106 : Fall 2004

Fermat

1. Method of infinite descent. One of the most enduring of Fermat's contributions to number theory is the so called *method of infinite descent.* It is a way of proving that certain diophantine equations do not have integral solutions. The idea is to show that if one starts with a solution, then one can find another solution which is *lower* than the first solution in a definite sense. This gives the *infinite descent* which will lead to a contradiction. We shall illustrate this method by proving that the equation

$$x^4 + y^4 = z^2, \qquad x \neq 0, y \neq 0, z \neq 0$$

has no solutions in integers. Notice that this will also prove that the Fermat equation

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

also has no solutions in integers all nonzero. Since for any solution (x, y, z), $(\pm x, \pm y, \pm z)$ are also solutions, it is enough to prove that

$$x^4 + y^4 = z^2, \qquad x > 0, y > 0, z > 0$$

has no solutions. Hereafter, by a solution we mean a solution in integers all positive. We shall show that if there is a solution (x, y, z) there is a solution (x_1, y_1, z_1) with $x_1 > 0, y_1 > 0, z_1 > 0$ and $0 < z_1 < z$. Since there are only finitely many integers that are > 0 and $< z_1$ we get a contradiction. We shall therefore start with a solution (x, y, z) and show that there is another solution (x_1, y_1, z_1) with $0 < z_1 < z$.

We consider 2 cases:

Case I : $(x, y) > 1$. Then there is a prime p that divides both x and y. So p^4 divides x^4, y^4 and hence $z^2 = x^4 + y^4$. This means p^2 divides z (supply argument). We can therefore write

$$x = px_1, \ y = py_1, \ z = p^2 z_1$$

to get

$$p^4(x_1^4 + y_1^4) = p^4 z_1^2$$

or

$$x_1^4 + y_1^4 = z_1^2.$$

Since

$$z_1 = \frac{z}{p^2} < z$$

we have completed the descent step.

Case II : $(x, y) = 1$. The equation says that (x^2, y^2, z) is a Pythagorean triple, and the condition that $(x, y) = 1$ which implies that $(x^2, y^2) = 1$, means that the PT is primitive. We now appeal to the theorem that describes all primitive PT's. By interchanging x and y we may assume that y^2 is even. Then we can write

$$x^2 = m^2 - n^2, \ y^2 = 2mn, \ z = m^2 + n^2$$

where

$$(m, n) = 1, m > n > 0, m, n \text{ of opposite parity.}$$

This gives

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

We assert that n is even and hence m is odd. Suppose that n is odd so that m is even. Then x^2 is odd, so x is odd, and so $x^2 \equiv 1 \bmod 4$. As n^2 is also $\equiv 1 \bmod 4$, we see that $m^2 \equiv 2 \bmod 4$ which is a contradiction, because, as m is even, m^2 is divisible by 4.

This means that (x, n, m) is a primitive PT with n even and so we can write

$$x = a^2 - b^2, \ n = 2ab, \ m = a^2 + b^2$$

where

$$(a, b) = 1, \ a > b > 0, \ a, b \text{ of opposite parity.}$$

Since m and n are mutually prime and m is odd, m and $2n$ are mutually prime. But $y^2 = m(2n)$ and so each of m and $2n$ is a perfect square (why?). Hence $m = m_1^2, \ 2n = n_1^2$. So $n_1 = 2n_2$ giving us $n = 2n_2^2 = 2ab$ or $n_2^2 = ab$. Again, as a and b are mutually prime, they both must be perfect squares. Hence $a = x_1^2, b = y_1^2$. Then

$$m = m_1^2 = a^2 + b^2 = x_1^4 + y_1^4$$

showing that $(x_1, y_1, z_1 = m_1)$ is a solution in positive integers for the original equation. If we prove that $m_1 < z$ we are done. But

$$z = m^2 + n^2 > m^2 = m_1^4 \geq m_1 = z_1.$$

Fermat applied his method of infinite descent to many problems. He was aware that one cannot apply it to all diophantine equations and knew roughly the class of problems to which it can be applied. A real understanding of his method is possible only with the help of the theory of elliptic curves.

2. Fermat principle for the propagation of light. Consistent with the picture we have of Fermat as an innovator and creator of ideas is the fact that he discovered what is called the *Fermat principle of least time.* This asserts that *light always travels by the path which takes the least amount of time.* This is an absolutely remarkable principle, for *all* the laws of geometrical optics can be derived from it. I give a discussion of a few examples; for more details one should refer to the *Feynman Lectures in Physics, Vol I, §26.* In following the discussion below it is better to refer to the accompanying set of figures at the appropriate places.

The law of reflection. When light falls on a mirror, the incident angle and the reflected angle are the same:

$$\theta_i = \theta_r.$$

All angles are measured as angles made with the normal to the surface at the point of incidence.

