THE REPRESENTATION OF PRIMES OF THE FORM 4n + 1 AS THE SUM OF TWO SQUARES

Autor(en): **Barnes, C. W.**

Objekttyp: **Article**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **18 (1972)**

Heft 1: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **19.04.2024**

Persistenter Link: https://doi.org/10.5169/seals-45379

THE REPRESENTATION OF PRIMES OF THE FORM $4n + 1$ AS THE SUM OF TWO SQUARES

by C. W. Barnes

1. *Introduction.* Fermat stated that every prime of the form $4n + 1$ can be uniquely expressed as the sum of two squares. The first proof was given by Euler. There are four constructions for the integers x and y such that $p = x^2 + y^2$ where p is a prime of the form $4n + 1$; these are due to Legendre, Gauss, Serret, and Jacobsthal. We are concerned here with the work of Legendre and the variation of Serret's construction given by Smith.

Readable accounts of the proof of the Fermat theorem are readily available. The constructions, in particular those which use continued fractions, have not received much attention in the current literature. Davenport [1] and Olds [3] give a summary of Legendre's method.

Using results of Legendre [2] on solutions of the equation $u^2 - pv^2 = -1$, which give the characterization of the continued fraction for $\sqrt{p}$, we have obtained a construction similar to that of Legendre. It has the two advantages of giving explicit expressions for x and y, where $x^2 + y^2 = p$, in terms of the approximants of the continued fraction for $\sqrt{p}$, and of using fewer terms of that continued fraction.

Our method can be applied to obtain the continued fraction of Smith [6], the existence of which he established in his proof of the Fermat theorem.

A method of representing p as a sum of squares of rational numbers is given. We conclude with some examples.

2. *Continued Fractions.* The results we need can be found in Perron [4]. We denote the simple continued fraction

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

by $[a_0, a_1, \ldots, a_n]$. For $0 \leqq m \leqq n$ we denote the numerator and denominator of the mth approximant of $[a_0, a_1, \ldots, a_n]$ by A_m and B_m respectively. Hence

$$A_m = a_m A_{m-1} + A_{m-2}$$
$$B_m = a_m B_{m-1} + B_{m-2}$$

and

(2) $$A_m B_{m-1} - A_{m-1} B_m = (-1)^{m-1} .$$

Frequently we need to specify the terms of the continued fraction which are present in a numerator or a denominator of an approximant. To achieve this we use the standard notation $K(a_0, a_1, \ldots, a_m)$ to denote the numerator of the mth approximant to (1). Therefore

$$A_m = K(a_0, a_1, \ldots, a_m)$$
$$B_m = K(a_1, a_2, \ldots, a_m) .$$

That is, [4], page 7, the numerator of $[a_1, a_2, \ldots, a_m]$ is equal to the denominator of $[a_0, a_1, \ldots, a_m]$ and $K(a_1, a_2, \ldots, a_m)$ denotes this common value. Therefore, [4], page 12,

$$K(a_0, a_1, \ldots, a_m) = K(a_m, a_{m-1}, \ldots, a_0) .$$

With reference to (1) set

$$A_n = K(a_0, a_1, \ldots, a_n) = K(0, n) ,$$
$$B_n = K(a_1, a_2, \ldots, a_n) = K(1, n) ,$$

and

$$K(r, s) = K(a_r, a_{r+1}, \ldots, a_s) .$$

Then

$$K(0, n)\, K(l, m) - K(0, m)\, K(l, n)$$
(3) $$= (-1)^{m-l+1}\, K(0, l-2)\, K(m+2, n)$$

when

$$0 < l < m < n .$$

The limit of a periodic continued fraction is a quadratic irrational; conversely, the continued fraction of a quadratic irrational is periodic. If t has a purely periodic continued fraction then $t > 1$ and t', the conjugate of t, satisfies $-1 < t' < 0$. The conjugate of t is the negative reciprocal of the limit of the continued fraction obtained from that of t by reversing the period. This characterization of periodic continued fractions is a theorem of Lagrange, ([4], Satz 2, page 74), while the result on purely periodic continued fractions is due to Galois ([4], Satz 6, page 83).

