PYTHAGOREAN PRIMES AND PALINDROMIC CONTINUED FRACTIONS

Arthur T. Benjamin
Department of Mathematics, Harvey Mudd College, Claremont, CA 91711
benjamin@math.hmc.edu

Doron Zeilberger
Department of Mathematics, Rutgers University, Piscataway, NJ 08854-8019
zeilberg@math.rutgers.edu

Received: 4/12/05, Revised: 11/23/05, Accepted: 11/29/05, Published: 12/9/05

Abstract

In this note, we prove that every prime of the form $4m + 1$ is the sum of the squares of two positive integers in a unique way. Our proof is based on elementary combinatorial properties of continued fractions. It uses an idea by Henry J. S. Smith ([3], [5], and [6]) most recently described in [4] (which provides a new proof of uniqueness and reprints Smith's paper in the original Latin). Smith's proof makes heavy use of nontrivial properties of determinants. Our purely combinatorial proof is self-contained and elementary.

For $n \geq 1$ and positive integers $a_0, \ldots, a_n$, let $[a_0, \ldots, a_n]$ denote the finite continued fraction

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \frac{1}{a_n}}}}, \tag{1}$$

which simplifies to a unique rational number $r/s > 1$ in lowest terms. Conversely, for every rational number $r/s > 1$, there is a unique continued fraction $[a_0, \ldots, a_n] = r/s$ where $a_0 \geq 1, \ldots a_{n-1} \geq 1$, and $a_n \geq 2$. (It happens that r/s has one other continued fraction representation, namely $[a_0, \ldots, a_{n-1}, a_n - 1, 1]$, but we will not use this.)

Continued fractions have a simple combinatorial interpretation, which we describe here. For positive numbers $a_0, \ldots, a_n$, define the *continuant* $K(a_0, \ldots, a_n)$ to be the number of ways to tile a strip of length n with dominoes (of length two) and *stackable* squares (of length one). For $1 \leq i \leq n$, if cell i is covered by a square, then the number of squares that may be stacked on the ith cell is at most a_i; if cell i is covered by a domino, then nothing is stacked on top of that domino.

For example, $K(3,7,15) = 333$ since a strip of length three can be tiled as "domino-square" in 15 ways (by choosing how many squares to stack on the third cell) , as "square-domino" in 3 ways (by choosing how many squares to stack on the first cell), or as "square-square-square" in $3 \times 7 \times 15 = 315$ ways (by choosing how many squares to stack on each cell). Thus for nonnegative integers a and b,

$$K(a) = a, \qquad K(a,b) = ab + 1. \tag{2}$$

For example, $K(7,15) = 106$. For the empty set, we define $K(\ ) = 1$. For $n \geq 2$,

$$K(a_0, \ldots, a_n) = a_n K(a_0, \ldots, a_{n-1}) + K(a_0, \ldots, a_{n-2}), \tag{3}$$

since the first term counts tilings that end with a stack of squares and the second term counts those that end with a domino. More generally, observe that for $n \geq 1$ and $0 \leq \ell \leq n-1$,

$$K(a_0, \ldots, a_n) = K(a_0, \ldots, a_\ell)K(a_{\ell+1}, \ldots, a_n) + K(a_0, \ldots, a_{\ell-1})K(a_{\ell+2}, \ldots, a_n). \tag{4}$$

The first summand counts tilings that do not have a domino covering cells ℓ and $\ell+1$, while the second summand counts those that do.

Finally, we observe that for any nonnegative $a_0, \ldots, a_n$,

$$K(a_n, \ldots, a_0) = K(a_0, \ldots, a_n) \tag{5}$$

since any length n tiling that satisfies the conditions on the right (at most a_i squares stacked on cell i) can be reversed to satisfy the tiling conditions on the left (at most a_{n-i} squares stacked on cell i), and vice versa.

Using the initial conditions and recurrence in equations (2) and (3), it follows that

$$[a_0, \ldots, a_n] = \frac{K(a_0, \ldots, a_n)}{K(a_1, \ldots, a_n)}, \tag{6}$$

in lowest terms. (See [1], [2] for more details.) That is, if the continued fraction $[a_0, \ldots, a_n] = \frac{p_n}{q_n}$, in lowest terms, then $p_n = K(a_0, \ldots, a_n)$ and $q_n = K(a_1, \ldots, a_n)$. Thus, for example, the continued fraction $[3,7,15] = K(3,7,15)/K(7,15) = 333/106$.

