circ}{M}\overline{\iota\varsigma}$ A 1ª m. 20 et 21 Denominatores hic, ut ubique infra, nisi contrarium adnotatum fuerit, om. A 1ª m., post numeratores (non supra lineam) add. 2ª m.; εἰκοστοπέμπτων scripsit B post $\overline{\sigma\nu\varsigma}$ et $\overline{\rho\mu\delta}$, εἰκοστόπεμπτα post $\bar{\upsilon}$. 21 ἤτοι add. A 2ª m.

Ponatur differentia ipsorum esse 2 et minor $= x$; ergo maior erit $= x + 2$. Oportebit igitur $4x + 4$ esse 3^{plum} 2 et adhuc superare 10. Ergo

$$3 \times 2 + 10 = 4x + 4.$$

Sed

$$3 \times 2 + 10 = 16.$$

Ista aequantur $4x + 4$ et fit $x = 3$.

Erit minor numerus $= 3$, maior $= 5$, et problema solvunt.

VIII.

Propositum quadratum partiri in duos quadratos.

Proponatur iam 16 partiri in duos quadratos.

Ponatur primus $= x^2$, alter erit igitur $16 - x^2$, et oportebit esse

$$16 - x^2 = \square.$$

Quadratum formo a quotlibet x minus tot unitatibus quot est radix 16. Esto a $2x - 4$, cuius quadratus erit

$$4x^2 + 16 - 16x.$$

Ista aequantur

$$16 - x^2.$$

Utrimque addantur negata et a similibus similia. Ergo

$$5x^2 = 16x \quad \text{et fit} \quad x = \frac{16}{5}.$$

Erit alter $\frac{256}{25}$, alter $\frac{144}{25}$, quorum summa facit $\frac{400}{25} = 16$, et uterque quadratus est.

4

Some background and history for Fermat's Last Theorem

Robert Osserman

I. Pythagorean triples

1. The basic fact is the so-called "Pythagorean Theorem": if a and b are the lengths of the sides of a right triangle and c is the length of the hypotenuse, then

$$a^2 + b^2 = c^2.$$

The followers of Pythagoras—the "Pythagoreans"—were particularly interested in right triangles where a, b, and c are whole numbers. (by "whole numbers", I shall mean the ordinary "counting numbers": $1, 2, 3, \ldots$.) In other words, they were looking for triples of whole numbers a, b, and c satisfying $a^2 + b^2 = c^2$. Any such a, b, c is called a *Pythagorean triple*.

2. The Pythagoreans discovered that it is impossible to have a Pythagorean triple with $a = b$. If there were such a triple, then

$$c^2 = a^2 + b^2 = 2a^2$$

or

$$c^2/a^2 = 2, \quad c/a = \sqrt{2}.$$

The Pythagoreans proved that $\sqrt{2}$ cannot be written as a fraction —the ratio of two whole numbers. (The expression "rational number" is also used for a fraction. Hence $\sqrt{2}$ is *irrational*.)

3. Diophantus of Alexandria is believed to have lived in the third century A.D.. His book *Arithmetica* is really about what we would call "algebra": solving equations—but he is specifically interested in solutions that are whole numbers or fractions (which often amounts to the same thing, since for example, if x, y, z, are fractions satisfying $x^n + y^n = z^n$, then multiplying through by the product of the denominators gives a solution of $a^n + b^n = c^n$ in whole numbers.) We now call a *Diophantine equation* any algebraic equation with whole-number coefficients where we look for whole-number solutions. Finding Pythagorean triples is the same as solving the Diophantine equation

$$x^2 + y^2 = z^2.$$

4. The Big Facts about Pythagorean triples are that first: you can write down an infinite number of essentially different ones (where "essentially different" means not just multiples of each other, such as $3, 4, 5$; $6, 8, 10$; $9, 12, 15$, etc.) by using the formula

$$a = m^2 - n^2, \quad b = 2mn, \quad c = m^2 + n^2,$$

where m and n are any whole numbers and m is bigger than n. (You can check for yourself that for any m and n, $a^2 + b^2 = c^2$.) Second: every Pythagorean triple is either one of those given by the formula, or else a multiple of one of them.

For example, if we choose $m = n + 1$ in our general formula, then we get the following table.

n	$a = 2n+1$	$b = 2n(n+1)$	$c = b+1$
1	3	4	5
2	5	12	13
3	7	24	25
4	9	40	41

which gives a Pythagorean triple a, b, c where a can be any odd number starting at 3. Each row of the table gives an infinite number of triples that are not essentially different, by just multiplying through by $2, 3$, etc.: $5, 12, 13$; $10, 24, 26$; $15, 36, 39$; ... .

In the same way we get more examples by taking other specific value of m and n, and then all multiples; for $m = 4$, $n = 1$, we get $8, 15, 17$; $16, 30, 34$ etc. The important fact for later applications is that there are no other Pythagorean triples besides those obtained in this manner. In other words, this method gives a complete solution to the Diophantine equation $x^2 + y^2 = z^2$ in whole numbers.

II. Enter Fermat

1. Chronology

1601 Pierre de Fermat born in a small town near Toulouse, in the south of France.

1621 Diophantus' *Arithmetica* published in the original Greek with a Latin translation and extensive commentary by Bachet. (Note: The tradition of the time was for all mathematics (and other science) articles and books to be written in Latin—allowing mathematicans and scientists from all countries to read the work of each other without learning a new language.)

1638–1659 Fermat's letters to various of his correspondents mention the two problems:

(1) show that $x^3+y^3 = z^3$ and $x^4+y^4 = z^4$ have no solutions in whole numbers,

(2) show that if the sides of a right triangle are whole numbers, then the area cannot be a perfect square.

These problems are often presented as challenges, seeming to imply that Fermat has his own solutions.

1665 Fermat dies. His eldest son, Samuel, decides to publish Fermat's many unpublished mathematical writings, including the notes in the margins of Fermat's copy of Bachet's translation of Diophantus' *Arithmetica*.

2. The notes in the margin

One of the notes contains the main ideas for answering the question posed by Fermat of whether a right triangle with whole-number sides can have area equal to a perfect square. The answer, says Fermat, is "no", and he fits enough of the proof in the margin to make his reasoning clear, although he concludes "there is not enough room in the margin for a complete proof".

Another of the notes is the one that came to be called "Fermat's Last Theorem", stating that $x^n + y^n = z^n$ has no solution in whole numbers x, y, z if n is a whole number greater than 2. In this case, Fermat says he has a marvelous proof, but the margin

is too narrow to contain it, and he does not include even a hint of how it might go. However, the case $n = 4$ can be proved in a manner very like the one used in the other marginal note described above, and there is no doubt that Fermat had done that.

3. The proof of the case n = 4 of Fermat's Last Theorem

Fermat's method for settling the case $n = 4$, like Wiles' method for proving the complete result for all n is an argument by contradiction. In each case, you start by assuming that there is a solution to the equation $x^n + y^n = z^n$, and then by a sequence of mathematical deductions, arrive at a contradiction. The conclusion is that the original assumption that the equation had a solution must have been false.

The key idea in Fermat's argument is this: suppose there do exist whole numbers x, y, z satisfying $x^4 + y^4 = z^4$. Fermat shows that then there must be smaller whole numbers satisfying the same equation. But you can then repeat the argument with the smaller numbers and get still smaller ones. The process could be repeated indefinitely since the same argument applies to each new set of whole-number solutions. But that is clearly absurd, since whatever the original value of z, if it gets reduced by at least 1 in each step, then after z steps you will have run out of whole numbers. So the original assumption that a solution exists to $x^4 + y^4 = z^4$ leads to a contradiction, and we are forced to conclude that no solution exists.

Here is the key step: suppose that x and y are whole numbers, with $x^4 + y^4$ a perfect square; that is

$$x^4 + y^4 = c^2,$$

where c is a whole number. Then we can find new whole numbers u, v, with

$$u^4 + v^4 = d^2,$$

where d is a whole number smaller than c. (Repeating the argument c times clearly leads to a contradiction.)

The proof of this statement is as follows. Let $a = x^2$, $b = y^2$. Then $x^4 + y^4 = c^2$ means

$$a^2 + b^2 = c^2.$$

But we know that every solution of this equation is of the form

$$a = k \cdot 2mn, \quad b = k(m^2 - n^2), \quad c = k(m^2 + n^2),$$

where k, m, and n are whole numbers. We can also assume that m and n have no common factor, because if they did, we could factor it out and include it in the factor "k". We now prove the statement in each of two cases.

