Lecture 8: Factorizing special numbers

We have seen that factorizing numbers is a time-consuming business. However, in certain circumstances it is possible to obtain better results. In this lecture, I shall describe two such circumstances.

Fermat factorization

This method arose from correspondence between Fermat and Mersenne in the seventeenth century. It is based on the well-known identity

$$a^2 - b^2 = (a+b)(a-b).$$

Thus if a number can be represented as the difference of two perfect squares it can be factorized. For example,

$$20 = 6^2 - 4^2$$

and

$$20 = (6+4)(6-4) = 10 \times 2.$$

This might look unlikely but in fact is always possible as we shall now see.

Theorem 0.1. *Let n be an odd natural number. Then there is a bijective correspondence between factorizations $n = ab$ where $a \geq b > 0$ and representations of n of the form $u^2 - v^2$. The bijection takes the following form*

$$(a,b) \mapsto u^2 - v^2 \text{ where } u = \frac{a+b}{2} \text{ and } v = \frac{a-b}{2}.$$

Proof. Suppose first that

$$n = u^2 - v^2.$$

Then $n = (u+v)(u-v)$. Put $a = u+v$ and $b = u-v$ then $u = \frac{a+b}{2}$ and $v = \frac{a-b}{2}$.

Suppose that $n = ab$. Since n is odd it follows that both a and b are odd and so $u = \frac{a+b}{2}$ and $v = \frac{a-b}{2}$ are both integers. Observe that

$$u^2 - v^2 = \frac{a^2 + 2ab + b^2}{4} - \left(\frac{a^2 - 2ab + b^2}{4}\right) = ab.$$

□

Let's now see how the above result can be turned into a method for factorizing numbers. Our goal is to try and write $n = u^2 - v^2$. Observe that $u^2 = n + v^2$. It follows that $u \geq \sqrt{n}$. Put $t = \lfloor \sqrt{n} \rfloor + 1$. We calculate in turn the numbers $t^2 - n$, $(t+1)^2 - n$, $(t+2)^2 - n$, ... and at each attempt we check to see if the number is a perfect square. This procedure will be fast when a and b are close together.

For example, let's use this method to factorize $200,819$. Here $t = 449$. Here are the calculations.

t	$t^2 - n$
449	782
450	$1681 = 41^2$

Thus $450^2 - n = 41^2$. It follows that

$$n = 450^2 - 41^2 = (450 + 41)(450 - 41) = 491 \cdot 409.$$

Generalized Fermat factorization

The above method is only practical when there are two factors close together. If you apply it and you have no success after a number of steps you can try the following varation. Choose a small number k and calculate $t = \lfloor \sqrt{kn} \rfloor + 1$. We then compute $t^2 - kn$ incrementing t by one at each step until we get a perfect square s^2. For that value of t we have that $t^2 - kn = s^2$ and so $(t + s)(t - s) = kn$. If $\gcd(t + s, n)$ is non-trivial it will be a non-trivial factor of n.

For example, we use this method on the number 141467 and choose $k = 3$. We find that $t = 652$. We now calculate.

t	$t^2 - 3 \cdot n$
652	703
653	2008
654	3315
655	$4624 = 68^2$

Thus $655^2 - 3 \cdot 141467 = 68^2$. Hence $655^2 - 68^2 = 3 \cdot 141467$ and so $(655+68)(655-68) = 3 \cdot 141467$. We calculate $\gcd(723, 141467) = 241$ and $141467 = 241 \cdot 587$.

The choice of k is down to trial and error and the benificence of the exam setter.

Sums of squares

This method is due to Euler. We assume that we are given an odd number n that can be written in two different ways as a sum of squares

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

where a and c are both odd and b and d are both even. Thus

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

and so

$$a^2 - c^2 = d^2 - b^2.$$

We therefore may factorize

$$(a-c)(a+c) = (d-b)(d+b).$$

Let

$$k = \gcd(a-c, d-b).$$

Then

$$a-c = kl, \quad d-b = km, \quad \gcd(l,m) = 1.$$

Observe that $a-c$ and $d-b$ are both even and so k is even. Substituting back and cancelling the k we get that

$$l(a+c) = m(d+b).$$

But l and m are coprime. Thus by Gauss's Lemma, $a+c$ must be divisible by m. Put

$$a+c = mn$$

for some n. We also get that

$$b+d = ln.$$

It follows that $n = \gcd(a+c, d+b)$ and so is also even.