If D is a positive integer which is not the square of an integer, then [5], Theorem 3, page 294,

$$\sqrt{D} = [a_0, \overline{a_1, a_2, ..., a_2, a_1, 2a_0}] ,$$

in the usual notation for a periodic continued fraction. The continued fraction has the form $\sqrt{D} = [a_0, \overline{2a_0}]$ if and only if $D = a^2 + 1$, as established in [5], page 298. We will not consider this case since no construction is necessary.

Legendre's construction depends on the fact that if p is a prime of the form $4n + 1$ then

$$\sqrt{p} = [a_0, \overline{a_1, ..., a_m, a_m, ..., a_1, 2a_0}] , \tag{4}$$

that is, the symmetric part of the period does not have a central term. This follows at once from the following two lemmas, both of which are proved in [4]. The first is Satz 18, page 103; the second is Satz 22, page 107.

LEMMA 1. The equation $x^2 - Dy^2 = -1$ has a solution if and only if the number of terms in the period of the simple continued fraction for $\sqrt{D}$ is odd.

LEMMA 2. If p is a prime of the form $4n + 1$, the equation $x^2 - py^2 = -1$ has a solution.

3. *The Construction for $p \neq a^2 + 1$.* We make use of the fact that the number t defined by

$$t = [\overline{a_m, ..., a_1, 2a_0, a_1, ..., a_m}]$$

is a reduced quadratic irrational and moreover (see [1], page 121, for instance) $t' = -\frac{1}{t}$.

In terms of the continued fraction for $\sqrt{p}$ we have

$$\sqrt{p} = [a_0, a_1, ..., a_m, \overline{a_m, ..., a_1, 2a_0, a_1, ..., a_m}] ,$$

that is, $\sqrt{p} = [a_0, a_1, ..., a_m, t]$, or $\sqrt{p} = \dfrac{tA_m + A_{m-1}}{tB_m + B_{m-1}}$. Hence

$$t = \frac{-A_m A_{m-1} + pB_m B_{m-1} + (-1)^{m-1}\sqrt{p}}{A_m^2 - pB_m^2}$$

where we used (2).

Therefore

$$t' = \frac{-A_m A_{m-1} + pB_m B_{m-1} + (-1)^{m}\sqrt{p}}{A_m^2 - pB_m^2},$$

and

$$-1 = tt' = \frac{(-A_m A_{m-1} + pB_m B_{m-1})^2 - p}{(A_m^2 - pB_m^2)^2}.$$

If we set

$$x = pB_m B_{m-1} - A_m A_{m-1} \tag{5}$$

$$y = A_m^2 - pB_m^2 \tag{6}$$

then $p = x^2 + y^2$. Hence (5) and (6) are the required integers. The work is valid for every prime of the form $4n + 1$. We note that these may be constructed from $m + 1$ terms of (4) instead of the usual $2m + 1$ terms.

We can now establish

THEOREM 1. Let p be a prime of the form $4n + 1$, and suppose

$$\sqrt{p} = [a_0, \overline{a_1, \ldots, a_m, a_m, \ldots, a_1, 2a_0}]$$

where $m > 0$. If

$$\frac{2B_m B_{m-1}}{B_m^2 - B_{m-1}^2} = [b_0, b_1, \ldots, b_n]$$

then

$$p = K(b_0, b_1, \ldots, b_{n-1})^2 + K(b_1, b_2, \ldots, b_{n-1})^2.$$

Proof. It follows from (5), (6), and the equation $p = x^2 + y^2$ that

$$p = \frac{A_m^2 + A_{m-1}^2}{B_m^2 + B_{m-1}^2}.$$

We see this as follows. Solving the equation $\sqrt{p} = \dfrac{tA_m + A_{m-1}}{tB_m + B_{m-1}}$ for t we get $t = \dfrac{A_{m-1} - B_{m-1}\sqrt{p}}{-A_m + B_n\sqrt{p}}$. Hence