CASE 1: $k > 1$. In this case, let p be any prime factor of k. Then p divides into both a and b. Since $a = x^2$, $b = y^2$, it follows that p must be a factor of both x and y. Hence p^4 divides into x^4 and y^4, and therefore also into $c^2 = x^4 + y^4$. Then let $u = x/p$, $v = y/p$, $d = c/p^2$. We get

$$u^4 + v^4 = \frac{x^4}{p^4} + \frac{y^4}{p^4} = \frac{c^2}{p^4} = d^2.$$

where

$$d = \frac{c}{p^2} < c.$$

This proves the statement in Case 1.

CASE 2: $k = 1$. Then

$$a = 2mn, \quad b = m^2 - n^2, \quad c = m^2 + n^2.$$

In this case, we notice first that m and n cannot both be even, since they are assumed to have no common factor. On the other hand, one of them has to be even, since $a = 2mn$ is an even number, and $a = x^2$, so that x must be even, and that means that a is divisible by 4, or that $a/2 = mn$ is an even number. It follows that m and n consist of one even and one odd number. Then $b = m^2 - n^2$ and $c = m^2 + n^2$ must both be odd.

What we have just proved can be stated this way; if a, b, c, is a Pythagorean triple with no common factor (that is, $k = 1$) then c must be an odd number.

Now in our case, we have

$$b = m^2 - n^2, \qquad \text{where } b = y^2.$$

That means $m^2 - n^2 = y^2$, or

$$n^2 + y^2 = m^2.$$

Now if n and y had any common divisor p, then p would also divide into m. But m and n have no common factors, so that n

and y cannot have any either. That means that n, y, m form a Pythagorean triple with no common divisor, and therefore, as we have just proved, m has to be an odd number. Since m and n cannot both be odd, n is even. Finally since $n^2 + y^2 = m^2$, we can write them in the form

$$n = 2rs, \quad y = r^2 - s^2, \quad m = r^2 + s^2,$$

where r and s are whole numbers with no common factor. But then

$$x^2 = a = 2mn = 4rs(r^2 + s^2).$$

CLAIM. *The numbers r, s and $r^2 + s^2$ are all perfect squares.*

PROOF. Let p be any prime number that divides r. Since r and s have no common factor, it follows that p does not divide s. Hence p is a factor of r^2 but not of s^2, so that it is not a factor of $r^2 + s^2$. On the other hand, p is a factor of the whole number

$$\left(\frac{x}{2}\right)^2 = rs(r^2 + s^2),$$

so that p is a factor of $x/2$. If p^k is a factor of $x/2$, for $k \geq 1$, then p^{2k} is a factor of $(x/2)^2$. Since the only term on the right containing p as a factor is r , it follows that p^{2k} is a factor of r. In other words, every prime factor of r occurs to an even power. Hence r is a perfect square. Exactly the same argument holds for s. It follows that

$$r^2 + s^2 = \left(\frac{x}{2}\right)^2 / rs$$

is also a perfect square. This proves the claim.

Now what the claim says is that we may set

$$r = u^2, \quad s = v^2, \quad r^2 + s^2 = d^2,$$

where u, v, and d are whole numbers. In other words,

$$u^4 + v^4 = r^2 + s^2 = d^2,$$

where

$$d < d^2 = r^2 + s^2 = m < m^2 + n^2 = c.$$

This completes the proof of Case 2 of our main statement: any solution of

$$x^4 + y^4 = c^2$$

leads to another solution

$$u^4 + v^4 = d^2$$

with $d < c$.

Finally, if there were any solutions of

$$x^4 + y^4 = z^4$$

in whole numbers, then we could let $c = z^2$ to get $x^4 + y^4 = c^2$, and applying the above argument would lead to a contradiction. We are forced to conclude that no such solution is possible, which proves Fermat's Last Theorem for the case $n = 4$.

4. Enter elliptic curves

The second of the two problems mentioned a number of times by Fermat was to show that if a right triangle has whole-number sides, then its area cannot be a perfect square. If the sides are a, b, c, then the area is $ab/2$. The statement then is that the pair of equations

$$a^2 + b^2 = c^2$$

and

$$ab = 2d^2$$

have no whole number solutions a, b, c, d.

Since every solution of the first equation is of the form

$$a = k \cdot (m^2 - n^2), \quad b = k \cdot 2mn, \quad c = k \cdot (m^2 + n^2),$$

the second equation takes the form:

$$2d^2 = ab = k^2(m^2 - n^2) \cdot 2mn = 2k^2mn(m^2 - n^2)$$

or

$$d^2 = k^2(m^3n - mn^3).$$

We now let

$$x = \frac{m}{n}, \quad y = \frac{dn^2}{k}.$$

Then

$$x^3 - x = \frac{m^3}{n^3} - \frac{m}{n} = \frac{m^3 n - m n^3}{n^4} = \frac{n^4 d^2}{k^2} = y^2.$$

Proving that there are no whole number solutions a, b, c, d to the original equations is the same as proving that there is no solution x, y to the elliptic equation

$$y^2 = x^3 - x,$$

where x and y are fractions m/n, r/s with whole numbers m, n, r, s.

5
Sophie Germain's contribution

Joe Buhler

Perhaps the first truly general results of any depth on Fermat's Last Theorem were due to Sophie Germain in the early nineteenth century. Fermat's assertion was also verified for specific small exponents around this time, but the next general results were obtained by Kummer in the middle of the century. In order to state Sophie Germain's most famous theorem on this subject it is useful to first make some algebraic observations.

Suppose that there is a solution to Fermat's equation

$$x^n + y^n = z^n$$

for some exponent $n > 2$. Any number n larger than 2 is either divisible by an odd prime p or by 4. (Recall that a *prime* is a number such as 17, which is not divisible by any smaller number, and that to say that n is divisible by p means that $n = pm$ for some whole number m.) If $n = pm$ then

$$x^n + y^n = (x^m)^p + (y^m)^p = (z^m)^p = z^n.$$

This means that *if* there is a solution for the exponent n *then* there is a solution for the exponent p. By one of the logical transformations which mathematicians are so found of, this means that it suffices to prove Fermat's Last Theorem for the exponent 4 and for all primes $p > 2$. Fermat proved the result for $n = 4$ so, as was well known to eighteenth century mathematicians, it suffices to prove Fermat's Last Theorem for exponents p that are odd primes.

For various reasons it is also natural to subdivide the assertion of Fermat's Last Theorem into two cases called, naturally enough,

the *First Case* and the *Second Case*. The First case says that the equation

$$x^p + y^p = z^p$$

has no solution for which all of the whole numbers x, y, and z are *not* divisible by p. The *Second Case* says that the equation has no solutions in relatively prime whole numbers for which *at least one* of x, y, and z is divisible by p. At first glance this is a slightly peculiar distinction since it relates the exponents p with the solutions x, y, and z. However, it has turned out to be useful since various results can be proved for one or the other case, but not both. Note that Fermat's Last Theorem for the exponent p is equivalent to the statement that *both* the First Case and the Second Case are true for p.

Sophie Germain's most famous result on Fermat's Last Theorem was really a consequence of a more general technical theorem of hers. The "corollary" says that if p and $2p+1$ are both primes then the First Case of Fermat's Last Theorem is true. Thus the First Case is true for the exponent 5 since 5 and 11 are primes. Similarly, it is true for 11 since 11 and 23 are primes. However, this does not help for 7 since 15 is not a prime.

Further generalizations of this idea enabled Germain and Legendre, with whom she corresponded, to prove that the First Case of Fermat's Last Theorem was true for all primes (and hence all exponents) less than 100.

6

Rational and irrational: music and mathematics[1]

Robert Osserman

... some specifically human tendency to create and notice organized patterns, hierarchies, and sequences

—JOHN A. SLOBODA
The Musical Mind

... I still feel that mathematics, more than any other discipline, studies the fundamental, pervasive patterns of the universe .... To me, the deepest and most mysterious of all patterns is music

—DOUGLAS HOFSTADTER
Introduction to *Metamagical Themas*

The name of Pythagoras is most frequently asociated with the Pythagorean theorem, perhaps the best known and most often cited result in all of mathematics. But the appelation "Pythagorean" occurs in another context of probably far greater importance for science as a whole: the Pythagorean theory of music. The basic tenet of the theory is that music has profound underpinnings in mathematics. The reason for the powerful impact of an apparently narrowly-focused theory is that it was viewed as merely an illustration and confirmation of a broad doctrine, holding that mathematical relationships are at the heart of the physical world. That doctrine, in the hands of Kepler, Einstein, and many others, has led to major advances in our understanding of the universe.

