H. J. S. Smith and the Fermat Two Squares Theorem

F. W. Clarke, W. N. Everitt, L. L. Littlejohn, and S. J. R. Vorster

This article is dedicated to Professor P. R. Halmos

1. INTRODUCTION. In his remarkable book *A Mathematician's Apology*, G. H. Hardy wrote [**12**, p. 97]

> Another famous and beautiful theorem is Fermat's 'two square' theorem. The primes may (if we ignore the special prime 2) be arranged in two classes; the primes
>
> $$5, 13, 17, 29, 37, 41, \ldots$$
>
> which leave remainder 1 when divided by 4, and the primes
>
> $$3, 7, 11, 19, 23, 31, \ldots$$
>
> which leave remainder 3. All the primes of the first class, and none of the second, can be expressed as the sum of two squares: thus
>
> $$5 = 1^2 + 2^2, \quad 13 = 2^2 + 3^2$$
> $$17 = 1^2 + 4^2, \quad 29 = 2^2 + 5^2$$
>
> but 3, 7, 11 and 19 are not expressible in this way (as the reader may check by trial). This is Fermat's theorem, which is ranked, very justly, as one of the finest of arithmetic. Unfortunately there is no proof within the comprehension of anybody but a fairly expert mathematician.

The history of this theorem of Fermat is given in detail by Dickson [**7**, 224–237]. Dickson names the theorem after Girard, who discussed the result in 1632; however the common practice now is to attribute the result to Fermat, who stated in 1659 that he possessed an irrefutable proof by the method of infinite descent; see [**7**, p. 228] and [**2**, p. 89]. The first recorded proof is due to Euler given in 1749 [**7**, pp. 230–231]; Bell writes, "It was first proved by the great Euler in 1749 after he had struggled, off and on, for *seven years* to find a proof" [**2**, p.89]. The first proof that such prime numbers can be *uniquely* represented as the sum of squares of two positive integers was given by Gauss in 1801 [**7**, p. 233]. See also the account of the two squares theorem of Fermat in the books by Burton [**4**, Chapter 12, Section 2], and Hardy and Wright [**14**, Chapter XX].

The last sentence in the quotation from Hardy is significant. Hardy had an interest in the classification of proof; see, in particular, [**13**, p. 6] in connection with the "elementary" proof of inequalities. In this context the word *elementary* must not be confused with the words *obvious* or *easy*; many of the elementary proofs in [**13**] are subtle, ingenious, and far from obvious. When Hardy wrote [**12**] he was, more than likely, not aware that an elementary proof of this theorem of Fermat had been given in 1855 by H. J. S. Smith, one of his predecessors in the Savilian Chair of Geometry in the University of Oxford. This simple but remarkable proof of Smith is within the comprehension of those with knowledge of elementary

algebra, including simple properties of determinants, and the fundamental theorem of arithmetic [**6**, Chapter I, Section 4]. The proof is also remarkable for giving a construction that permits one to compute the integers of the two squares representation.

In this paper we give Smith's proof of the theorem of Fermat and present what is, possibly, a new elementary proof of the uniqueness of the two squares representation, but now using Smith's ideas and method. This uniqueness proof involves the Euler Criterion [**8**, Section 11] for solutions of the quadratic equation $x^2 \equiv -1 \pmod{p}$; we present a new existence proof that leads to a constructable solution of this equation.

The original paper of Smith [**20**] is (the good news) only 2 pages long but is (the bad news for most of us) written in Latin; see also the collected works of Smith [**21**], in which [**20**] appears as the second contribution. The Smith proof has not gone entirely without notice; Chrystal [**5**, p. 471] reproduces the proof in English, as does, in part, Dickson [**7**, pp. 240–241]; Davenport mentions the proof [**6**, p. 122] but does not give complete details. Barnes [**1**] gives an exposition of Smith's existence theorem, and establishes the connection between the Smith palindromic continuant and the Euler Criterion (see our Theorems 1 and 2 and their proofs).

