A brief history of elliptic curves

Paul Hewitt

December 5, 2005

Elliptic curves are beautiful and serendipitous, ancient and ubiquitous. Again and again they turn up in the most astonishing places. Here I offer the briefest historical survey, a mere hint of their surprising power and elegance.

Diophantus' world

Our story begins with Diophantus... or at least it should. Unfortunately we don't know much about him! We don't know where he came from and we aren't even sure when he lived. The best guess is that he lived and worked some time in the second or third century. We *do* know where he worked: the famed Library of Alexandria. So, we first put Diophantus' work in political and cultural context.

The city of Alexandria, at the mouth of the Nile, was founded by Alexander the Great, in 331 BCE. For three centuries it was ruled by the descendents of one of his Greek generals, Ptolemy. The Ptolemys built one of the most magnificent cities of the ancient world, a center for commerce, learning, and the arts. Part of this cultural flowering was the "Library", which was in fact a center for scientific research. Among the many famous researchers who were employed there were the mathematicians Euclid and Apollonius, and the astronomer Claudius Ptolemy (no relation to the rulers).

When Julius Caesar conquered Egypt in the year 31 BCE, he brought Alexandria under Roman control. He also set fire to the Library. However, the Library survived, and the Roman Emperors left most administration to the locals. Greek remained the language of the urban, learned classes for many centuries, but tensions ran deep.

After the crucifixion of Jesus and the destruction of the Temple in Jerusalem, Alexandria was home to both a burgeoning Christian movement, divided against one another in many sects, and violently persecuted by Rome; and also a flood of Jewish refugees from Palestine. It was at times a volatile center of religious conflict.

So Diophantus was probably part of the Greek-speaking, increasingly Christian, eastern Roman Empire. This was the Age of Empire: to the east of the Roman Empire lay the Parthian Empire of Iran and Afghanistan; to the east of that lay the Kushan Empire of Pakistan and northern India; and further east still lay the Chinese Empire (the Chin Dynasty). Altho the great empires were often at war, over all this was a period of relative stability. Hence this was also the first great era of international commerce. With the trade in commodities went the trade in ideas. (And also diseases: most of these societies were eventually weakened to the point of collapse by the epidemics emanating from India and southern China, and spread along the trade routes.)

Thus, we place Diophantus' work centuries after the "Golden Age" of Euclid, Apollonius, and Archimedes (who apparently studied in Alexandria, but then returned to his native Sicily). Diophantus may have been at the leading edge of what is sometimes called the "Silver Age", which culminated with the work of Pappus, perhaps a century later. A century after Pappus saw the effective end of the era of Greek mathematics, when Christianity became the official religion of the eastern Roman Empire. In 415 Hypatia, the last chief librarian in Alexandria, was set upon by a Christian mob and killed.

Diophantus' legacy

Mathematics at the Library during its seven centuries of existence focused on three areas: geometry, as typified by Euclid's *Elements*, masterwork of the most famous chief librarian; trigonometry, used mainly in astronomy; and mechanics. This makes the work of Diophantus so surprising, because he was interested in solving algebraic equations.

His work was so entirely out of the Greek mainstream that some have speculated that he must have come from the more eastern parts of the realm, perhaps as far east as Iraq. Here are some of the features of his work which set him apart:

- Diophantus developed a system of algebraic notation. While not fully symbolic, it was so advanced that his ideas on algebraic notation would not have much influence on mathematicians until his work was taken up by the Frenchmen Viète and Fermat, over one thousand years later.
- Diophantus studied indeterminate algebraic equations of degree up to 6. An "indeterminate" equation is one with more than one variable, and hence typically with infinitely many solutions. The apex of Alexandrian geometry was Apollonius' work on conics.[1] Conics can be described by equations of degree 2. Apollonius derived their defining equations (he called them the *symptoms*) from geometrical hypotheses, and then used the equations to further elucidate the geometry. By contrast, Diophantus studied algebraic equations outside of any geometric context, and all but ignored their geometric implications.
- Diophantus sought "rational" solutions — that is, solutions in the field of fractions. This was totally out of step with Greek understanding of number, which recognized only two types of quantities: positive whole numbers and continuous geometric measurements (lengths, areas, and so forth). Rather than develop a system of rational numbers, the Greeks elaborated a rather cumbersome theory of ratios of like quantities (numbers to numbers, lengths to lengths, areas to areas, . . .).
- Diophantus freely used negative quantities in his calculations, even tho he was ultimately interested in positive solutions. By contrast, other Greek mathematicians never mentioned negative quantities at all. A complete understanding of negative numbers was developed in medieval China and India, but Diophantus had developed a working understanding of these concepts centuries before.