[1]Reprinted from *Essays in Humanistic Mathematics*, MAA 1993

Our goal here is to describe the Pythagorean theory but with a kind of reverse twist: not to explain music in terms of mathematics, but, rather, to use music to motivate and clarify some elusive mathematical concepts. One might say that the goal is to prove that there really is rhythm in a logarithm.

Long before Pythagoras, it must have been observed that, if a string is held taut and either struck or plucked, it will produce a musical note: the shorter the string (assuming that is is maintained at the same tension), the higher the note. Furthermore, if two strings are struck simultaneously, the result will sometimes be perceived as discordant or dissonant, and sometimes as concordant or harmonious.

What the Pythagorean theory[2] states is that, provided the strings are both of the same material and at equal tension, they will sound harmonious precisely when the lengths of the strings form a simple mathematical ratio such as $2:1$, $3:2$, $4:3$—the more complicated the ratio, the more dissonant the sound. The simplest of these, the ratio $2:1$, produces sounds that are so concordant that we often describe them as "the same note, an octave apart". What we hear in each case is a pair of notes that we perceive as being a certain "distance" apart. That distance is referred to as a *musical interval*. The ratio $3:2$ corresponds to the interval called a *fifth*, the ratio $4:3$ to a *fourth*. The terminology derives simply from the position of the corresponding notes on the scale. If we start the scale on the note corresponding to the longer of the two strings, then the shorter one will be the

[2]I shall use the phrase "Pythagorean Theory" throughout, since that is the standard term in Western writing, although it cannot really be justified on historical grounds. Pythagoras himself was a shadowy and enigmatic figure who lived some two hundred years before Euclid. We have no clear evidence that he produced any written documents, and, in any case, none have come down to us. What he did have was a large following of "Pythagoreans" for many generations afterwards, and the general consensus of scholars today is that the proof of the "Pythagorean Theorem", as we know it from Euclid, is far more likely to have originated with them, perhaps a century after Pythagoras, than with Pythagoras himself. There is somewhat more reason to believe that Pythagoras did expound the "Pythagorean" theory of music, although it should be added that the first *written* record that we have occurs in Euclid: *Section of the Canon*. To be more precise, Euclid is the first written record in the *West*, but ancient Chinese historians document a much earlier development of the same theory, attributed to *Ling Lun*, perhaps two thousand years before Pythagoras.

fifth note up on the scale if their ratio is $3:2$. The word *octave* is simply Latin for "eighth", the position on the scale corresponding to the ratio $2:1$.

In Figure 1 we show how this process works when referred to a standard piano keyboard. The vertical line segment at the left represents a string length corresponding to the note "C" directly above it. As it is successively divided into 2, 3, 4, 5, and 6 equal parts, the resulting string lengths—the darkened part of each interval—correspond to the notes directly above them on the scale. The ratios of successive string lengths and the corresponding musical intervals are indicated below. For example, dividing the left-hand segment into two equal parts, we obtain a string half as long which will produce a note again designated "C", but an octave above the original. Dividing the original string into three equal parts, we obtain a string of a third the original length. Compared to the previous one—half the original length—we obtain two strings in the ratio $3:2$, producing the musical interval "a fifth", with the upper note designated by "G". Proceeding in the same fashion, we obtain the series of notes shown in Figure 1.

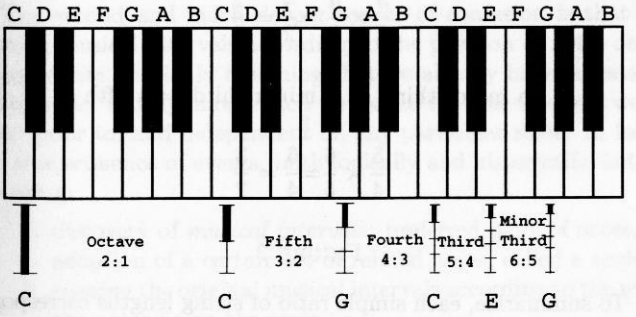


FIGURE 1

We now ask the following fundamental question: "How is the 'addition' of intervals expressed in the language of ratios?" The answer turns out to be as simple as it is fundamental: *adding musical intervals translates to multiplying the ratios*. Referring back to Figure 1, we see, for example, that if we wish to add a fifth and

a fourth, we must first take a string 2/3 the length of the original one, and then a string 3/4 the length of the second one. The resulting string will be 2/3 times 3/4 the original length, or half as long, and the corresponding interval is an octave. Musically, we have first moved up from C to G, and then up from G to the following C, a total of one octave. (See Figures 1 and 2.) The same process is completely general: moving from one note to a higher one through a certain interval means going from one string to a second one whose length is a certain fraction of the original length. Moving up from there through another interval, we arrive at a third string whose length is some fraction of the length of the second string. To find the relation of the third string to the first, we must take a *fraction of a fraction*; that is, we must multiply the corresponding fractions.

$$\begin{array}{ccc} \text{musical intervals} & & \text{adding intervals} \\ \updownarrow & & \updownarrow \\ \text{ratios of lengths} & & \text{multiplying ratios} \end{array}$$

$$\begin{array}{c} \text{a fifth} + \text{a fourth} = \text{an octave} \\ \updownarrow \\ \frac{3}{2} \times \frac{4}{3} = \frac{4}{2} = \frac{2}{1} \end{array}$$

$$\begin{array}{c} \text{a major third} + \text{a minor third} = \text{a fifth} \\ \updownarrow \\ \frac{5}{4} \times \frac{6}{5} = \frac{6}{4} = \frac{3}{2} \end{array}$$

FIGURE 2

To summarize, each simple ratio of string lengths corresponds to a musical interval, and *multiplying* two ratios corresponds to *adding* the musical intervals.

For someone with a mathematical background, such a description is bound to stike a responsive chord: the conversion of multiplication to addition is precisely the basic property of *logarithms*. Thus, the fact of logarithms is built into the physiology of sound perception, even though it was not until 1614 that the idea of logarithms was extracted and explicitly formulated by Napier.

One point that tends to obscure the relationship between numerical ratios and musical intervals is that there are two peculiarities in the *naming* of musical intervals. The first, and more superficial, is that an interval is named after the position on the scale of the upper of the two notes, assuming that the lower one is the first note on the scale. It would be much more natural to name an interval according to the space between the two notes, of the *difference* in position: that is, the number of steps on the scale needed to go from the first note to the second. For example, what we have called a "fifth", the interval from C to G, should rightly be assigned the number *four*, since there are four steps in between them: C to D, D to E, E to F, and F to G. Continuing from G up to the C above requires three more steps (see Figure 1), so that we should assign the number *three* to the interval we have called a "fourth". Then, adding the two musical intervals would correspond to simple numerical addition: $3 + 4 = 7$; it takes *seven* steps to go from C to the C above, and interval of an "octave". Renaming all musical intervals in this fashion might be a worthwhile reform, but it is unlikely to happen; the names of musical intervals are as firmly entrenched as the many bizarre spellings in our language, and appear equally resistant to rationalization.

The second and much deeper source of confusion is that in naming musical intervals according to the position of notes on a scale, we are implicitly assuming that we already have an established scale. But upon reflection we realize that the intervals were there prior to, and independent of, any particular scale. In fact, the true sequence of events, both logically and historically, is the following:

1. discovery of *musical intervals*: preferred pairs of notes,
2. adoption of a certain set of related notes, called a *scale*,
3. *naming* the original musical intervals according to the way the notes lie on the scale.

We are thus led to one of the most fundamental questions in music. How do we go about choosing, out of all the (infinite) possibilities for musical notes, the small number that will make up our "scale"?

The basic idea is simplicity itself. Out of all possible string lengths, the ones we wish to choose are those that sound most harmonious together. Following Pythagoras (or Ling Lun), we

may start with some arbitrary length, and then choose the others, using the simplest possible ratios. To begin with, we may restrict ourselves to the ratios 2 : 1 and 3 : 2. If our original string length corresponds to the note "C" on the piano, then using the ratio 3 : 2 will bring us up a "fifth" to the note designated by G. Repeating the process, we start on G and find that the fifth note up (corresponding to the ratio 3 : 2 once more) is the note "D". (Check once more on Figure 1.) We may now use our ratio 2 : 1 to move down an octave, and arrive at the note "D" just above our original "C".