Both Serret [**19**] and Hermite [**17**] use ideas similar to the Smith method [**20**] to give an algorithm for finding the integers in the two squares representation of the theorem of Fermat. This method was subsequently improved by Brillhart [**3**] to give an impressively fast numerical procedure to determine the representation; as an example the Brillhart method gives

$$10^{50} + 577 = 761106534380835424545040 1^2 + 648626890687392164224542 4^2. \tag{1.1}$$

The two squares theorem of Fermat continues to attract attention; see the recent contributions by Ewell [**9**], Heath-Brown [**16**], Wagon [**22**], and Zagier [**23**].

In Section 2 we give formal statements of the results to be proved by the Smith methods. In Section 3 we give a brief account of the life of Henry Smith. In Section 4 there is a definition and statement of the properties of continuants. The remaining sections are devoted to proofs of the results. Lastly, in an appendix, we reproduce the original Smith paper [**20**].

At the end of Sections 6, 7, and 8 we exemplify the general results by considering the case $p = 13$, and other cases including the example in (1.1).

2. STATEMENT OF RESULTS. Let $\mathbb{N} := \{1, 2, 3, \ldots\}$ and $\mathbb{P} := \{p \in \mathbb{N} : p \text{ is a prime number}\}$.

Theorem 1 [Fermat and Gauss]. *Let $p \in \mathbb{P}$ with $p \equiv 1 \pmod{4}$. Then there exist two unique, positive, co-prime integers $u, v \in \mathbb{N}$ such that*

$$p = u^2 + v^2.$$

Proof: See Sections 6 and 7. ∎

Theorem 2 [The Euler Criterion]. *Let $p \in \mathbb{P}$ with $p \equiv 1 \pmod{4}$. Then*

1. *The quadratic equation*

$$x^2 \equiv -1 \pmod{p} \tag{2.1}$$

has two unique solutions $x_0, x_1 \in \mathbb{N}$ such that

$$1 < x_0 < (p-1)/2 \quad \text{and} \quad (p-1)/2 < x_1 < p,$$

with $x_1 = p - x_0$.

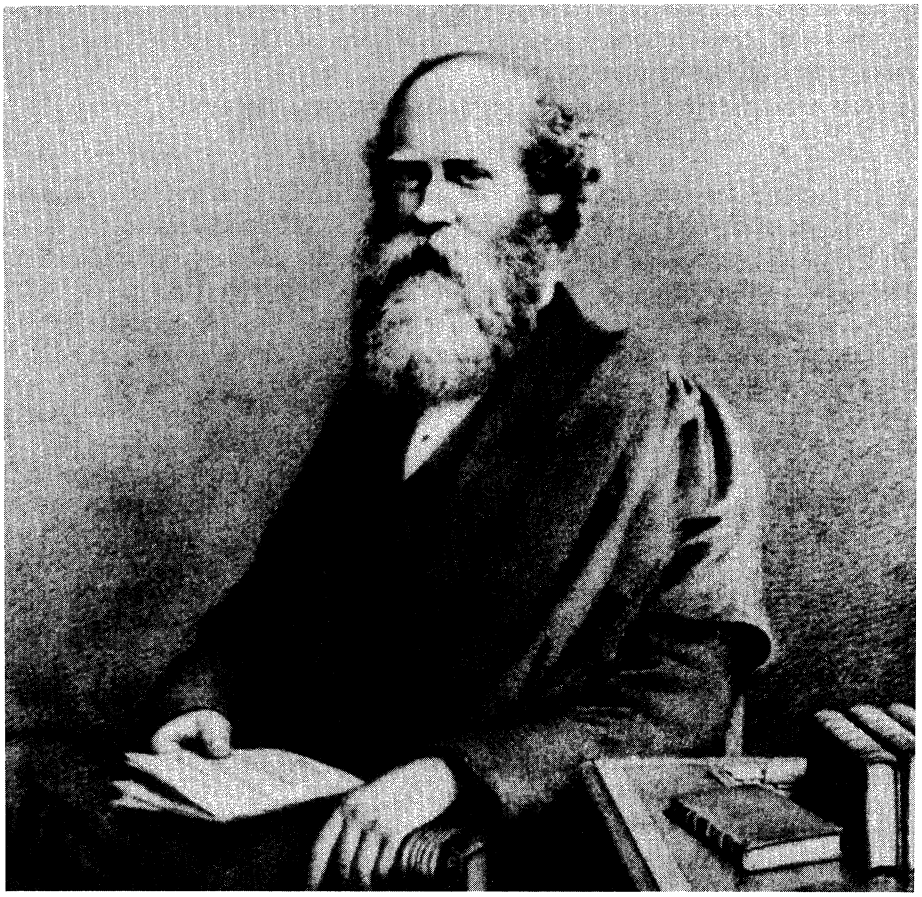
2. *All other solutions of (2.1) are congruent to x_0 or $x_1 \pmod{p}$.*