Only a few of Diophantus' works have survived, and these only in fragments. Many scholars believe that most of his works did not survive long after his death because they were not appreciated by his contemporaries. We do know that Hypatia wrote extensive commentaries on at least part of Diophantus' work, but unfortunately her writing has also not survived.

After Egypt fell to the Muslims around 640, the Greek mathematical traditions were continued by Islamic scholars. The eastern Roman Empire (also called Byzantium) shrank to a small territory surrounding its capital, Constantinople. This last outpost of classical Greek culture survived until 1453, when it was conquered by the Ottoman Turks, and renamed Istanbul.

What we have of Diophantus' works comes in part from fragments that survived in Byzantium and in part from fragments of Arabic translations and commentaries, especially by the Iranians Abu'l-Wafa and Qusta ibn Luqa. Some scholars believe that the latter's work is in fact based on Hypatia's commentaries on Diophantus, and not a translation of Diophantus' work itself.

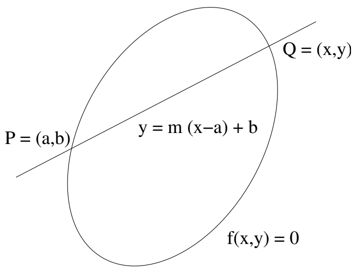
By contrast, medieval Europe had completely abandoned mathematics, and relearned only as Arabic texts were translated into Latin during the reconquest of Spain by the Christians. The capture of Toledo by Alfonso in 1085 brought a wealth of material to Christian Europe, but it would be another five centuries before Europeans would expand on Islamic scholarship.

Ultimately Diophantus exerted a profound influence on mathematicians, especially after they were ready to join algebra and geometry together. Today we refer to the search for rational solutions of algebraic equations as diophantine analysis, and regard this as a very geometric algebra.

Diophantus' work on quadratics

It was known since the time of Pythagoras, seven centuries before Diophantus, that some quadratics have no rational solutions. For example, $x^2 + y^2 = 3$ has no rational solutions. However, Diophantus observed that if a quadratic equation has at least one rational solution, then it has infinitely many. Indeed, if P is a solution to the quadratic equation $f(x, y) = 0$ then every line of rational slope thru P intersects the quadratic curve in a second point Q, whose coordinates must be rational. This gives a way to parametrize the rational solutions, using the slope m as parameter.

[1] Archimedes was the greatest mathematician of antiquity, and one of the greatest of all time. As noted above he did not work at Alexandria, but in the Greek city of Syracuse, in Sicily. Both in mathematical technique and in writing style his work was very un-Alexandrian.



For example, the quadratic

$$3x^2 + 2xy + y^2 - 4x - 1 = 0$$

has the solution $(0, 1)$. To find the point of intersection of this quadratic and the line $y = mx + 1$ we eliminate y and seek to solve the equation

$$3x^2 + 2x(mx + 1) + (mx + 1)^2 - 4x - 1 = [(3 + 2m + m^2)x + (2m - 2)]x = 0$$

It is no accident that x is a factor of the left-hand side of this equation, since we started with the solution $(x, y) = (0, 1)$. The roots of the second factor

$$(3 + 2m + m^2)x + (2m - 2)$$

give us the other solution. Thus, starting with one solution $(0, 1)$ we find all of the others:

$$x = \frac{2 - 2m}{3 + 2m + m^2},\ y = mx + 1 = \frac{3 + 4m - m^2}{3 + 2m + m^2}.$$

Exercise: Apply the diophantine method to find all "pythagorean triples" — positive integers (a, b, c) such that $a^2 + b^2 = c^2$. This is the same as finding rational solutions on the quadratic curve $x^2 + y^2 = 1$, where $(x, y) = (a/c, b/c)$. Use the solution $(0, 1)$ to find all of the other rational solutions.

Answer: $x = -2m/(m^2 + 1)$, $y = (1 - m^2)/(m^2 + 1)$. Note that $-1 < m < 0$ when $x, y > 0$! To avoid negative signs in the answer, Diophantus started with the negative solution $(0, -1)$.

