Zagier's One-Sentence Proof

A. Kolesnikov

When you think of a one-line proof, you probably picture a counterexample or a claim so trivial that one sentence is enough. That makes it all the more surprising that Don Zagier proved the harder half of Fermat's two-square theorem in a single sentence [1].

Fermat's two-square theorem. An odd prime is a sum of two integer squares exactly when it leaves remainder 1 upon division by 4.

An even square leaves remainder 0 upon division by 4, and an odd square leaves remainder 1. An odd sum of two squares must use one of each, so it leaves remainder 1. For the converse, fix a prime $p = 4k + 1$. We want to show that $p = x^2 + (2y)^2$ for some positive integers x and y.

Zagier starts with $x^2 + 4y^2 = p$ and replaces one copy of y with a new variable z. Let S_p be the set of triples of positive integers (x, y, z) that satisfy

$$x^2 + 4yz = p.$$

This set is finite. We have $x < \sqrt{p}$, and $yz \leq (p-1)/4 = k$ because $x^2 \geq 1$. Since both y and z are positive, each is at most k. A single triple in S_p with $y = z$ gives the two squares we want.

The advantage of introducing z is that swapping y and z keeps their product the same. Call this swap g. Applying g twice returns the original triple, so g is an *involution*, a rule that undoes itself. Whenever g moves a triple to a different one, it also moves that second triple back to the first. The only triples left unpaired are the ones g leaves unchanged. These are its *fixed points*, and they are exactly the triples with $y = z$. If S_p has an odd number of triples, at least one must be left unpaired under g.

Zagier proves that the total is odd by giving the same triples different partners. His second involution, which we will call f, leaves exactly one triple unpaired. In the original proof below, Zagier writes S for our set S_p.

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. □

Don Zagier's original proof, *The American Mathematical Monthly*, 1990 [1].

To justify the sentence, we need to check that f sends every triple in S_p to another triple in S_p, that applying it twice returns the starting triple, and that it has exactly one fixed point. We take these in that order.

First, every triple must fall into one of the three rows. Since $y - z < 2y$, the rows cover all values of x except the two boundaries $x = y - z$ and $x = 2y$. At the first boundary,

$$p = x^2 + 4yz = (y - z)^2 + 4yz = y^2 + 2yz + z^2 = (y + z)^2.$$

This would make a prime a square. At the second boundary,

$$p = x^2 + 4yz = 4y^2 + 4yz = 4y(y + z),$$

which would make an odd prime divisible by 4. Thus exactly one row applies to each triple.

Now check the coordinates that row produces. In the first row, $x + 2z$ and z are positive, while $x < y - z$ gives $y - x - z > 0$. In the middle row the two bounds $y - z < x < 2y$ give $x - y + z > 0$ and $2y - x > 0$. In the third row $x > 2y$ gives $x - 2y > 0$, and also $x - y + z > y + z > 0$. Every output is therefore a triple of positive integers.

It must also satisfy the same equation. For the first row, substitution and expansion give

$$\begin{aligned}(x + 2z)^2 + 4z(y - x - z) &= x^2 + 4xz + 4z^2 + 4yz - 4xz - 4z^2 \\ &= x^2 + 4yz = p.\end{aligned}$$

For the middle row the corresponding calculation is

$$\begin{aligned}(2y - x)^2 + 4y(x - y + z) &= 4y^2 - 4xy + x^2 + 4xy - 4y^2 + 4yz \\ &= x^2 + 4yz = p.\end{aligned}$$

The third row gives this same expression because $x - 2y$ and $2y - x$ have the same square, while swapping the last two coordinates leaves their product unchanged. Hence f sends every triple in S_p to another triple in S_p.

We can now check that f undoes itself. Write $(x', y', z') = f(x, y, z)$. If the original triple uses the first row, then $x' = x + 2z$, $y' = z$ and $z' = y - x - z$. Since

$$x' - 2y' = (x + 2z) - 2z = x > 0,$$

the new triple uses the third row. Applying that row returns the original coordinates, one by one:

$$\begin{aligned}x' - 2y' &= (x + 2z) - 2z = x, \\ x' - y' + z' &= (x + 2z) - z + (y - x - z) = y, \\ y' &= z.\end{aligned}$$

Conversely, starting in the third row gives $(x', y', z') = (x - 2y, x - y + z, y)$. Here

$$y' - z' - x' = (x - y + z) - y - (x - 2y) = z > 0,$$

so the new triple uses the first row. It returns

$$\begin{aligned}x' + 2z' &= (x - 2y) + 2y = x, \\ z' &= y, \\ y' - x' - z' &= (x - y + z) - (x - 2y) - y = z.\end{aligned}$$

The remaining possibility is the middle row, which gives

$$(x', y', z') = (2y - x,\ y,\ x - y + z).$$

Before using the rule a second time, we must check that this new triple is still in the middle case. Its upper bound holds because

$$2y' - x' = 2y - (2y - x) = x > 0.$$

Its lower bound holds because

$$x' - (y' - z') = (2y - x) - y + (x - y + z) = z > 0.$$

Thus $y' - z' < x' < 2y'$, as required. Applying the middle row again now gives

$$\begin{aligned} 2y' - x' &= 2y - (2y - x) = x, \\ y' &= y, \\ x' - y' + z' &= (2y - x) - y + (x - y + z) = z. \end{aligned}$$

All three cases therefore return the original triple after two applications. This proves that f is an involution. Every triple that moves under f belongs to a pair with the triple it moves to.

It remains to find the triples that stay put. The first and third rows move a triple to a different case, so a fixed point must use the middle row. There the first coordinate is unchanged exactly when

$$2y - x = x, \qquad \text{or equivalently} \qquad x = y.$$

When $x = y$, the third coordinate is unchanged too, because $x - y + z = z$. Thus being fixed by f is equivalent to having $x = y$.

Substituting $x = y$ into the equation for S_p gives

$$p = y^2 + 4yz = y(y + 4z).$$

Both factors are positive integers, and $y + 4z > y$. A prime has only the positive factors 1 and itself, so $y = 1$ and $y + 4z = p$. It follows that

$$x = y = 1, \qquad z = \frac{p - 1}{4} = k.$$

This identifies the only possible fixed point. It really is in S_p, since $1^2 + 4 \cdot 1 \cdot k = 4k + 1 = p$, and it lies in the middle case because $1 - k < 1 < 2$. Substitution into that row gives

$$f(1, 1, k) = (2 - 1,\ 1,\ 1 - 1 + k) = (1, 1, k).$$

There is therefore exactly one fixed point.

We can now finish the counting argument. Apart from $(1, 1, k)$, every triple in S_p has a distinct partner under f. If there are r such pairs, then

$$|S_p| = 2r + 1,$$

where $|S_p|$ means the number of triples in the set. So S_p has an odd number of triples.

Return to the original swap $g(x, y, z) = (x, z, y)$. Its pairs each use two triples, so an odd set must contain at least one triple that the swap fixes. For that triple $y = z$, and the defining equation becomes

$$\boxed{p = x^2 + 4y^2 = x^2 + (2y)^2.}$$

These are the two integer squares whose existence we wanted to prove.

References

[1] D. Zagier, A One-Sentence Proof That Every Prime $p \equiv 1 \pmod 4$ Is a Sum of Two Squares, *The American Mathematical Monthly* **97** (1990), 144.