$$-\frac{1}{t} = \frac{A_m A_{m-1} - pB_m B_{m-1} + (-1)^{m-1}\sqrt{p}}{A_{m-1}^2 - pB_{m-1}^2},$$

and also

$$t' = \frac{A_m A_{m-1} - pB_m B_{m-1} + (-1)^{m-1}\sqrt{p}}{-A_m^2 + pB_m^2},$$

where we rationalized each denominator and made use of (2). Finally, since $t' = -\dfrac{1}{t}$ we must have $A_{m-1}^2 - pB_{m-1}^2 = -A_m^2 + pB_m^2$ and the result follows. Using this, we can write x and y in the form

$$x = \frac{A_m B_m - A_{m-1}B_{m-1}}{B_m^2 + B_{m-1}^2}, \tag{7}$$

$$y = \frac{A_m B_{m-1} + A_{m-1}B_m}{B_m^2 + B_{m-1}^2}. \tag{8}$$

Solving (7) and (8) for A_m and A_{m-1} in terms of x and y we get

$$B_m x + B_{m-1} y = (-1)^{m-1} A_m \tag{9}$$

$$-B_{m-1}x + B_m y = (-1)^{m-1} A_{m-1}, \tag{10}$$

and therefore multiplying (9) by B_{m-1}, (10) by B_m, and using (2)

$$2B_m B_{m-1} x - (B_m^2 - B_{m-1}^2)\, y = 1. \tag{11}$$

Therefore $2B_m B_{m-1}$ and $B_m^2 - B_{m-1}^2$ are relatively prime.

We now, however, consider (11) as an indeterminate equation. It can be solved by the usual technique of converting $\dfrac{2B_m B_{m-1}}{B_m^2 - B_{m-1}^2}$ into a continued fraction. Since $p \neq a^2 + 1$, $B_{m-1} \neq 0$. Thus suppose

$$\frac{2B_m B_{m-1}}{B_m^2 - B_{m-1}^2} = [b_0, b_1, \dots, b_n].$$

Since $(2B_m B_{m-1}, B_m^2 - B_{m-1}^2) = 1$, the integers

$$x_0 = K(b_0, b_1, \ldots, b_{n-1})$$
$$y_0 = K(b_1, b_2, \ldots, b_{n-1}),$$

with a possible change of sign of one of them, form a solution of (11).

Denote the approximants of $[b_0, b_1, \ldots, b_n]$ by $\frac{P_k}{Q_k}$. Then

$$P_n = 2B_m B_{m-1},$$

and

$$Q_n = B_m^2 - B_{m-1}^2.$$

For every integer r, the integers

(12) $$u = Q_{n-1} - Q_n r, \; v = P_{n-1} - P_n r,$$

with an adjustment of sign, satisfy (11). We now determine the unique integer r such that $u^2 + v^2 = p$. We obtain

(13) $$(P_n^2 + Q_n^2) r^2 - 2(P_n P_{n-1} + Q_n Q_{n-1}) r + (P_{n-1}^2 + Q_{n-1}^2 - p) = 0.$$

Since (13) will have one integral root, the other root is rational. Let r_0 be the integral root and suppose that $r_0 \neq 0$. Then $P_{n-1}^2 + Q_{n-1}^2 - p \neq 0$ and

$$r_0 \mid (P_{n-1}^2 + Q_{n-1}^2 - p).$$

There is an integer $s \neq 0$ such that $P_{n-1}^2 + Q_{n-1}^2 - p = sr_0$. Therefore (13) vanishes when

$$r = \frac{P_{n-1}^2 + Q_{n-1}^2 - p}{s}$$

and thus the integer s is determined by

(14) $$s^2 - 2(P_n P_{n-1} + Q_n Q_{n-1}) s + (P_n^2 + Q_n^2)(P_{n-1}^2 + Q_{n-1}^2 - p) = 0.$$

The discriminant of (14) is $4\{(P_n^2 + Q_n^2) p - 1\}$ and is not zero. It follows that there exist two distinct integers s which satisfy (14). This gives rise to two distinct integral roots of (13) and contradicts the uniqueness of the representation of p as the sum of two squares.