Pythagoras tells us to continue in the same fashion, each time adding a new note to what is to become the privileged set of notes, called our "scale". If we carry out the process, we arrive successively at the notes C, G, D, A, E, B, then the five "black notes" lying between those with letters, then F, and then back to C. In other words, we obtain the *entire scale* on the piano just by using the ratios 2 : 1 and 3 : 2. If you have a piano handy, try it out starting at C near the bottom of the keyboard, and moving upwards in successive fifths. You will find that after twelve steps you will have struck each different note exactly once (where notes that differ by any number of octaves are counted as the same) and have come back to C, seven octaves above your starting point. We may express that fact as a musical equation:

$$12 \text{ fifths} = 7 \text{ octaves}$$

or

$$\text{a fifth} + \text{a fifth} + \cdots + \text{a fifth (12 times)} = \text{an octave} + \cdots + \text{an octave (7 times)}.$$

Since adding intervals corresponds to multiplying ratios, that musical equation converts to the mathematical one:

$$\text{“}(3/2)^{12} = 2^7\text{”}.$$

We have all at once arrive at a very suspicious looking "equation", and, indeed, a quick check on our calculator reveals that

$$(3/2)^{12} = 129.746\ldots, \qquad \text{while } 2^7 = 128.$$

So the two sides of our "equation" are close, but definitely not equal. What has gone wrong?

Before we answer that question, let us make a brief digression, to examine more closely the offending "equation". We may ask whether there is some way to see why the two sides cannot be equal, other than by a mindless calculation. One way to approach the problem would be to try to eliminate the fractions. In fact,

$$(3/2)^{12} = 3^{12}/2^{12}$$

so that the "equation" takes the form

$$\text{“}3^{12}/2^{12} = 2^7\text{”}$$

and multiplying through by 2^{12} leads to

$$\text{“}3^{12} = 2^7 \cdot 2^{12} = 2^{19}\text{”}.$$

Now each side is an integer and the two sides are transparently not equal. One could give a variety of reasons, the simplest being that the left hand side is odd, and the right is even. We therefore conclude that our original "equation" must also not be valid.

Is this more roundabout argument any better than a direct check with a calculator? That is clearly a matter of subjective judgment, but the abstract argument does yield more insight. In fact, the same argument holds much more broadly, and shows that *no* power of 3/2 can ever equal any power of 2. Converting back to music we conclude that no successive number of "pure fifths" (corresponding to the ratio 3 : 2) can ever equal any exact number of octaves. That simple fact has profound musical implications, and it leads us back to the problem faced by Pythagoras: what can be done about the frustrating fact that twelve fifths is ***almost*** but not quite, equal to seven octaves? The answer is the highly unsatisfying but inevitable one: you have to stretch things a bit here and there, and hope that no one will notice. Over the centuries many modifications of the Pythagorean approach have been proposed and hotly debated. One simple one was suggested by the famous Greek astronomer Ptolemy, who also devoted considerable energy to musical theory. He noted that the interval of a major third (from C to E) would correspond in the strict Pythagorean theory to the ratio

$$(3/2)^4/2^2 = 81/64$$

obtained by going up four successive fifths and down two octaves, but that a far simpler (and hence more harmonious) ratio would be

$$80/64 = 5/4,$$

which is the one we have used in Figure 1. Ptolemy proposed in general using the simplest ratios that approximated the notes obtained from the Pythagorean scheme. The advantage of Ptolemy's approach was that it produced simpler ratios, and presumably, as a consequence, more pleasing sounds. The disadvantage was that what formerly were equal intervals, such as from C to D, and from D to E, were now different, corresponding to the ratios 10/9 and 9/8, respectively.

And so the debate continued from generation to generation, with new proposals and compromises. A leading figure in the sixteenth century was Vincenzo Galilei, who suggested a method for placing frets on a lute to divide the octave into twelve equal intervals. Those intervals are called *semitones* and correspond to the interval between adjacent white and black notes on the piano. Vincenzo's method amounted to a choice of the ratio 18/17 for each semitone, and his claim that twelve of those intervals made up an octave translates to the equation

$$(18/17)^{12} = 2.$$

But that cannot be exactly correct, for much the same reasons we gave in the earlier case (or again, by a direct computation, as was rapidly pointed out by his contemporaries). In any case Vincenzo's main claim to immortality is as the father of Galileo Galilei, who, among his other accomplishments, became in turn the father of Physics, with the publication in the Spring of 1638 of his book *Two New Sciences*, a document very much in the Pythagorean tradition, using precise mathematical reasoning to analyze the working of the physical world.

As time went on, the debate began to focus more and more on the conflict between simple ratios and equal intervals, with the gradual realization that the two were incompatible. Finally, in a book called *Harmonie Universelle*, published in 1636 (just two years before Galileo's book), the French mathematician Marin Mersenne spelled out the details of what has since become known as *equal temperament*: the division of an octave into twelve exactly equal intervals. From our modern standpoint the solution is

almost trivial. If r denotes the ratio of strings needed to produce an interval which repeated twelve times gives an octave, then, since the ratio for an octave is $2:1$, our basic correspondence principle tells us that necessarily

$$r^{12} = 2,$$

and therefore,

$$r = \sqrt[12]{2}.$$

Mersenne also gave an approximate value for r. To four decimal places, it is

$$r \sim 1.0595.$$

For historal completeness we should mention that once again the West was anticipated by a Chinese scholar, Prince Chu Tsai-yu, who in 1596 explained the principles of equal temperament, and gave the correct value of r to nine decimal places!

With these discoveries the terms of the debate shifted. Now the argument was between *just intonation*, where intervals were determined by simple ratios of whole numbers, and equal temperament, where simple ratios were abandoned, but all intervals were equal in size. After Bach took up the banner of equal temperament, displaying its great versatility in his collection of preludes and fugues called *The Well-Tempered Clavichord*, there has been virtually no contest, with almost universal acceptance (at least in the West) of the equal-tempered scale. However, there have always been and continue to be holdouts. To understand the reasons for the objections, consider just one interval on the well-tempered scale—the one that divides the octave precisely in half. Since there are twelve semi-tones to the octave, the interval in question would consist of six semi-tones. Starting at C and moving up six semi-tones, we arrive at the black note between F and G. That note is called F-sharp, and written $F^{\#}$. The corresponding interval from C to $F^{\#}$ is called an *augmented fourth*. (See Figure 3.) The interval from $F^{\#}$ to the C above is again an augmented fourth, as we may check by confirming that it consists also of six semi-tones.

We now ask the question: What is the ratio of string lengths needed to produce the musical interval of an augmented fourth on the well-tempered scale? Since that interval added to itself

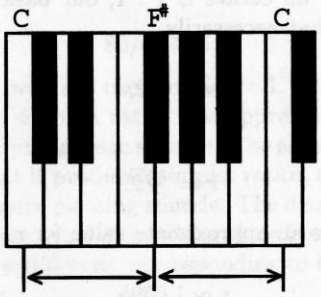


FIGURE 3

gives an octave, we deduce immediately that the corresponding ratio, call it m/n, must satisfy the equation

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

But, as the Pythagoreans discovered, *this equation has no solution*: there are no whole numbers m and n satisfying it. In the terminology of the time, the strings needed to produce the interval of an augmented fourth must be *incommensurable*. In modern notation, a number whose square is two would be written as $\sqrt{2}$, and we call it an *irrational* number, since it cannot be expressed as the ratio of two whole numbers.

How are we to interpret this situation? Looked at from the Pythagorean point of view, where the most harmonious intervals correspond to the simplest ratios, and more complicated ratios become more and more discordant, an interval corresponding to *no* ratio at all should be maximally discordant. Interestingly enough, this musical interval, also called the *tritone*, was considered terribly jarring, and referred to as *Diabolus in Musica* in the Middle Ages[3].

To be quite fair, we should not single out the tritone since on the well-tempered scale, *all* the intervals correspond to irrational

[3]According to the Harvard Dictionary of Music, "It has always been considered a 'dangerous' interval, to be avoided or treated with great caution", while the Grove Dictionary states, "In nineteenth century Romantic opera the tritone regularly portrays the ominous or evil."

numbers, starting with the semi-tone: the twelfth root of two. The difference, presumably, is that the other intervals are quite close to simple ratios of whole numbers, whereas the augmented fourth is not. For example, a fourth consists of five semi-tones, and a fifth consists of seven. The corresponding numerical values are therefore

	"just" intonation	well-tempered
a fourth	$4/3 \sim 1.3333$	$\left(\sqrt[12]{2}\right)^5 \sim 1.3348$
a fifth	$3/2 = 1.5$	$\left(\sqrt[12]{2}\right)^7 \sim 1.4983$

both remarkably close to the exact ratios. Nevertheless, some musicians experience well-tempered intervals as jarring, and continue to press for just intonation. Some, like Harry Partch, build their own musical instruments, based on one or another modification of the Pythagorean ideal, and compose and perform music for those instruments. Incidentally, Partch's book, *Genesis of a Music*, is a wonderful source for history and commentary on the matters we have been discussing. His chapter on "The Language of Ratios" is particularly appropriate.