Proof: See Section 8. ■

For a detailed discussion of the Euler Criterion see the book by Dudley [**8**, pp. 85-86].

2. HENRY JOHN STEPHEN SMITH. Henry Smith was born on 2 November 1826 in Dublin, Ireland. His father died soon afterwards and the widow moved with her family to England. Smith was educated first by his mother and then by a succession of private tutors, before spending three years at Rugby School; from this School he gained entry to the University of Oxford, in 1844, by winning the top scholarship to Balliol College. In 1848 at Oxford he gained first class honours in both classics and mathematics; he also won the major University prizes in both these subjects, the Ireland scholarship in classics, and the Senior Mathematical Scholarship in mathematics.

In 1849 the Balliol College fellowships in classics and mathematics fell vacant; until this time Smith seems to have been undecided as to whether to follow a career in classics or mathematics, but seems to have settled at this time on mathematics. His first paper, on geometry, dates from the next year.



Henry John Stephen Smith (1826–1883).

Smith completed his first paper on number theory in 1854 and published it the following year in *Crelle's Journal* [**20**]. Unusually, even for that time, it was written in Latin, perhaps in homage to Carl Friedrich Gauss, whose *Disquisitiones arithmeticae* served as an inspiration.

In 1859 Smith was elected to Fellowship of the Royal Society of London and then, in 1860, to the Savilian Chair of Geometry in the University of Oxford; two of his eventual successors to this Chair were G. H. Hardy and E. C. Titchmarsh.

Henry Smith died in Oxford in 1883.

It is puzzling that Henry Smith's work and name are so little known, even amongst those who make regular use of the ideas that he introduced; this point is discussed by Keith Hannabuss in his paper *The mathematician the world forgot* [**10**]. Several historians of mathematics have ranked Smith with Cayley and Sylvester among the great pure mathematicians of the nineteenth century. In 1861 Smith proved the existence and uniqueness of what is now called the *Smith Normal Form* of a matrix with integer entries. This result has subsequently been used to prove the cyclic decomposition for modules, but Smith's first application was to determine when linear Diophantine equations admit solutions, settling a longstanding problem first studied by Greek mathematicians. His remarkable contributions to, and his panoramic knowledge of, the theory of numbers can be seen in the monumental *Report on the theory of numbers*, reproduced in [**21**]. In this area, in 1868, he shared the Steiner Prize of the Royal Academy of Sciences in Berlin for his solution of a geometric problem that involved the representation of integers as a sum of squares.

Not so well known is Smith's early contribution to measure theory and integration in his paper of 1875 *On the integration of discontinuous functions*; see [**21**, paper 25]. There, Smith introduced the first example of what is now called a Cantor set; Cantor's own example appeared eight years later and was not presented as his own discovery. Smith's example divides an interval into $m > 2$ subintervals, and then keeps repeating this process to each remaining subinterval, except the last. Smith also seems to have been the first mathematician to perceive the connection between measure and integral. However, his paper received less attention than it deserved, owing to an inaccurate review in the *Fortschritte der Mathematik*. In his history of integration, see [**15**, pp. 37, 40], Thomas Hawkins has remarked:

> Probably the development of a measure-theoretic viewpoint within integration theory would have been accelerated had the contents of Smith's paper been known to mathematicians whose interest in the theory was less tangential than Smith's.

For an informed discussion on the contents of this paper of Smith, and for the development of the ideas therein to higher dimensions, see [**10**] and, especially, [**11**].