Diophantus' work on cubics

Diophantus applied this same "secant line method" to cubic equations, and obtained his most spectacular results. He discovered that altho the method does not give a rational parametrization of the solutions of most cubics, it does allow you to start with two solutions and produce a third. Centuries later it was realized that a slight modification to this diophantine secant procedure endows the cubic with what we now call the structure of an abelian group.

Diophantus made no such claims, of course. He was interested in algorithmic rather than structural algebra. That is, he was interested in formulas and procedures for solving equations, not abstract structures for describing the set of solutions. However on one other point he again pushed his method to the edge of modern mathematics. Centuries before the development of calculus he used the "limiting case" of the secant line — that is, the tangent line. Of course you do not need the notion of a derivative to find the tangent line of an algebraic curve, nor even the picture of what the tangent line looks like. This is because when you eliminate a variable between the equations of an algebraic curve C and of a line L then L is tangent to C precisely when you obtain an equation for x with a repeated root.

One final remark about the diophantine tangent-and-secant method. He observed that the rational solutions of certain plane quartics can found by making a change of variables which transform the quartic to a cubic. For example, the substitution

$$x = \frac{u + 1}{u}, y = \frac{v}{u^2}$$

transforms

$$y^2 = (x - 1)(x^3 + 2)$$

into

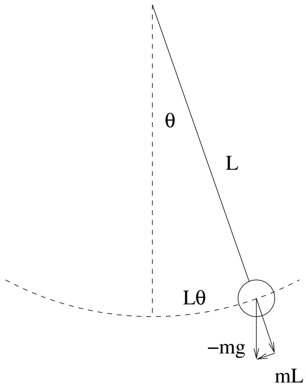
$$v^2 = 3u^3 + 3u^2 + 3u + 1.$$

(See chapter 3 of [1] for more details.) This algebraic trick has no simple geometric meaning, which serves to exhibit the power of Diophantus' algebra.

The pendulum

Quite unexpectedly, the group law on a cubic curve turns up in the solution of many differential equations that arise in engineering applications, including electrostatics, fluid flow, and materials science. Here we discuss only one application: the pendulum.

If θ is the angle a frictionless pendulum of mass m and length L makes with the vertical axis then $L\theta$ is the length of the arc swept out by the pendulum. Using Newton's Law $F = ma = mL\theta''$ and some elementary trigonometry we find that $mL\theta'' = -mg\sin\theta$, where g denotes the acceleration of gravity:



If we multiply both sides of this equation by θ' and divide by mg then we obtain an integrable equation:

$$\frac{L}{g}\,\theta'\theta'' = -\theta'\,\sin\theta$$

$$\frac{L}{2g}(\theta')^2 = \int \frac{L}{g}\,\theta'\theta'' = -\int \theta'\,\sin\theta = C + \cos\theta$$

$$\sqrt{\frac{L}{2g}}\,\theta' = \sqrt{\frac{L}{2g}}\,\frac{d\theta}{dt} = \sqrt{C+\cos\theta}$$

$$\frac{d\theta}{\sqrt{C+\cos\theta}} = \sqrt{\frac{2g}{L}}\,dt$$

$$\int \frac{d\theta}{\sqrt{C+\cos\theta}} = \sqrt{\frac{2g}{L}}\,t$$

Thus, to determine θ as a function of t we need to evaluate the integral on the left and then invert the resulting function of θ. (That is, we find the inverse function, not the reciprocal!)

When we change variables we encounter an old friend: if $x = \cos\theta$ then $dx = -\sin\theta\,d\theta = -\sqrt{1-x^2}\,d\theta$, whence

$$\int \frac{d\theta}{\sqrt{C+\cos\theta}} = -\int \frac{dx}{\sqrt{(1-x^2)(C+x)}} = -\int \frac{dx}{y},$$

where $y^2 = (1-x^2)(C+x)$. Since the natural domain of this integral is a plane cubic, we can derive an addition formula for these integrals:

$$\int_\infty^{x_1} \frac{dx}{y} + \int_\infty^{x_2} \frac{dx}{y} = \int_\infty^{x_3} \frac{dx}{y},$$

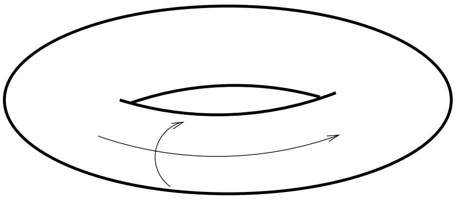
where x_1, x_2, and x_3 are the x-coordinates of the intersection of a line with the cubic. In particular, x_3 is a rational function of x_1 and x_2. (We write down this rational function explicitly in a separate note, where we study the group law in some detail.)