One last comment before turning to another aspect of our subject: one might ask, in view of the incommensurable lengths of the strings needed to produce, say, an augmented fourth on the well-tempered scale, how one could ever construct such a pair of strings. There are at least two answers. First, the distinction between rational and irrational is not physically meaningful. In practice, all lengths are approximate, and we can as easily approximate $\sqrt{2}$ as any other number. Second, there is a simple geometric construction of the desired creation. Just stretch a string around a right isosceles triangle. Then by the Pythagorean theorem, the ratio of the hypotenuse to either side is precisely $\sqrt{2}$, so that plucking the corresponding strings should produce a perfect tritone.

A striking feature of this whole story is the leading role played by mathematicians and scientists in uncovering and elaborating the deep bonds between mathematics and the musical notions of consonance and dissonance. There is another connection between mathematics and a different aspect of music that lies much nearer the surface, and that can be described much more briefly. That aspect is the use of rhythm, and more particularly of *polyrhythm*.

By polyrhythm, we mean the practice of playing two lines of music in two different rhythms. For example, Brahms was especially fond of the combination "three against two", where one instrument plays three notes in the same time that another plays just two. Chopin frequently wrote pieces pitting four against three, and occasionally five against three.

Modern composers have ventured further and further in the same direction. One of them, Conlon Nancarrow, finally despaired of finding players who could handle the complexities of his rhythms and turned to composing for a pair of player pianos. Once freed from human limitations, he could make really daring leaps, culminating in his Study #33: "Canon—$\sqrt{2}/2$". It was his first one involving irrational tempo relations; each piano is moving along at a fixed speed, where the ratio of the two speeds is $\sqrt{2}/2$!

To understand better what that means, let us think of polyrhythms as being represented by a pair of metronomes, one of which we keep always at the same speed. The second one is then adjusted to give us the desired rhythm. For example, if we want a Brahmsian three against two, we simple set the second metronome to strike three times for each pair of beats of the first. Similarily, we can produce four against three, and so on. In each case, we start the two metronomes together, and we see that they repeatedly come back together after a certain number of beats. Note that the more complicated the ratios, the longer we have to wait before the two metronomes come back together. What happens when the tempos are related irrationally, as in Nancarrow's piece, is that, if they start together, then they will *never* come back together, no matter how long we wait.

The metronome analogy may be the most illuminating way there is to illustrate the difference between rational and irrational numbers. That the distinction should arise in both aspects of music—in harmony, comparing just intonation with equal temperament, and in rhythm, contrasting standard polyrhythms with Nancarrow's $\sqrt{2}/2$—may seem initially surprising. What is even more surprising is that both of these aspects turn out to be separate manifestations of one and the same phenomenon. To explain how that comes about, we must return once more to the subject of consonance/dissonance.

For the ancient Pythagoreans, numbers possessed a semi-mystical significance. The fact that numbers were revealed to underlie as distant-seeming a human activity as music greatly added to their sense of mystery and power. The Pythagoreans reveled in that sense of mystery, as did many scientists subsequently, most notably, perhaps, Johann Kepler. But other scientists, more in the Galilean tradition, wanted to find *physical* causes that would dispel the mystery and reveal the reasons behind the role played by numbers. In fact, Galileo himself, in *Two New Sciences*, was one of the first to explain the physical origins of mathematical ratios occurring in music. He pointed out that what really counts in determining what we perceive as the pitch of a note is the number of vibrations per second made by the string. If one divides a string in half, each half will vibrate twice as fast (assuming always that the tension remains the same) and we hear the note we identify as an octave above the original. In general, all other factors being equal, the number of vibrations will be proportional to the reciprocal of the length. Thus a ratio of lengths $3:2$ corresponds to the same ratio of frequencies of vibrations. Galileo uses this physical fact as the basis for a theory of consonance and dissonance. The musical interval of a fifth sounds consonant to our ear because the two series of vibrations in the air set up by the strings will stike our eardrums in a regular fashion, coming together on every third beat. The more complicated the ratios, the more irregular the effect on our eardrum, and the more dissonant our perception of the tones.

We now easily see the connection between musical intervals and polyrhythms. Two tuning forks tuned a fifth apart are nothing but a pair of metronomes beating three against two, speeded up by a factor of several hundred. Complicated rhythms representing ratios of large numbers, when speeded up produce more dissonant intervals, and Nancarrow's canon on $\sqrt{2}/2$ turns into a perfect augmented fourth.

As a final note, we should point out that the perception of consonance and dissonance changes from generation to generation and from culture to culture, and that we should not make the mistake of assuming that consonance is "good" or "beautiful" in music, and dissonance "bad" or undesirable. In fact, what we seem to respond to is a deliberate tension between the two. Perhaps it is appropriate to let Galileo have the last word. Writing

in *Two New Sciences*, he points out that *too* much consonance strikes us as "too bland, and lacks fire." What we seek in our music is just the right amount of dissonance. "This produces a tickling and teasing of the cartilage of the eardrum, so that the sweetness is tempered by a sprinkling of sharpness, giving the impression of being simultaneously sweetly kissed, and bitten."

Acknowledgement

I would like to thank Nancy Hechinger and Tom Lehrer, who read an earlier version of this paper and made helpful suggestions that have been incorporated here, and also Brian Osserman for preparing the illustrations.

7
Notes on elliptic curves

Karl Rubin

I. Elliptic curves

An *elliptic curve* is an equation

$$y^2 = x^3 + ax^2 + bx + c$$

where a, b, and c are rational numbers and the cubic polynomial

$$x^3 + ax^2 + bx + c$$

is not equal to $(x - r)^2(x - s)$ for any rational numbers r and s. A solution (x, y) of the equation is called a point on the elliptic curve, and it is called a *rational point* if x and y are rational numbers.

Examples

1. Take $a = 0$, $b = -1$, and $c = 0$. This gives the elliptic curve $y^2 = x^3 - x$ which was studied by Pierre de Fermat in the seventeenth century. Fermat proved that the only rational points on this elliptic curve are $(0, 0)$, $(1, 0)$, and $(-1, 0)$.

2. If $a = -1$, $b = 0$, and $c = 0$ then the equation $y^2 = x^3 - x^2$ is *not* an elliptic curve because $x^3 - x^2 = x^2(x - 1)$. Similarly $y^2 = x^3$ is *not* an elliptic curve.

3. Pick two rational numbers A and B and take the equation $y^2 = x(x - A)(x - B)$, which is the same as $y^2 = x^3 + (-A - B)x^2 + ABx$. This equation is an elliptic curve as long as A and B are different and not zero. If we take $A = 1$ and $B = -1$ we get Fermat's example again.

An elliptic curve is not an ellipse. The name comes from the fact that these equations first arose when people used calculus to measure the circumference of an ellipse.

II. Finding rational points on elliptic curves

There are several ways to use rational points on an elliptic curve to find more rational points. Fix an elliptic curve and call it E.

a) Suppose $P = (x, y)$ is a rational point on E. Then $Q = (x, -y)$ is also a rational point on E.

b) Suppose P_1 and P_2 are two different points on E. Draw the line through P_1 and P_2 (see Figure 1). This line intersects E at exactly one additional rational point Q, and if P_1 and P_2 have rational coordinates then Q will also have rational coordinates. (Actually, there are possible exceptions: if the line is tangent to E at either P_1 or P_2, or if the line is vertical, then there will be no third point.)

c) Suppose P is a rational point on E. Draw the line which is tangent to E at P (see Figure 1). This line will intersect the elliptic curve at exactly one other point Q, and if P has rational coordinates then Q will also have rational coordinates. (Again there are some exceptional situations where there will be no additional intersection point Q. For example, this happens if the x-coordinate of P is 0, so the tangent line is vertical.)