Now suppose that $p = 4m+1$ is prime. We shall consider the continued fraction expansions of the numbers $p/1, p/2, \ldots, p/(2m)$. For each j between 1 and $2m$, we have $p/j > 2$ and is in lowest terms. Thus $p/j = [a_0, \ldots, a_n]$ where $a_0 \geq 2$ and $a_n \geq 2$. By equations (6) and (5),

$$p = K(a_0, \ldots, a_n) = K(a_n, \ldots, a_0).$$

Thus $[a_n, \ldots, a_0]$ also has numerator p, and since $a_n \geq 2$, $[a_n, \ldots, a_0] = p/i$ for some i between 1 and $2m$. Thus each fraction p/j can be paired up with its "reversed" fraction p/i.

Now $p/1 = [p]$ is *palindromic*; it is its own reversal. Thus since $2m$ is even, there must be at least one other fraction p/j^* that is palindromic. That is, for some j^* between 2 and $2m$,

$$[a_0, \dots, a_{n^*}] = p/j^* = [a_{n^*}, \dots, a_0].$$

For example, when $p = 5$, $5/1 = [5]$ and $5/2 = [2, 2]$ are both palindromic. When $p = 13$, $13/1 = [13]$, $13/2 = [6, 2]$, $13/3 = [4, 3]$, $13/4 = [3, 4]$, $13/5 = [2, 1, 1, 2]$, $13/6 = [2, 6]$; so $13/1$ and $13/5$ are palindromic.

We claim that n^* must be even. For suppose, to the contrary, that $n^* = 2\ell + 1$, for some $\ell \geq 0$. Then $p/j^* = [a_0, \dots, a_\ell, a_{\ell+1}, a_\ell, \dots, a_0]$. Thus by applying equations (6), (4), and (5), we have

$$\begin{aligned} p &= K(a_0, \dots, a_\ell, a_{\ell+1}, a_\ell, \dots, a_0) \\ &= K(a_0, \dots, a_\ell)K(a_{\ell+1}, \dots, a_0) + K(a_0, \dots, a_{\ell-1})K(a_\ell, \dots, a_0) \\ &= K(a_0, \dots, a_\ell)[(K(a_0, \dots, a_{\ell+1}) + K(a_0, \dots, a_{\ell-1})]. \end{aligned}$$

But then p is composite (both factors are at least two, since $a_0 \geq 2$), which is a contradiction.

Thus $n^* = 2\ell$ for some $\ell \geq 1$. Consequently, $p/j^* = [a_0, \dots, a_l, a_l, \dots, a_0]$, and

$$\begin{aligned} p &= K(a_0, \dots, a_\ell, a_\ell, \dots, a_0) \\ &= K(a_0, \dots, a_\ell)K(a_\ell, \dots, a_0) + K(a_0, \dots, a_{\ell-1})K(a_{\ell-1}, \dots, a_0) \\ &= (K(a_0, \dots, a_\ell))^2 + (K(a_0, \dots, a_{\ell-1}))^2 \end{aligned}$$

is the sum of two squares, as desired.

For example, when $p = 13$, the palindromic $13/5$ leads to

$$13 = K(2, 1, 1, 2) = K(2, 1)K(1, 2) + K(2)K(2) = 3^2 + 2^2.$$

For a larger example, when $p = 1069$, the palindromic $1069/249 = [4, 3, 2, 2, 3, 4]$ leads to

$$1069 = K(4, 3, 2, 2, 3, 4) = (K(4, 3, 2))^2 + (K(4, 3))^2 = 30^2 + 13^2.$$

Following the strategy in [4], we combinatorially prove uniqueness of the sum of squares using one more elementary fact about continued fractions: If $[a_0, \dots, a_n] = p_n/q_n$ in lowest terms, then for $n \geq 2$,

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^n}{q_n q_{n-1}}.$$

Equivalently, $p_n q_{n-1} - p_{n-1} q_n = (-1)^n$, or by equation (6), for $n \geq 2$,

$$K(a_0, \dots, a_n)K(a_1, \dots, a_{n-1}) - K(a_0, \dots, a_{n-1})K(a_1, \dots, a_n) = (-1)^n. \tag{7}$$

For a direct combinatorial proof of equation (7), see [1] or [2].

Now suppose that $p = r^2 + s^2$ and $p = u^2 + v^2$ for positive integers $r > s$ and $u > v$. Since p is prime, $\gcd(r, s) = 1 = \gcd(u, v)$. Thus r/s and u/v are fractions in lowest terms, and there exist unique positive integers $r_0, \dots, r_t$ and $u_0, \dots, u_w$ such that $r/s = [r_0, \dots, r_t]$ and $u/v = [u_0, \dots, u_w]$, where $r_t \geq 2$ and $u_w \geq 2$.