The addition formula for these *elliptic integrals* is due to the Swiss Euler. The inverses of these integrals are *elliptic functions*, and were studied extensively by the Norwegian Abel, and the Germans Gauss and Jacobi.

Euler probably was not aware of the connection with Diophantus' work. Rather, he proceeded by analogy: the inverse of

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin(x)$$

is the familiar sine function, which has a well-known addition formula, and is periodic. An elliptic function $\wp$ is *doubly periodic*: $\wp(x+\tau) = \wp(x)$ for 2 independent periods τ, one of them complex. This reflects a *topological* difference: the set of complex solutions of a quadratic is a (possibly deformed) sphere, whereas the set of complex solutions of an elliptic curve is a torus (the surface of a doughnut).



Functions defined on a (complex) elliptic curve
must be periodic in two different directions

Ellipses?

Why on earth are they called "elliptic" this and "elliptic" that? Ellipses are quadratic curves, whereas elliptic curves are either cubic or quartic. The reason for the name is that these integrals first appeared in the work of the Englishman Wallis, who was trying to determine the arclength of an ellipse. If the ellipse is parametrized by the equations $x = a\cos\theta$, $y = b\sin\theta$, where $a > b$, then the arclength is computed by the integral

$$\int ds = \int \sqrt{dx^2 + dy^2} = \int \sqrt{a^2 \sin^2\theta + b^2 \cos^2\theta}\, d\theta = a \int \sqrt{1 - e^2 \cos^2\theta}\, d\theta,$$

where $1 - b^2/a^2 = e^2$. When we substitute $x = \cos\theta$ we transform the integral to $-a\int y\,dx$, where $y^2(1-x^2) = 1 - e^2x^2$. This quartic can be transformed into a cubic, using a diophantine substitution. Hence it is an elliptic curve, with a group law, which provides this (truly!) elliptic integral with an addition formula.

Exercise: Find a diophantine substitution which transforms the quartic $y^2(1-x^2) = 1 - e^2x^2$ to a cubic.

Answer: Rewrite the equation as $y^2(1-x)^2/(1+x)^2 = (1-x)(1-e^2x^2)/(1+x)^3$, then set $u = 1/(1+x)$ and $v = y(1-x)/(1+x)$. This technique is described at the end of chapter 3 in [1].

Fermat's Last Theorem

There's an interesting appendix to this story. Once, when Fermat was reading his copy of the recent Latin translation of Diophantus' *Arithmetica*, he had a flash of insight. He scribbled in the margin of the text the statement that when $n > 2$ there is no positive rational solution to the equation $x^n + y^n = z^n$. He added that he had a marvelous proof which, sadly, was too long to fit into the margin of the book. Many years later, after Fermat died, his nephew discovered the marginal remark and included it in the posthumous collection of his uncle's mathematical works. It became known as Fermat's Last Theorem, even tho there was no proof.

It is likely that Fermat's "proof" fell apart when he thought about it in more detail the next day, because there is no other mention of this "theorem" anywhere else in his work. Fermat did not always publish proofs of his theorems, but he invariably circulated at least their statements, as a boast and a challenge. Fermat was a prolific letter-writer, but his "last theorem" appears nowhere in his correspondence.

At any rate, the problem stumped the best minds for over 300 years, until it was solved by the Englishman Wiles, in 1994, fulfilling his boyhood dream. In fact, Wiles attacked the Fermat "enigma" by solving a completely different problem, concerning elliptic curves, which was first studied by the Japanese Taniyama and Shimura. A few years before Wiles began his work, the German Frey had shown how to connect Fermat's Last Theorem to the Taniyama-Shimura Conjecture. Once again diophantine addition imposes itself.

References

[1] Isabella Grigoryevna Bashmakova, *Diophantus and Diophantine Equations*, MAA, 1997.

[2] Simon Singh, *Fermat's Enigma*, Knopf, 1998.

[3] *The MacTutor History of Mathematics:* `http://www-groups.dcs.st-and.ac.uk/~history/`.