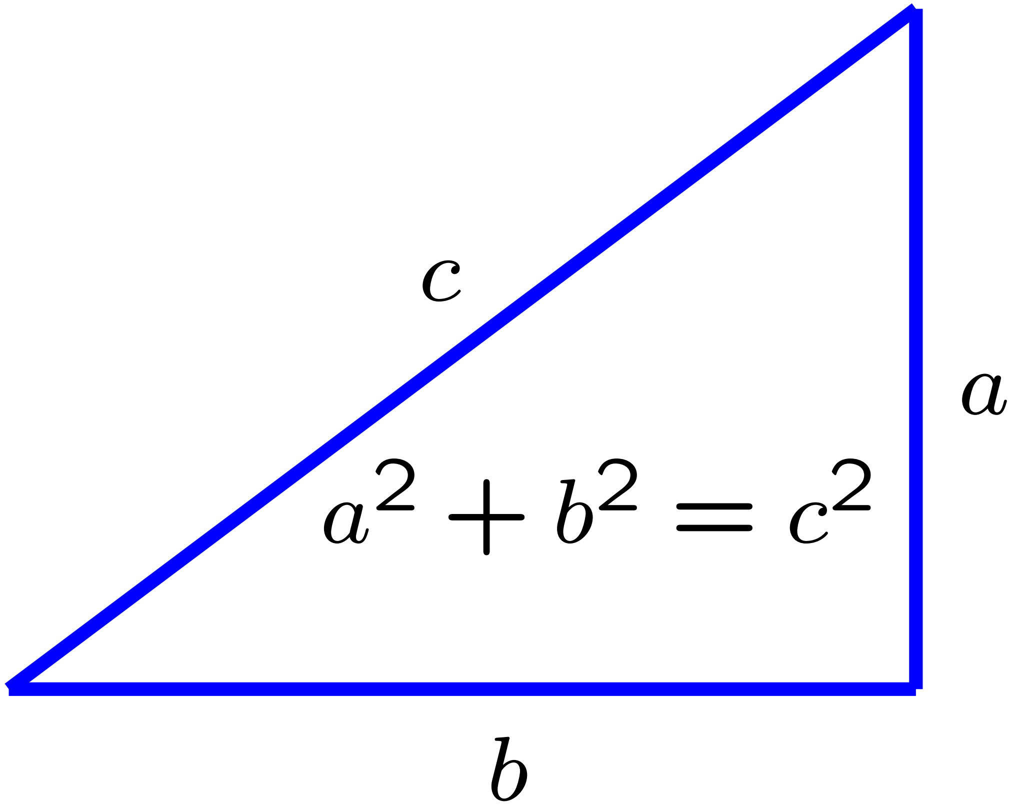
An Introduction to the Modular Forms Database Project:

My Dream Computation (not a toy problem!)

William A. Stein
Associate Professor of Mathematics
University of California, San Diego

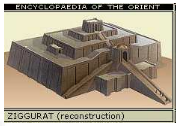
San Diego Supercomputing Center: August 3, 2005