Examples

1. Take E to be the elliptic curve $y^2 = x^3 - 36x$, which has rational points $P_1 = (0, 0)$ and $P_2 = (-3, 9)$. If we apply method (a) to P_1 we get the same point back. If we apply (a) to P_2 we get the new point $(-3, -9)$. Now let's use method (b). The equation of the line through P_1 and P_2 is $y = -3x$. If we solve for the intersection of this line with E we get

$$(-3x)^2 = x^3 - 36x, \quad \text{or} \quad x^3 - 9x^2 - 36x = 0,$$
$$\text{or} \quad x(x+3)(x-12) = 0.$$

This solutions $x = 0, -3$ correspond to P_1 and P_2, but the third solution $x = 12$ gives us a new rational point $Q = (12, -36)$ on E. If we apply method (c) with P_1 we get no new point because the tangent line is vertical. But we can apply (c) with P_2 to get a new point. Calculus shows that the slope of the tangent line to E at P_2 is $-1/2$, so the equation of the tangent line is $y = -x/2 + 15/2$.

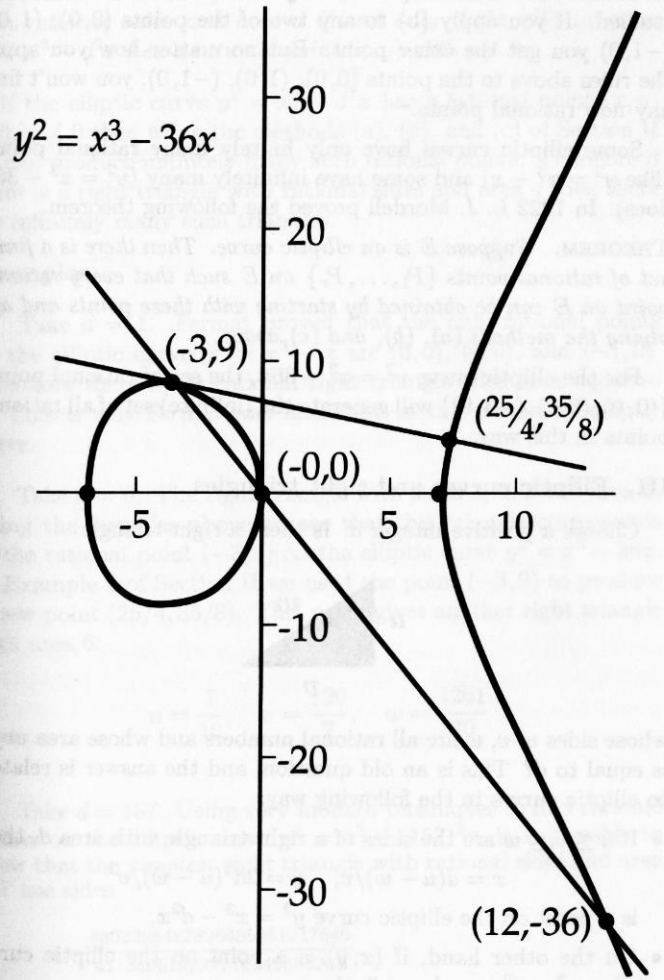


FIGURE 1. $y^2 = x^3 - 36x$

Solving as above we see that this line intersects E at P_2 and at the new rational point $(25/4, 35/8)$.

2. Take E to be the elliptic curve $y^2 = x^3 - x$, the curve Fermat studied. If you apply (b) to any two of the points $(0,0)$, $(1,0)$, $(-1,0)$ you get the other point. But no matter how you apply the rules above to the points $(0,0)$, $(1,0)$, $(-1,0)$, you won't find any new rational points.

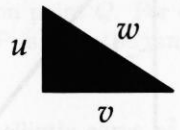
Some elliptic curves have only finitely many rational points (like $y^2 = x^3 - x$) and some have infinitely many ($y^2 = x^3 - 36x$ does). In 1922 L. J. Mordell proved the following theorem.

THEOREM. *Suppose E is an elliptic curve. Then there is a finite set of rational points $\{P_1, \dots, P_r\}$ on E such that every rational point on E can be obtained by starting with these points and applying the methods (a), (b), and (c) above.*

For the elliptic curve $y^2 = x^3 - 36x$, the set of rational points $\{(0,0), (6,0), (-3,9)\}$ will generate the (infinite) set of all rational points in this way.

III. Elliptic curves and right triangles

Choose a positive integer d. Is there a right triangle



whose sides u, v, w are all rational numbers and whose area $uv/2$ is equal to d? This is an old question, and the answer is related to elliptic curves in the following way:

- If $u \le v \le w$ are the sides of a right triangle with area d, then
$$x = d(u-w)/v, \quad y = 2d^2(u-w)/v^2$$
is a point on the elliptic curve $y^2 = x^3 - d^2x$.
- On the other hand, if (x,y) is a point on the elliptic curve $y^2 = x^3 - d^2x$, and $y \neq 0$, then
$$u = |(x^2 - d^2)/y|, \quad v = |2dx/y|, \quad w = |(x^2 + d^2)/y|$$
are the sides of a right triangle with area d. (These facts are not hard to check if you remember that the triangle with sides $u \le v \le w$ is a right triangle if and only if $u^2 + v^2 = w^2$.)

This proves the following statement:

THEOREM. *Choose a positive integer d. There is a right triangle with rational sides and area d if and only if the elliptic curve $y^2 = x^3 - d^2x$ has a rational point (x,y) with $y \neq 0$.*

If the elliptic curve $y^2 = x^3 - d^2x$ has a rational point (x,y) with $y \neq 0$ then using the methods (a), (b), and (c) of Section II we can produce infinitely many such rational points. Therefore if there is a right triangle with rational sides and area d then there are infinitely many such triangles.

Examples

1. Take $d = 1$. Fermat proved that the only rational points on the elliptic curve $y^2 = x^3 - x$ are $(0,0)$, $(1,0)$, and $(-1,0)$. Therefore there is *no* rational right triangle with area equal to 1. This is why Fermat was interested in this particular elliptic curve.

2. Take $d = 6$. The right triangle with sides 3, 4, 5 has area 6. Using the formulas above we see that this triangle corresponds to the rational point $(-3,9)$ on the elliptic curve $y^2 = x^3 - 36x$. In Example 1 of Section II we used the point $(-3,9)$ to produce a new point $(25/4, 35/8)$. This point gives another right triangle with area 6:
$$u = \frac{7}{10}, \quad v = \frac{120}{7}, \quad w = \frac{1201}{70}$$

3. Take $d = 157$. Using very modern techniques to find rational points on the elliptic curve $y^2 = x^3 - (157)^2x$, it is possible to show that the *simplest* right triangle with rational sides and area 157 has sides
$$u = \frac{6803298487826435051217540}{411340519227716149383203},$$
$$v = \frac{411340519227716149383203}{2166655569371476130961 0},$$
$$w = \frac{224403517704336969924557513090674863160948472041}{8912332268928859588025535178967163570016480830}.$$

IV. Reduction modulo primes

Return now to a general elliptic curve. That is, we fix integers a, b, and c and consider the elliptic curve $y^2 = x^3 + ax^2 + bx + c$ which we call E. In general it is very difficult to decide if E has any rational points, or whether the number of rational points on E is finite or infinite. It turns out that one can try to answer these questions by asking a lot of simpler questions. Instead of asking

how often is $y^2 - (x^3 + ax^2 + bx + c)$ equal to zero?

we can ask

how often is $y^2 - (x^3 + ax^2 + bx + c)$ a multiple of 5?

There is nothing special about 5; we could ask the same question for any positive integer. We will be interested in the answer with 5 replaced by any prime number. For every prime number p we define

N_p = the number of pairs of integers x, y in the range $0 \le x, y \le p-1$ such that $y^2 - (x^3 + ax^2 + bx + c)$ is a multiple of p.

We only consider x and y in this range because if you add a multiple of p to x or y it also changes $y^2 - (x^3 + ax^2 + bx + c)$ by a multiple of p.

Examples

1. Take E to be $y^2 = x^3 - 36x$ and p=5. We can make a table of values

y \ x	0	1	2	3	4
0	**0**	**35**	64	81	**80**
1	1	36	**65**	82	81
2	4	39	68	**85**	84
3	9	44	73	**90**	89
4	16	51	**80**	97	96

$y^2 - (x^3 - 36x)$

Seven of the entries in the table are multiples of 5, so $N_5 = 7$. One of the entries is equal to zero, corresponding to the fact that $(0,0)$ is actually a point on E. Also, $(1,0)$, $(4,0)$, and $(2,4)$ give multiples of 5; this is not surprising because the rational points $(6,0)$, $(-6,0)$, and $(-3,9)$ (which would give zero if we extended the table that far) differ from these three points by multiples of 5.