Thus the integral root of (13) is zero, and

$$p = P_{n-1}^2 + Q_{n-1}^2.$$

That is

$$p = K(b_0, b_1, \dots, b_{n-1})^2 + K(b_1, b_2, \dots, b_{n-1})^2 .$$

When $2B_mB_{m-1} < B_m^2 - B_{m-1}^2$ we alter the construction and obtain the continued fraction for $\dfrac{B_m^2 - B_{m-1}^2}{2B_mB_{m-1}}$.

COROLLARY 1. If r_1 is the rational root of (13) and u_1 and v_1 are the corresponding rational numbers in (12) then

$$p = u_1^2 + v_1^2$$

is a representation of p as the sum of squares of two rational numbers.

Sierpiński [5], page 353, gives a method for obtaining infinitely many such rational representations once an initial pair u_1, v_1 is known.

We also have

COROLLARY 2. In the representation $p = x^2 + y^2$, one of the integers x and y is a quadratic residue modulo p.

Proof. This follows directly from (6).

If $p = x^2 + y^2$ and y is the integer which is a quadratic residue, x may or may not be a quadratic residue modulo p. For example, we have $13 = 2^2 + 3^2$ and 2 is a quadratic nonresidue modulo 13. However, $41 = 4^2 + 5^2$ and each of 4 and 5 is a quadratic residue modulo 41.

4. *The Continued Fraction of Smith.* Smith [6] proved the Fermat theorem by means of a finite continued fraction. Suppose that $c_1, c_2, \dots, c_l$ are the positive integers such that $(c_i, p) = 1$ and $c_i < \dfrac{p}{2}$ where $1 \leqq i \leqq l$.

Construct the continued fraction for $\dfrac{p}{c_i}$ where we require that the last term in each is greater than 1. The last approximant of each of these continued fractions has the form $\dfrac{K(d_0, d_1, \dots, d_m)}{K(d_1, d_2, \dots, d_m)}$ where $d_0 > 1$, $d_m > 1$. Hence the number of ways of expressing p in the form $p = K(d_0, d_1, \dots, d_m)$ where $d_0 > 1$ and $d_m > 1$ is the number of integers less than $p/2$ and relatively prime to p. However, since $K(d_0, d_1, \dots, d_m) = K(d_m, d_{m-1}, \dots, d_0)$, we see that $K(d_m, d_{m-1}, \dots, d_0)$ must be present as a numerator in one of the continued fractions for some p/c_i. That is, for each i, $1 \leq i \leq l$, there is some j in this same set such that the terms in the continued fraction for p/c_j are obtained from those for the continued fraction for p/c_i by reversing their order.

But we have $[p/2] = 2n$, so the number of equations of the form $K(d_0, d_1, \ldots, d_m) = p$ must be even. Inasmuch as $K(p)$ is among them we see that among those which remain, there is one which does not change when we reverse the order of its terms. That is, the terms in the particular numerator are symmetric. Hence in the notation of section 2 we have

$$p = K(d_0, d_1, \ldots, d_j, d_j, \ldots, d_1, d_0) . \tag{15}$$

Otherwise, we would have

$$p = K(d_0, d_1, \ldots, d_{j-1}, d_j, d_{j-1}, \ldots, d_1, d_0) .$$

We note, however, by (3), where l, m, and n of that equation have values as indicated

$$d_0, \ldots, \overset{(l)}{d_{j-1}}, \overset{(m)}{d_j}, d_{j-1}, \ldots, \overset{(n)}{d_0} ,$$

that

$$\begin{aligned} pK(d_j, d_{j-1}) &- K(d_0, \ldots, d_j, d_{j-1})\, K(d_j, d_{j-1}, \ldots, d_0) \\ &= (-1)^{j+1-j+1}\, K(d_0, \ldots, d_{j-2})\, K(d_{j-3}, \ldots, d_0) . \end{aligned}$$