4. CONTINUANTS. Continuants are closely connected with continued fractions, as noted by Smith at the beginning of [**20**]. There is a detailed and elegant account of this connection in Chrystal [**5**, Chapter XXXIV, Sections 4–11]. However Smith uses only continuants in his paper and uses determinants to define them; for this definition see [**5**, Chapter XXXIV, Section 11] and the reference therein to the

remarkable history of determinants by Muir and Metzler [**18**, Chapters III and XIII]. We follow Smith and make the

Definition 1. For $n \in \mathbb{N}$ let $q_r \in \mathbb{N}$ $(r = 1, 2, \ldots, n)$; then define $[\cdot]: \mathbb{N}^n \to \mathbb{N}$ by the determinant

$$[q_1, q_2, q_3, \ldots, q_{n-1}, q_n] := \begin{vmatrix} q_1 & 1 & 0 & \cdots & 0 & 0 \\ -1 & q_2 & 1 & \cdots & 0 & 0 \\ 0 & -1 & q_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & q_{n-1} & 1 \\ 0 & 0 & 0 & \cdots & -1 & q_n \end{vmatrix} \tag{4.1}$$

We note that

$$[q_1] = q_1, \quad [q_1, q_2] = q_1 q_2 + 1, \quad \text{and } [q_1, q_2, q_3] = q_1 q_2 q_3 + q_1 + q_3. \tag{4.2}$$

Lemma 1. *Let $n \in \mathbb{N}$ with $n \geq 2$. Then*

1. $[q_1, q_2, \ldots, q_n] = [q_1][q_2, q_3, \ldots, q_n] + [q_3, \ldots, q_n]$
2. $[q_1, q_2, \ldots, q_n] \in \mathbb{N}$
3. $[q_1, q_2, \ldots, q_n] = [q_n, \ldots, q_2, q_1]$
4. $[q_2, q_3, \ldots, q_n] < [q_1, q_2, \ldots, q_n]$
5. $[q_2, q_3, \ldots, q_n]$ *and* $[q_1, q_2, \ldots, q_n]$ *are co-prime integers*
6. $$[q_1, \ldots, q_{s-1}, q_s, q_{s+1}, q_{s+2}, \ldots, q_n] = [q_1, \ldots, q_{s-1}, q_s][q_{s+1}, q_{s+2}, \ldots, q_n] + [q_1, \ldots, q_{s-1}][q_{s+2}, \ldots, q_n].$$

Proof: Note that if in any formula in Lemma 1 an empty continuant appears then it is convenient, and consistent, to give such a continuant the value 1.

1. Expand the determinant (4.1) by the first row.
2. Use (4.2), property 1 and mathematical induction.
3. Standard property of determinants.
4. Use properties 1 and 2.
5. Use property 1.
6. Use the Laplace expansion on (4.1) centred on row s; see also the proof in [**5**, Chapter XXXIV, Section 6].

5. THE EUCLIDEAN ALGORITHM

Algorithm 1. *Let $r, s \in \mathbb{N}$ be co-prime with $s < r$ and write*

$$\frac{r}{s} = q_1 + \frac{t}{s} \quad (0 < t < s), \frac{s}{t} = q_2 + \frac{u}{t} \quad (0 \leq u < t), \ldots, \frac{v}{w} = q_n + \frac{0}{w} = q_n \tag{5.1}$$

for some $n \in \mathbb{N}$ with $n \geq 2$, $q_i \in \mathbb{N}$ $(i = 1, 2, \ldots, n)$, and $q_n \geq 2$.