How does one derive this from Fermat's principle of least time? If A, B are two points and we seek to find out how light travels from A to B, it is clear that it has to be along the straight line AB. For, as the speed of light is constant, transit time is proportional to distance, and so we seek the *shortest path* between A and B. This is clearly the line joining A and B. Suppose now that we have a mirror which is represented by a line L, A, B are both on the same side of L, and we want the path from A to B with *the condition that the path has to hit L first.* What is the shortest path? Let Z be an arbitrary point on L and consider the path AZB (see figure 1). Let B' be the reflection of B in L, join B' to A and let AB' meet L at P. We claim that APB is the shortest path. Let BB' meet L at R. The triangles $ZBR, ZB'R$ are congruent because they have a common side ZR, equal sides $BR = B'R$, and the included angle is a right angle. So $ZB = ZB'$. Hence, $|\cdot|$ denoting length, $|AZB| = |AZB'|$. But the length $|AZB'|$ is a minimum only when A, Z, B' are all on a straight line, i.e., when $Z = P$. Thus APB is the shortest path.

Let us derive the law of reflection from this. Suppose the light ray goes via the path AZB from A to B. Then we saw that $Z = P$. Arguing as before we know that the triangles $PBR, PB'R$ are congruent. Hence the angles $BPR, B'PR$ are equal. But APB' is a straight line. Hence the angles $B'PR$ and CPA are equal where C is a point on the line L to the left of P (see figure 1). Hence the angles CPA and

RPB are equal. But these are complementary to the angles θ_i, θ_r respectively (see figure). Hence $\theta_i = \theta_r$.

Snell's law of refraction. When going from one medium (say air) into another (say water), light rays bend at the surface of separation of the two media. The precise amount of bending is determined by *Snell's law of refraction.* If the velocity in the second medium is n^{-1} times the velocity in the first medium, where $n > 1$ is a constant, then the angles of incidence and refraction are related by

$$\sin\theta_i = n\cdot\sin\theta_r.$$

We shall now deduce this from the principle of least time.

For simplicity let us assume that the boundary separating the two media is a line which we take as the x-axis. Let A and B be on opposite sides of the x-axis which we denote by L. As before, let Z be a point on L with coordinates $(u, 0)$. We want to find Z so that the path AZB takes the least amount of time. Now transit time is *not* proportional to distance because the speeds in the two media are different. But if the velocities are v above L and $n^{-1}v$ below L, the transit time for the path AZB is

$$\frac{|AZ|}{v} + n\cdot\frac{|ZB|}{v}$$

and hence is proportional to

$$|AZ| + n\cdot|ZB| = t(Z)$$

which is the expression we want to minimize. Now, as $Z = (u, 0)$, $t(Z)$ is a function of u, say $T(u)$. Let

$$A = (a_1, a_2), \qquad B = (b_1, b_2).$$

Then

$$T(u) = \sqrt{(u-a_1)^2 + a_2^2} + n\sqrt{(u-b_1)^2 + b_2^2}.$$

We must solve the equation

$$\frac{dT}{du} = 0$$

for u; in any case u must obey this equation. So

$$\frac{dT}{du} = \frac{u-a_1}{\sqrt{(u-a_1)^2 + a_2^2}} + n\frac{u-b_2}{\sqrt{(u-b_1)^2 + b_2^2}} = 0.$$

If (see figure 2) M is the line through Z perpendicular to L and R, S are the feet of the perpendiculars on M from A, B respectively, the above equation can be written

$$\frac{|AR|}{|AZ|} = n\frac{|BS|}{|BZ|}$$

which is just

$$\sin\theta_i = n\sin\theta_r.$$

Focusing. Suppose we want to insert in the path of light an optical system so that light from A always passes through B, i.e., we want to *focus light from A at B.* How should we do it? Here again the principle of least time furnishes the answer. Suppose we have a family of light rays starting from A, passing through the optical system, and then emerging to pass through B, the paths depending on a parameter α (figure 3). Let $t(\alpha)$ be the transit time for the path corresponding to α. Since each path must have shortest transit time, we must have

$$\frac{dt}{d\alpha} = 0$$

for all α. But then this means that $t(\alpha)$ *must be a constant.* In other words, *all the paths from A to B through the optical system must have the same transit time.* This is the *principle of focusing.*

As a first illustration let us find the shape of a mirror M so that light from A gets reflected and always passes through B (see figure 4). Since there is no change of media, transit times are proportional to lengths of paths and so, by the focusing principle, we need the mirror to be such that if Z is the point where the light ray strikes M,

$$|AZ| + |ZB| = \text{ a constant.}$$

In other words, M should be an elliptic mirror with its foci at A and B!

As a second illustration, let us ask for the shape of the mirror so that starlight gets reflected and focuses at a point B, where the astronomer sits presumably squinting at the stars (see figure 5). This is the same problem as before except that the source of light is infinitely distant (in comparison with the dimensions of the mirror and other things in the problem). If we draw two parallel lines L, L' above and below the mirror M, perpendicular to all the light rays from infinity, then a typical light ray takes the path AZB where B is the focusing point. If we choose the curve representing the mirror as the locus of points Z such that $|A'Z| = |ZB|$, then

$$|AZ| + |ZB| = |AZ| + |ZA'| = \text{ a constant.}$$

Thus the curve representing the mirror is the locus of the point which moves such that its distance from B is equal to its distance from the line L'. This is just a parabola with focus at B. These discussions make clear why these points are called foci in geometry.

The Fermat principle of least time was the first principle where the actual path of a moving system was singled out as the solution to a variational problem. This theme would be taken up by Lagrange, Hamilton, and then, most spectacularly, by Feynman himself in his great researches in quantum theory.