We may now factorize n as follows

$$n = \left(\left(\frac{k}{2}\right)^2 + \left(\frac{n}{2}\right)^2\right)\left(l^2+m^2\right).$$

To prove this, multiply out the righthand side.

Observe that

$$k = \gcd(a-c, d-b) \text{ and } n = \gcd(a+c, d+b).$$

For example, we may write

$$221 = 10^2 + 11^2 = 5^2 + 14^2.$$

Following the above method we get that

$$221 = (1+4^2)(2^2+3^2) = 17 \cdot 13.$$

Patterns within the primes: primes that can be written as sums of squares

There are interesting patterns within the prime numbers. Apart from 2, all prime numbers are odd. This means they leave the remainder 1 when divided by 2.

We now look at the possible remainders when an odd prime is divided by 4. Any number can be written as $4q$, $4q+1$, $4q+2$ or $4q+3$ for some q. This means that each odd prime is either of the form $4q+1$ or $4q+3$ for some q. The first few odd primes are

$$3, \mathbf{5}, 7, 11, \mathbf{13}, \mathbf{17}, 19, 23, \mathbf{29}, 31, \mathbf{37}, \mathbf{41}, 43, 47, \mathbf{53}, 59, \mathbf{61}, \ldots$$

where the ones in bold leave the remainder 1 when divided by 4.

One obvious question is: are there infinitely many of each type of prime?

Focus now on the primes in bold above, and observe the following

$$5 = 2^2 + 1^2, 13 = 2^2 + 3^2, 17 = 4^2 + 1^2, 29 = 5^2 + 2^2$$
$$37 = 6^2 + 1, 41 = 4^2 + 5^2, 53 = 7^2 + 2^2, 61 = 5^2 + 6^2.$$

That this is no accident is the substance of the following theorem.

Theorem 0.2. *An odd prime can be written uniqueley as the sum of two squares if and only if when divided by 4 it leaves the remainder 1.*

Proof. We prove the easy direction first. Let $p = a^2 + b^2$ where p is an odd prime. Since the sum of two even numbers or of two odd numbers is even, it follows that one square is odd and the other even. Without loss of generality, we may suppose that a^2 is even and b^2 is odd. From the fundamental property of primes, since $2 \mid a^2$ we must have that $2 \mid a$. Thus a is even. We may therefore write $a = 2r$ for some r. One the other hand b^2 odd implies that b is odd. Thus we may write $b = 2s + 1$. It follows that

$$a^2 + b^2 = 4r^2 + 4s^2 + 4s + 1 = 4q + 1$$

for some q. We have shown that p leaves the remainder 1 when divided by 4.

To prove the hard direction, we shall use the paper attached.

Uniqueness follows from the section above: if the decomposition were not unique the prime would be composite which is nonsense. □

THE TEACHING OF MATHEMATICS

EDITED BY MELVIN HENRIKSEN AND STAN WAGON

A One-Sentence Proof That Every Prime $p \equiv 1 \pmod 4$ Is a Sum of Two Squares

D. ZAGIER
Department of Mathematics, University of Maryland, College Park, MD 20742

The involution on the finite set $S = \{(x, y, z) \in \mathbb{N}^3 : x^2 + 4yz = p\}$ defined by

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

has exactly one fixed point, so $|S|$ is odd and the involution defined by $(x, y, z) \mapsto (x, z, y)$ also has a fixed point. □

This proof is a simplification of one due to Heath-Brown [1] (inspired, in turn, by a proof given by Liouville). The verifications of the implicitly made assertions—that S is finite and that the map is well-defined and involutory (i.e., equal to its own inverse) and has exactly one fixed point—are immediate and have been left to the reader. Only the last requires that p be a prime of the form $4k + 1$, the fixed point then being $(1,1,k)$.

Note that the proof is not constructive: it does not give a method to actually find the representation of p as a sum of two squares. A similar phenomenon occurs with results in topology and analysis that are proved using fixed-point theorems. Indeed, the basic principle we used: "The cardinalities of a finite set and of its fixed-point set under any involution have the same parity," is a combinatorial analogue and special case of the corresponding topological result: "The Euler characteristics of a topological space and of its fixed-point set under any continuous involution have the same parity."

For a discussion of constructive proofs of the two-squares theorem, see the Editor's Corner elsewhere in this issue.

REFERENCE

1. D. R. Heath-Brown, Fermat's two-squares theorem, *Invariant* (1984) 3–5.

Taken from The American Mathematical Monthly
97 (1990), 144.