The Pythagorean Theorem



Pythagoras
Approx 569–475BC

Pythagorean Triples

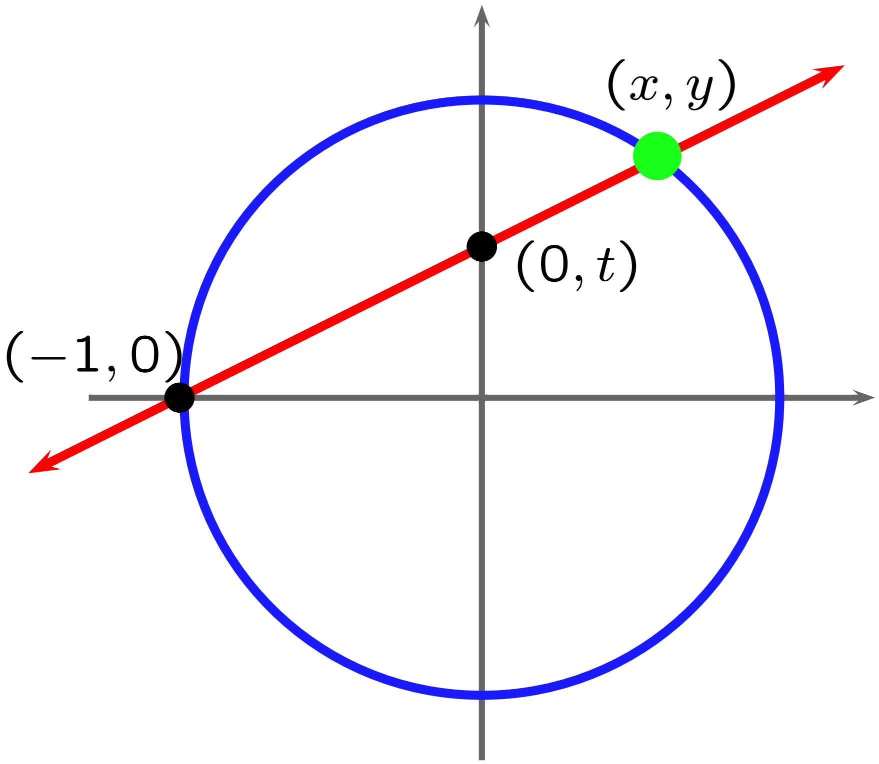


$(3, 4, 5)$
$(5, 12, 13)$
$(7, 24, 25)$
$(9, 40, 41)$
$(11, 60, 61)$
$(13, 84, 85)$
$(15, 8, 17)$
$(21, 20, 29)$
$(33, 56, 65)$
$(35, 12, 37)$
$(39, 80, 89)$
$(45, 28, 53)$
$(55, 48, 73)$
$(63, 16, 65)$
$(65, 72, 97)$
$(77, 36, 85)$
$\vdots$