2. For the elliptic curve $y^2 = x^3 - x$, C. F. Gauss proved in 1814 the following formula for N_p:

   (1) if p is one less than a multiple of 4 then $N_p = p$;

   (2) if p is one more than a multiple of 4 then write $p = r^2 + s^2$ with integers r and s. Switch r and s if necessary so that r is odd, and replace r by $-r$ if necessary so that $r + s - 1$ is a multiple of 4 (we can always do this). Then $N_p = p - 2r$.

(For example: for $p = 29$, take $r = -5$ and $s = 2$, so $N_{29} = 29 - 2(-5) = 39$.)

The first example above illustrates an idea which goes back to two British mathematicians B. J. Birch and H. P. F. Swinnerton-Dyer in the late 1950's. Namely, the more rational points E has, the larger the numbers N_p should be *on average*. They did a lot of computer calculations and were led to the following precise conjecture. If M is a positive integer write $S(M)$ for the product of the values $(N_p + 1)/p$ as p runs through all primes less than or equal to M:

$$S(M) = \prod_{p \le M} \frac{N_p + 1}{p}.$$

CONJECTURE (Birch and Swinnerton-Dyer). *If E has only finitely many rational points then $\lim_{M\to\infty} S(M) < \infty$ (that is, $S(M)$ approaches a finite limit as M gets large). If E has infinitely many rational points then*

$$\lim_{M\to\infty} S(M) = \infty$$

(that is, $S(M)$ can be made arbitrarily large by taking M large enough).

This statement is still a conjecture because no one has been able to prove it yet. But there is a great deal of evidence for

it, and some related statements have been proved. If true, this conjecture tells us that the answer to the question of whether the number of rational points on E is finite or infinite is contained in the sequence of numbers $N_2, N_3, N_5, \ldots$.

EXAMPLE. In Figure 2 we plot the values $S(M)$ for the curves $y^2 = x^3 - d^2x$ with $d = 1, 2, 3, 5, 6$, and 7 and M up to 15,000,000. The values are very irregular, but it is true that E has only finitely many rational points for $d = 1, 2, 3$ and infinitely many for $d = 5, 6, 7$. (We skipped $d = 4$ because it is the same, after a simple change of variables, as the curve with $d = 1$.)

Things To Try Yourself

1. We saw in Section I that $y^2 = x^3 - x^2$ is not an elliptic curve. Draw the graph of this equation, and pay particular attention to the graph near the point $(0,0)$. This equation is called singular because the graph crosses itself at the point $(0,0)$. Now draw the graph of $y^2 = x^3 - x^2 + c$ for some small (positive and negative) values of c, and see what happens. If c is not zero then $y^2 = x^3 - x^2 + c$ is an elliptic curve.

2. Use a computer and the methods of Section II to find more rational points on the elliptic curve $y^2 = x^3 - 36x$. Using these new points find more right triangles with rational sides and area 6.

3. Find some right triangles with rational sides and area 5 and some with area 7.

4. (This is harder) Show that the only pairs of *integers* (positive or negative whole numbers) (x, y) satisfying $y^2 = x^3 - x$ are $(0,0)$, $(1,0)$, and $(-1,0)$. (For rational numbers x, y the argument is much more difficult.) Can you find all the integer solutions for some other elliptic curves?

5. Use Gauss' formula from Section IV to compute the numbers N_p for the elliptic curve $y^2 = x^3 - x$ with some primes $p > 100$. Write a computer program to compute N_p using Gauss' formula and another to compute N_p using the counting definition, and compare the amount of time the two methods take for some very large primes p.

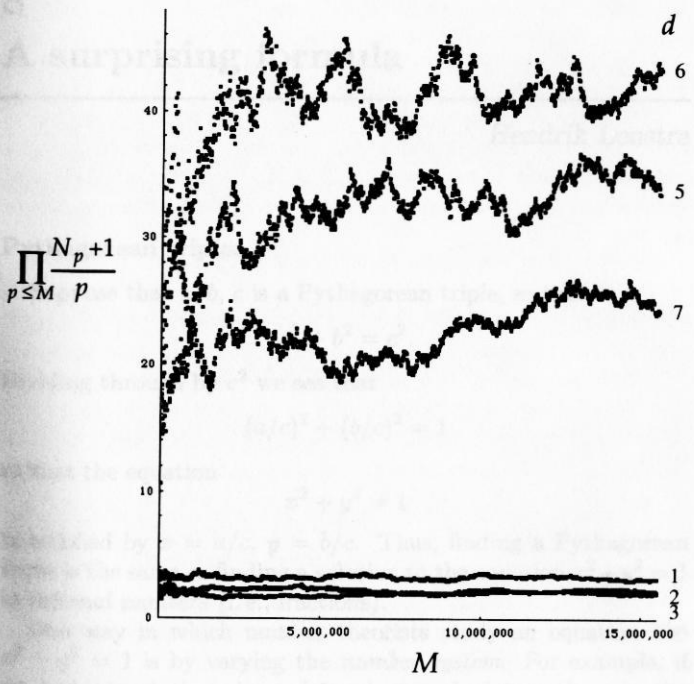


FIGURE 2

References for further reading

- CASSELS, J. W. S., *Lectures on elliptic curves*, London Math. Soc. Student Texts, vol. 24, Cambridge University Press, Cambridge, 1991.
- KOBLITZ, N, *Introduction to elliptic curves and modular forms*, Graduate Texts in Math., vol. 97, Springer-Verlag, New York, 1992.

• SILVERMAN, J., AND TATE, J., *Rational points on elliptic curves*, Undergraduate Texts in Math., Springer-Verlag, New York, 1992.

8
A surprising formula

Hendrik Lenstra

Pythagorean triples

Suppose that a, b, c is a Pythagorean triple, so that

$$a^2 + b^2 = c^2.$$

Dividing through by c^2 we see that

$$(a/c)^2 + (b/c)^2 = 1$$

so that the equation

$$x^2 + y^2 = 1$$

is satisfied by $x = a/c$, $y = b/c$. Thus, finding a Pythagorean triple is the same as finding a solution to the equation $x^2+y^2 = 1$ in *rational numbers* (i. e., fractions).

One way in which number theorists study an equation like $x^2 + y^2 = 1$ is by varying the *number system*. For example, if we look at solutions in *real numbers* and picture them in the (x, y)-plane, then we obtain a *circle*.

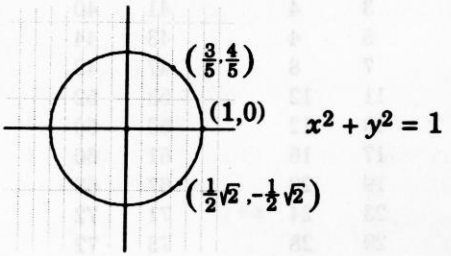


FIGURE 1

Examples of points on this circle are

$$x = 1,\ y = 0; \qquad x = 3/5,\ y = 4/5;$$
$$x = \frac{1}{2}\sqrt{2},\ y = -\frac{1}{2}\sqrt{2}.$$

Other number systems in which we can solve the equation are the *numbers modulo p*, for prime numbers p. Take for example $p = 5$. One obtains numbers modulo 5 by ignoring multiples of 5 ("keep the pennies, throw out the rest"). There are just five numbers modulo 5, namely 0, 1, 2, 3, 4. One does addition and multiplication by discarding multiples of 5. So for these numbers one has

$$2 + 4 = 6 = 1, \qquad 3 \times 4 = 12 = 2.$$

The equation $x^2 + y^2 = 1$ is now satisfied by $x = 4$, $y = 0$:

$$4^2 + 0^2 = 16 + 0 = 1 + 0 = 1.$$

Checking all possibilities for x and y one finds that the only other solutions are given by

$$x = 1,\ y = 0; \qquad x = 0,\ y = 1; \qquad x = 0,\ y = 4.$$

Altogether there are 4 solutions.

In the same way one can count the solutions in numbers modulo p for other prime numbers p. Say that there are N_p solutions, so that $N_5 = 4$. One finds the following table:

p	N_p	p	N_p
2	2	37	36
3	4	41	40
5	4	43	44
7	8	47	48
11	12	53	52
13	12	59	60
17	16	61	60
19	20	67	68
23	24	71	72
29	28	73	72
31	32	79	80

Looking at this table one discovers a regularity. It seems to be the case that for every prime number $p > 2$ one has $N_p = p - 1$ or $N_p = p + 1$, whichever of these is divisible by 4. It can be proved that this is actually true for all prime numbers p.