We use the facts that

$$\begin{aligned} K(d_j, d_{j-1}) &= d_j d_{j-1} + 1 , \\ K(d_{j-3}, \ldots, d_0) &= K(d_0, \ldots, d_{j-3}) , \\ K(d_0, \ldots, d_{j-1}) &= d_{j-1}\, K(d_0, \ldots, d_{j-2}) + K(d_0, \ldots, d_{j-3}) , \end{aligned}$$

and the above becomes

$$\begin{aligned} p(d_j d_{j-1} + 1) = (d_j d_{j-1} + 1)\{ & K(d_0, \ldots, d_{j-1})\, K(d_0, \ldots, d_{j-2}) \\ & + K(d_0, \ldots, d_j)\, K(d_0, \ldots, d_{j-1}) \} . \end{aligned}$$

That is,

$$p = K(d_0, \ldots, d_{j-1}) \{ K(d_0, \ldots, d_j) + K(d_0, \ldots, d_{j-2}) \} .$$

By the conditions on the terms in the continued fraction, this contradicts the hypothesis that p is a prime.

Therefore, by (3), with values of l, m, and n as indicated,

$$\overset{}{d_0}, d_1, \ldots, \overset{(l)}{d_j}, \overset{(m)}{d_j}, \ldots, d_1, \overset{(n)}{d_0},$$

we have from (15)

$$K(d_0, \ldots, d_j, d_j, \ldots, d_0)\, K(d_j, d_j) - K(d_0, \ldots, d_j, d_j)\, K(d_j, d_j, \ldots, d_0)$$
$$= (-1)^{j+1-j+1}\, K(d_0, \ldots, d_{j-2})\, K(d_{j-2}, \ldots, d_0).$$

Now

$$K(d_j, d_j) = d_j^2 + 1, K(d_j, d_j, \ldots, d_0) = K(d_0, \ldots, d_j, d_j),$$
$$K(d_{j-2}, \ldots, d_0) = K(d_0, \ldots, d_{j-2}),$$

and thus

$$p(d_j^2 + 1) - K(d_0, \ldots, d_j, d_j)^2 = K(d_0, \ldots, d_{j-2})^2.$$

Substituting $K(d_0, \ldots, d_{j-2}) = K(d_0, \ldots, d_j) - d_j K(d_0, \ldots, d_{j-1})$
we obtain

$$p(d_j^2 + 1) = K(d_0, \ldots, d_j)^2 (d_j^2 + 1) + K(d_0, \ldots, d_{j-1})^2 (d_j^2 + 1),$$

and we have

$$p = K(d_0, \ldots, d_j)^2 + K(d_0, \ldots, d_{j-1})^2.$$

We can apply Theorem 1 to obtain explicitly the continued fraction with this property and also the integer c, $c < \frac{p}{2}$, $(c, p) = 1$ such that

$$\frac{p}{c} = [d_0, d_1, \ldots, d_j, d_j, \ldots, d_1, d_0].$$

THEOREM 2. Let p be a prime of the form $4n + 1$, and such that $p \neq a^2 + 1$.

If

$$\sqrt{p} = [a_0, \overline{a_1, \ldots, a_m, a_m, \ldots, a_1, 2a_0}]$$

where $m > 0$,

$$B_m = K(a_1, a_2, \ldots, a_m)$$
$$B_{m-1} = K(a_1, a_2, \ldots, a_{m-1}),$$

and

$$\frac{2B_m B_{m-1}}{B_m^2 - B_{m-1}^2} = [b_0, b_1, \dots, b_n],$$

then

$$[b_{n-1}, \dots, b_0, b_0, \dots, b_{n-1}]$$

is the continued fraction of Smith.