Triples of integers a, b, c such that

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

Enumerating Pythagorean Triples



$$\text{Slope} = t = \frac{y}{x+1}$$

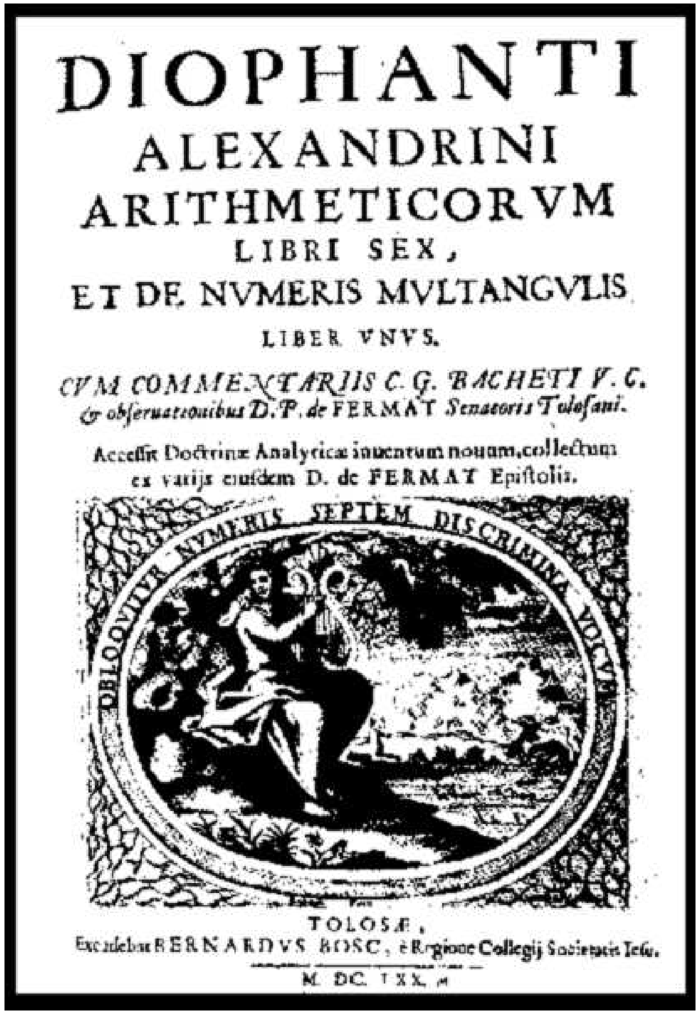
$$x = \frac{1-t^2}{1+t^2}$$

$$y = \frac{2t}{1+t^2}$$

If $t = \frac{r}{s}$, then $a = s^2 - r^2, \quad b = 2rs, \quad c = s^2 + r^2$
is a Pythagorean triple, and all primitive unordered triples arise in this way.

Fermat's "Last Theorem"

No "Pythagorean triples" with exponent 3 or higher.

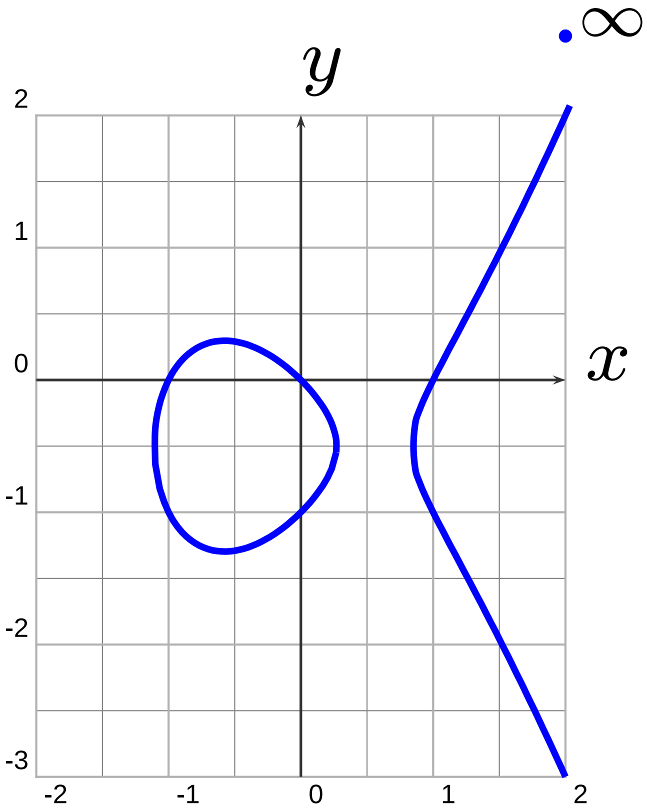


DIOPHANTI
ALEXANDRINI
ARITHMETICORVM
LIBRI SEX,
ET DE NVMERIS MVLTANGVLIS
LIBER VNVS.
CVM COMMENTARIIS C. G. BACHETI V. C.
& obseruationibus D. P. de FERMAT Senatoris Tolosani.
Accessit Doctrinæ Analyticæ inuentum nouum, collectum
ex varijs eiusdem D. de FERMAT Epistolis.

TOLOSÆ,
Excudebat BERNARDVS BOSC, è Regione Collegij Societatis Iesu.
M. DC. LXX.

Wiles's Proof of FLT Uses Elliptic Curves

An **elliptic curve** is a nonsingular plane cubic curve with a rational point (possibly "at infinity").



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

EXAMPLES

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

$$x^3 + y^3 = 1 \text{ (Fermat cubic)}$$

$$y^2 = x^3 + ax + b$$

$$\xcancel{3x^3 + 4y^3 + 5 = 0}$$

The Frey Elliptic Curve

Suppose Fermat's conjecture is **FALSE**. Then there is a prime $\ell \geq 5$ and coprime positive integers a, b, c with $a^\ell + b^\ell = c^\ell$.

Consider the corresponding Frey elliptic curve:

$$y^2 = x(x - a^\ell)(x + b^\ell).$$

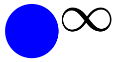
Ribet's Theorem: This elliptic curve is not *modular*.