There is a surprising connection between the "sizes" of the solution sets in the different number systems that we considered. The size of the set of real solutions is to be interpreted as the length of the circumference of the circle, which is 2π. For a prime number p, we measure the size of the solution set by means of the number N_p/p, which is the fraction by which N_p differs from p. If we multiply all these fairly complicated numbers together one finds an unexpectedly simple answer:

$$2\pi \cdot \frac{N_2}{2} \cdot \frac{N_3}{3} \cdot \frac{N_5}{5} \cdot \frac{N_7}{7} \cdot \cdots = 2\pi \cdot 1 \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \frac{8}{7} \cdot \cdots = 8.$$

That is to say, if we take more and more factors in the product we get closer and closer to 8. For example, if one stops at $p = 13$, then one gets $7.713\ldots$, and if one stops at $p = 79$ one gets $7.953\ldots$. If a single prime number > 2 is omitted then the formula fails to hold! The formula reflects that the solution sets in all those different number systems are linked in some mysterious manner.

One may wonder what the significance of the number 8 is. It is due to a symmetry in the equation: a typical solution x, y to $x^2 + y^2 = 1$ gives rise to *eight* solutions:

$$x, y; \quad -x, y; \quad x, -y; \quad -x, -y;$$
$$y, x; \quad -y, x; \quad y, -x; \quad -y, -x.$$

There are many types of equations for which similar phenomena are known or expected to take place.

9
References for further reading

Robert Osserman and Ken Ribet

Popular Math Books

- JOHN D. BARROW, *Pi in the Sky: Counting, Thinking, and Being*, Clarendon Press, Oxford, 1992.

  A cultural/historical approach to matters mathematical, with interesting sidelights.

- E. T. BELL, *Men of Mathematics*, Simon & Schuster, New York, 1986, Touchstone paperback.

  The classical collection of biographies of outstanding mathematicians through the ages, including Fermat, Euler, Gauss, and Kummer.

- E. T. BELL, *The Last Problem*, revised and updated by Underwood Dudley, MAA, 1990.

  A wide-ranging discourse on topics related one way or another to Fermat's Last Theorem.

- PHILIP J. DAVIS AND REUBEN HERSH, *The Mathematical Experience*, Birkhäuser, Boston, 1981.

  An overview of the practice and uses of mathematics, as well as underlying issues.

- KEITH DEVLIN, *Mathematics: the New Golden Age*, Penguin, London, New York, 1988.

  Covers a number of topics of current interest, including a section on Fermat's Last Theorem (before the proof).

- WILLIAM DUNHAM, *Journey Through Genius: The Great Theorems of Mathematics*, Wiley, New York, 1990.

Twelve chapters, each describing a major advance in mathematics, along with historical and cultural remarks setting each one in the context of its time.

- MARTIN GARDNER, *Penrose Tiles to Trapdoor Ciphers*, W. H. Freeman, New York, 1988.

  One of the many collections from Martin Gardner's long-running column in the Scientific American. This one is particularly good, and has additional chapters updating the original columns on Penrose tilings and public key cryptography. An earlier collection by Martin Gardner, *Wheels, Life and other Mathematical Amusements*, W. H. Freeman, New York, 1983, includes a chapter "Diophantine Analysis and Fermat's Last Theorem" and three chapters on "The Game of Life", invented by John Conway.

- TERI PERL, *Math Equals: Biographies of Women Mathematicians & Related Activities*, Addison-Wesley, Menlo Park, 1978.

  Recounts the lives and work of a number of eminent women mathematicians, including Sophie Germain.

Books on number theory, with a historical approach

- HAROLD M. EDWARDS, *Fermat's Last Theorem: A genetic introduction to algebraic number theory*, Springer-Verlag, New York, 1977.
- ANDRÉ WEIL, *Number Theory: an approach through history from Hammurapi to Legendre*, Birkhäuser, Boston, 1983.

The computational approach

- J. BUHLER, R. CRANDALL, R. ERNWALL, AND T. METSÄNKYLÄ, *Irregular primes and cyclotomic invariants to four million*, Mathematics of Computation **61** (July 1993), 151–153.

  This is the last of the giant computations carried out before Wiles proved Fermat's Last Theorem for all exponents n. In this paper, the theorem is verified for all n up to $4,000,000$. An earlier paper does it for n up to $1,000,000$ and gives more information:

- J. BUHLER, R. CRANDALL, AND R. W. SOMPOLSKI, *Irregular primes to one million*, Mathematics of Computation **59** (October 1992), 717–722.

More Technical Books and Articles

- J. W. S. CASSELS, *Diophantine equations with special reference to elliptic curves*, Survey article, Journal London Math. Soc. **41** (1966), 193–291.
- J. W. S. CASSELS, *Lectures on elliptic curves*, Cambridge Univ. Press, Cambridge and New York, 1991.
- J. CHAHAL, *Topics in number theory*, Plenum Press, New York and London, 1988.
- D. A. COX, *Primes of the form $x^2 + ny^2$: Fermat, class field theory, and complex multiplication*, Wiley, New York, 1989.
- D. HUSEMOLLER, *Elliptic curves*, Graduate Texts in Math., vol. 111, Springer-Verlag, Berlin and New York, 1987.
- K. F. IRELAND AND M. I. ROSEN, *A classical introduction to modern number theory*, Graduate Texts in Math., second edition, vol. 84, Springer-Verlag, Berlin and New York, 1990.
- A. W. KNAPP, *Elliptic curves*, Mathematical Notes, vol. 40, Princeton University Press, Princeton, 1992.
- N. KOBLITZ, *Introduction to elliptic curves and modular forms*, Graduate Texts in Math., vol. 97, Springer-Verlag, Berlin and New York, 1984.
- S. LANG, *Elliptic curves diophantine analysis*, Grundlehren der mathematischen Wissenschaften, vol. 231, Springer-Verlag, Berlin and New York, 1978.
- J. H. SILVERMAN, *The arithmetic of elliptic curves*, Graduate Texts in Math., vol. 106, Springer-Verlag, Berlin and New York, 1986.
- J. H. SILVERMAN AND J. T. TATE, *Rational points on elliptic curves*, Undergraduate Texts in Math., Springer-Verlag, Berlin and New York, 1992.
- J. T. TATE, *The arithmetic of elliptic curves*, Inventiones Math. **23** (1974), 179–206.
- M. WALDSCHMIDT ET AL., editors., *From number theory to physics*, Lectures given at the meeting "Number Theory and

Physics" held at the Centre de Physique, Les Houches, 1989, Springer-Verlag, Berlin and New York, 1992.

Here is the article in which Frey associates elliptic curves with solutions of Fermat's equation:

- G. FREY, *Links between stable elliptic curves and certain diophantine equations*, Annales Universitatis, Saraviensis **1** (1986), 1–40.

Conjectures in number theory and their relation with Fermat's Last Theorem:

- S. LANG, *Old and new conjectured Diophantine inequalities*, Bull. Amer. Math. Soc. **23** (1990), 37–75.

Two articles by K. Ribet proving that Frey's elliptic curves cannot be modular:

- K. A. RIBET, *On modular representations of* $\mathrm{Gal}(\bar{\mathbf{Q}}/Q)$ *arising from modular forms*, Inventiones Math. **100** (1990), 431–476.
- K. A. RIBET, *From the Taniyama-Shimura Conjecture to Fermat's Last Theorem*, Annales de la Faculté des Sciences de l'Université de Toulouse **11** (1990), 116–139.

An explanation of Taniyama's conjecture from a down-to-earth point of view:

- B. MAZUR, *Number theory as gadfly*, Amer. Math. Monthly **98** (1991), 593–610.

Authors' Afiiliations

Joe Buhler
Reed College

Hendrik Lenstra
University of California at Berkeley

Robert Osserman
Stanford University and Mathematical Sciences Research Institute

Ken Ribet
University of California at Berkeley

Karl Rubin
Ohio State University

Credits

Supplement Editor:
Robert Osserman

Fermat Fest and Video Production Director:
Arlene Baxter

Production Assistant:
Stephen R. Stapleton

Video Producer:
gayle k. yamada

Fermat Fest Planning Committee:
Arlene Baxter, Lenore Blum, Joe Buhler, Lee Dembart, Will Hearst, Thomas Humphrey, Hendrik Lenstra, Robert Osserman, Ken Ribet, Karl Rubin, Alice Silverberg, and William P. Thurston

The closing song "That's Mathematics" written and performed by Tom Lehrer.

Special thanks to:

Irving Kaplansky
The support staff of MSRI
The shop crew of the Exploratorium for props

MSRI wishes to acknowledge generous support from:

The Paul and Gabriella Rosenbaum Foundation for production of the video and supplement
Will Hearst for closed captioning

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX : 3 January 1994
Prepared by [QUANTEXT, Intl.]