Hence, by equation (6),

$$\frac{r}{s} = \frac{K(r_0, \dots, r_t)}{K(r_1, \dots, r_t)}$$

in lowest terms. Now by equations (4) and (5),

$$p = r^2 + s^2 = K(r_t, \dots, r_0)K(r_0, \dots, r_t) + K(r_t, \dots, r_1)K(r_1, \dots, r_t) = K(r_t, \dots, r_0, r_0, \dots, r_t).$$

Now let $x = K(r_t, \dots, r_0, r_0, \dots, r_{t-1})$. By equation (3),

$$p = K(r_t, \dots, r_0, r_0, \dots, r_t) = xr_t + K(r_t, \dots, r_0, r_0, \dots, r_{t-2}) \geq 2x + 1.$$

Thus $2 \leq x \leq (p-1)/2$. Also, by equation (7),

$$K(r_t, \dots, r_0, r_0, \dots, r_t)K(r_{t-1}, \dots, r_0, r_0, \dots, r_{t-1}) \\ - K(r_t, \dots, r_0, r_0, \dots, r_{t-1})K(r_{t-1}, \dots, r_0, r_0, \dots, r_t)$$

equals $(-1)^{2t} = 1$. Hence, $pK(r_{t-1}, \dots, r_0, r_0, \dots, r_{t-1}) - x^2 = 1$, and therefore x satisfies $x^2 \equiv -1 \pmod{p}$.

By the same argument, we also have $u/v = K(u_0 \dots, u_w)/K(u_1 \dots, u_w)$ in lowest terms, $p = K[u_w, \dots, u_0, u_0, \dots, u_w]$, and $y = K(u_w, \dots, u_0, u_0, \dots, u_{w-1})$ satisfies $2 \leq y \leq (p-1)/2$ and $y^2 \equiv -1 \pmod{p}$.

Thus $x^2 \equiv y^2 \pmod{p}$, and so p divides $x^2 - y^2 = (x+y)(x-y)$. Since p is prime it follows that $x \equiv y$ or $x \equiv -y \pmod{p}$. But since x and y are both between 2 and $(p-1)/2$, we must have $x = y$. But then, by equation (6), the continued fraction

$$\begin{aligned} [r_t, \dots, r_0, r_0, \dots, r_t] &= \frac{K(r_t, \dots, r_0, r_0, \dots r_t)}{K(r_{t-1}, \dots, r_0, r_0, \dots r_t)} = \frac{p}{x} = \frac{p}{y} \\ &= \frac{K(u_w, \dots, u_0, u_0, \dots u_w)}{K(u_{w-1}, \dots, u_0, u_0, \dots u_w)} = [u_w, \dots, u_0, u_0, \dots, u_w], \end{aligned}$$

and by the uniqueness of finite continued fraction representations (with $r_t \geq 2$ and $u_w \geq 2$), we must have $t = w$ and $r_i = u_i$ for all $1 \leq i \leq t$. Consequently,

$$\frac{r}{s} = [r_0, \dots, r_t] = [u_0, \dots, u_w] = \frac{u}{v}$$

in lowest terms. Thus $r = u$ and $s = v$ as desired.

In summary, every prime number $p = 4m+1$ is the unique sum of two squares as follows. Let x be the unique solution to $x^2 \equiv -1 \pmod{p}$, where $2 \leq x \leq 2m$. [By Wilson's

Theorem, x will be the smallest positive integer congruent to $\pm(2m)!$ (mod p). We note that if a is any quadratic nonresidue of p, then x can be efficiently calculated. From Euler's criterion, $a^{(p-1)/2} \equiv -1 \pmod{p}$. Thus we can set x equal to the smallest positive integer congruent to $\pm a^{(p-1)/4} \pmod{p}$.] Then p/x has palindromic continued fraction representation $[r_t, \ldots, r_0, r_0, \ldots, r_t]$, and $[r_0, \ldots, r_t] = r/s$, where $r^2 + s^2 = p$.

References

[1] Arthur T. Benjamin and Jennifer J. Quinn, *Proofs that Really Count: The Art of Combinatorial Proof*, MAA, Washington, DC, 2003.

[2] Arthur T. Benjamin, Jennifer J. Quinn, and Francis Edward Su, Counting on Continued Fractions, *Mathematics Magazine,* Vol 73, No. 2, pp. 98-104, 2000.

[3] G. Chrystal, *Algebra: an elementary text-book for the higher classes of secondary schools and for colleges, Part II*, page 499, reprinted by Chelsea, N.Y., N.Y., 1964.

[4] F. W. Clarke, W. N. Everitt, L. L. Littlejohn, and S. J. R. Vorster, H. J. S. Smith and the Fermat Two Squares Theorem, *Amer. Math. Monthly* **106** (1999) 652–665.

[5] L. E. Dickson, *History of the Theory of Numbers, Vol. II*, American Mathematical Society, Providence, pp. 240-241.

[6] Henry J. S. Smith, De Compositione Numerorum Primorum $4\lambda + 1$ Ex Duobus Quardatis, *Crelle's Journal* V. 50 (1855), 91–92.