Wiles's Theorem: This elliptic curve is *modular*.

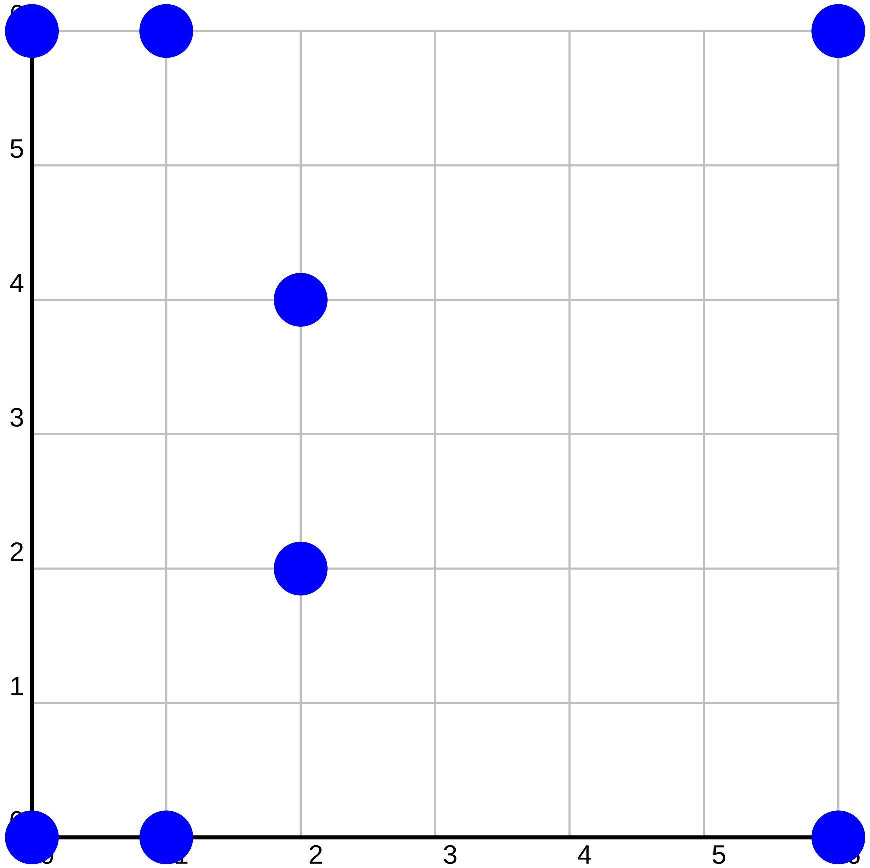
Conclusion: Fermat's conjecture is true.