Proof. As in the previous application of (3) we have

$$\begin{aligned} K(b_{n-1}, \dots, b_0, b_0, \dots, b_{n-1}) &= K(b_{n-1}, \dots, b_0)^2 + K(b_{n-1}, \dots, b_1)^2 \\ &= K(b_0, \dots, b_{n-1})^2 + K(b_1, \dots, b_{n-1})^2 = p, \end{aligned}$$

by Theorem 1. The result now follows by the uniqueness of the representation of p as a sum of two squares.

We now have

COROLLARY 3. $K(b_{n-2}, \dots, b_0, b_0, \dots, b_{n-1})$ is the integer c such that $\frac{p}{c}$ gives rise to Smith's continued fraction.

We remark that if $p = a^2 + 1$ then $p = K(a, a)$ and $\frac{p}{a} = [a, a]$. Thus we constructed Smith's continued fraction in all cases.

5. *Examples.* Let $p = 53$. Then

$$\sqrt{53} = [7, \overline{3, 1, 1, 3, 14}].$$

The appropriate continued fraction to take is [7, 3, 1]; $m = 2$, $B_2 = 4$, $B_1 = 3$.

$$\frac{2B_2 B_1}{B_2^2 - B_1^2} = \frac{24}{7} = [3, 2, 7], n = 2, P_1 = K(3, 2) = 7.$$

$Q_1 = K(2) = 2$ and $53 = 2^2 + 7^2$. Also $53 = K(2, 3, 3, 2)$. The equation (13) becomes

$$2826r^2 - 774r = 0,$$

since $P_2 = 51$, $Q_2 = 15$. The rational root is $\frac{43}{157}$.

Set $u_1 = 2 - 15\frac{43}{157} = -\frac{331}{157}$ and $v_1 = 7 - 51\frac{43}{157} = -\frac{1094}{157}$. Then $u_1^2 + v_1^2 = 53$, illustrating Corollary 1.

Finally consider $p = 61$.

$$\sqrt{61} = [7, \overline{1, 4, 3, 1, 2, 2, 1, 3, 4, 1, 14}] .$$

Form [7, 1, 4, 3, 1], where $m = 5$. $B_5 = 58$, $B_4 = 21$, $2B_4B_5 = 2436$, and $B_5^2 - B_4^2 = 2923$. Since $2B_4B_5 < B_5^2 - B_4^2$, we form the continued fraction for $\frac{B_5^2 - B_4^2}{2B_4B_5}$. Thus $\frac{2923}{2436} = [1, 5, 487]$. We must compute $K(1, 5) = 6$ and $K(5) = 5$. Then $61 = 5^2 + 6^2$ and also $61 = K(5, 1, 1, 5)$.

REFERENCES

[1] Davenport, H. The Higher Arithmetic, Hutchinson's University Library, London, 1952.

[2] Legendre, A. M. Théorie des Nombres. Troisième édition, Paris, 1830.

[3] Olds, C. D. Continued Fractions. Random House, New York, 1963.

[4] Perron, Oskar. Die Lehre von den Kettenbrüchen. Chelsea, New York, 1951.

[5] Sierpiński, Waclaw. Elementary Theory of Numbers. Państwowe Wydawnictwo Naukowe, Warsaw, 1964.

[6] Smith, J. De Composition Numerorum Primorum Formae $4\lambda+1$ Ex Duobus Quadratis. *Journal für die Reine und Angewandte Mathematik*, 50 (1855), pp. 91-92.

(Reçu le 20 septembre 1972)

C. W. Barnes
The University of Mississippi
Department of Mathematics
University, Mississippi, 38677
USA.

vide-leer-empty