Counting Solutions Modulo p

$$N(p) = \#\text{ of solutions } \pmod{p}$$



$$y^2 + y = x^3 - x \pmod{7}$$

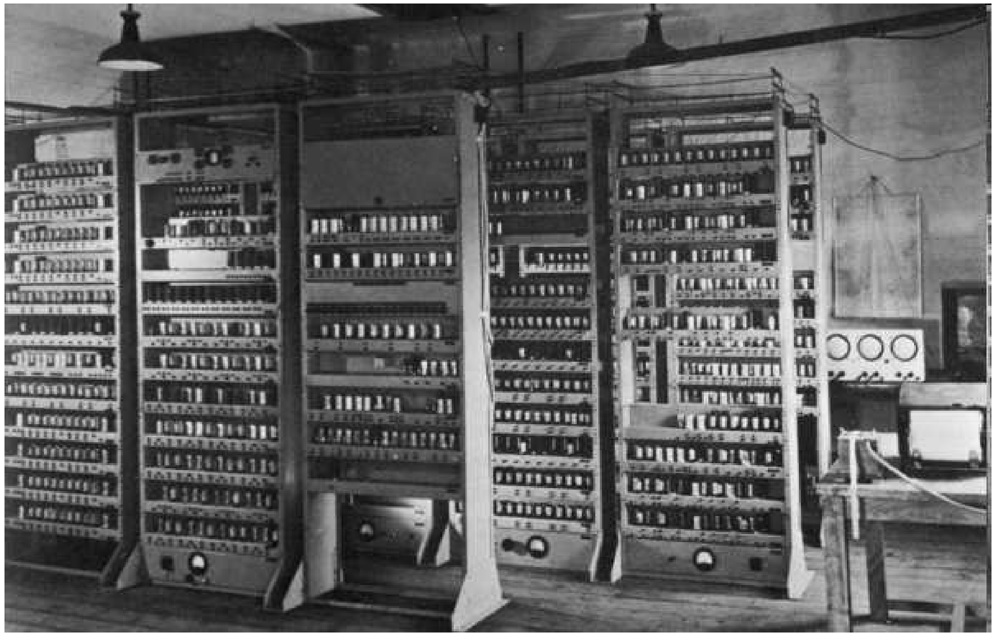


$$N(7) = 9$$

Counting Points

Cambridge EDSAC: The first point counting supercomputer...

Birch and Swinnerton-Dyer



The Hecke Eigenvalue

Hasse

Let

$$a_p = p + 1 - N(p).$$

Hasse proved that

$$|a_p| \leq 2\sqrt{p}.$$

For $y^2 + y = x^3 - x$:

$$a_2 = -2, \quad a_3 = -3, \quad a_5 = -2, \quad a_7 = -1, \quad a_{11} = -5, \quad a_{13} = -2,$$

$$a_{17} = 0, \quad a_{19} = 0, \quad a_{23} = 2, \quad a_{29} = 6, \quad \ldots$$

Elliptic Curves are "Modular"

An elliptic curve is **modular** if the numbers a_p are coefficients of a "modular form".

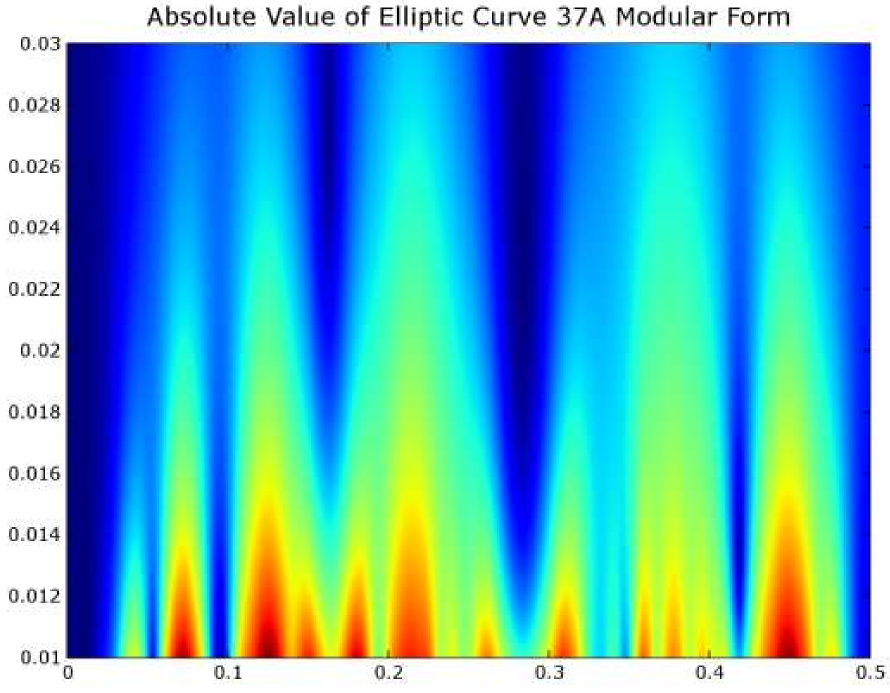
Theorem (Wiles et al.): *Every elliptic curve over the rational numbers is modular.*

Wiles at the Institute for Advanced Study

Modular Forms

The definition of modular forms as holomorphic functions satisfying a certain equation is very abstract.

I will skip the abstract definition, and instead give you an explicit "engineer's recipe" for producing modular forms. In the meantime, here's a picture:



Computing Modular Forms: Motivation

Motivation: Data about modular forms is **extremely** useful to many research mathematicians (e.g., number theorists, cryptographers). This data is like the astronomer's telescope images.

I want to compute modular forms on a **huge** scale using the SDSC resources, and make the resulting database widely available. I have done this on a small scale during the last 5 years — see `http://modular.fas.harvard.edu/Tables/`

What to Compute: Newforms

For each positive integer N there is a finite list of **newforms** of level N. E.g., for $N = 37$ the newforms are

$$f_1 = q - 2q^2 - 3q^3 + 2q^4 - 2q^5 + 6q^6 - q^7 + \cdots$$
$$f_2 = q + q^3 - 2q^4 - q^7 + \cdots,$$

where $q = e^{2\pi i z}$.

The newforms of level N determine all the modular forms of level N (like a basis in linear algebra). The coefficients are algebraic integers. *Goal: compute these newforms.*

Bad idea – write down many elliptic curves and compute the numbers a_p by counting points over finite fields. No good – this misses most of the interesting newforms, and gets newforms of all kinds of random levels, but you don't know if you get everything of a given level.

An Engineer's Recipe for Newforms

Fix our positive integer N. For simplicity assume that N is prime.

1. Form the $N+1$ dimensional $\mathbf{Q}$-vector space V with basis the symbols $[0],\ldots,[N-1],[\infty]$.

2. Let R be the suspace of V spanned by the following vectors, for $x = 0,\ldots,N-1,\infty$:

$$[x] - [N-x]$$
$$[x] + [x.S]$$
$$[x] + [x.T] + [x.T^2]$$

$S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $T = \left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$, and $x.\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = (ax+c)/(bx+d)$.

3. Compute the quotient vector space $M = V/R$. This involves "intelligent" **sparse Gauss elimination** on a matrix with $N+1$ columns.

4. Compute the matrix T_2 on M given by

$$[x] \mapsto [x.\left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)] + [x.\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1 \end{smallmatrix}\right)] + [x.\left(\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right)] + [x.\left(\begin{smallmatrix} 1 & 0 \\ 1 & 2 \end{smallmatrix}\right)].$$

This matrix is unfortunately not sparse. Similar recipe for matrices T_n for any n.

5. Compute the **characteristic polynomial** f of T_2.

6. **Factor** $f = \prod g_i^{e_i}$. Assume all $e_i = 1$ (if not, use a random linear combination of the T_n.)

7. Compute the **kernels** $K_i = \ker(g_i(T_2))$. The **eigenvalues** of T_3, T_5, etc., acting on an **eigenvector** in K_i give the coefficients a_p of the newforms of level N.

Implementation

- I implemented code for computing modular forms that's included with **MAGMA**:
`http://magma.maths.usyd.edu.au/magma/`.

- Unfortunately, MAGMA is expensive and closed source, so I'm reimplementing everything as part of **SAGE**:
`http://modular.fas.harvard.edu/sage/`.

- I'm teaching a **course** on this topic at UCSD this Fall.

- I'm finishing a **book** on these algorithms that will be published by the American Mathematical Society.

The Modular Forms Database Project

- Create a database of all newforms of level N for each $N < 100000$. This will require many gigabytes to store. (50GB?)

- So far this has only been done for $N < 7000$ (and is incomplete), so 100000 is a **major challenge**.

- Involves sparse linear algebra over $\mathbf{Q}$ on spaces of dimension up to 200000 and dense linear algebra on spaces of dimension up to 25000.

- Easy to parallelize – run one process for each N.

- Will be very useful to number theorists and cryptographers.

- John Cremona has done something similar but only for the newforms corresponding to elliptic curves (he's at around 84000 right now).