Thus a rational number $r/s > 1$ is associated with a set of positive integers $\{q_1, q_2, \ldots, q_n\}$ satisfying the properties in (5.1). Conversely we have

Lemma 2. *Let a set of positive integers* $\{q_1, q_2, \ldots, q_n\}$ *be given with* $n \geq 2$ *and* $q_n \geq 2$. *Then there is a unique rational number* $r/s > 1$ *whose Euclidean algorithm yields the set* $\{q_1, q_2, \ldots, q_n\}$; *moreover,* r/s *is determined by*

$$\frac{r}{s} = \frac{[q_1, q_2, \ldots, q_n]}{[q_2, q_3, \ldots, q_n]}. \tag{5.2}$$

Here r *and* s *are co-prime and given by*

$$r = [q_1, q_2, \ldots, q_n] \quad \textit{and} \quad s = [q_2, q_3, \ldots, q_n]. \tag{5.3}$$

Proof: Define r/s by (5.2) and apply property 1 of Lemma 1 n times. The result (5.3) follows from property 5 of Lemma 1. ∎

6. THE SMITH PROOF OF THE FERMAT THEOREM

Proof: Let $p \in \mathbb{P}$ with $p \equiv 1 \pmod 4$ and write $p = 4r + 1$. Let the number μ be taken arbitrarily from the set of positive integers $\{1, 2, \ldots, 2r\}$ and consider the corresponding set of rational numbers $\{p/\mu\}$, noting that $2 < p/\mu \leq p$. If we apply Algorithm 1 to p/μ we obtain a representation of the form

$$\frac{p}{\mu} = \frac{[q_1, q_2, \ldots, q_n]}{[q_2, \ldots, q_n]}. \tag{6.1}$$

Of course, the integer n and the set $\{q_1, q_2, \ldots, q_n\}$ depend upon the particular choice of μ. From property 5 of Lemma 1 we obtain

$$p = [q_1, q_2, \ldots, q_n] \quad \text{and} \quad \mu = [q_2, \ldots, q_n]. \tag{6.2}$$

From Algorithm 1 and from property 1 of Lemma 1, since $p/\mu > 2$, it follows that, in the representation (6.2),

$$q_1 \geq 2 \quad \text{and} \quad q_n \geq 2. \tag{6.3}$$

Now take one of the rational numbers p/μ with

$$\mu \in \{2, 3, \ldots, 2r\};$$

then we have the following chain of argument, using property 3 of Lemma 1 and (6.1),

$$\frac{p}{\mu} = \frac{[q_1, q_2, \ldots, q_n]}{[q_2, \ldots, q_n]} \Rightarrow [q_1, q_2, \ldots, q_n] = p = [q_n, q_{n-1}, \ldots, q_1]$$

$$\Rightarrow \frac{[q_n, q_{n-1}, \ldots, q_1]}{[q_{n-1}, \ldots, q_1]} = \frac{p}{\nu}, \tag{6.4}$$

(say). It follows from (6.3), Lemma 2, and property 1 of Lemma 1 that $1 < \nu < p/2$, so $\nu \in \{2, 3, \ldots, 2r\}$. Thus the chain of argument that gave (6.4) can be reversed, starting with ν and finishing with μ.

This argument pairs off the elements of the set $\{2, 3, \ldots, 2r\}$ and gives each member μ of the set a unique mate ν in the set. However this set contains an odd number of elements so there must exist at least one member, say λ, that mates with itself in the chain (6.4). For this λ we obtain from (6.4)

$$\frac{[q_1, q_2, \ldots, q_n]}{[q_2, \ldots, q_n]} = \frac{p}{\lambda} = \frac{[q_n, q_{n-1}, \ldots, q_1]}{[q_{n-1}, \ldots, q_1]}. \tag{6.5}$$

Now apply Algorithm 1 to both sides of (6.5) to give a representation

$$p = [q_1, q_2, \ldots, q_n], \tag{6.6}$$

with the palindromic property, and with (6.3) holding,

$$q_i = q_{n+1-i} \quad (i = 1, 2, \ldots, n). \tag{6.7}$$

If, in (6.7), $n = 2t + 1$ is odd then $n \geq 3$ and the representation (6.6) takes the form, for $s \geq 2$,

$$p = [q_1, \ldots, q_{s-1}, q_s, q_{s-1}, \ldots, q_1].$$

Now apply property 6 of Lemma 1 to give

$$p = [q_1, \ldots, q_{s-1}, q_s][q_{s-1}, \ldots, q_1] + [q_1, \ldots, q_{s-1}][q_{s-2}, \ldots, q_1].$$

Other properties of Lemma 1 permit us to write

$$p = [q_1, \ldots, q_{s-1}]\{[q_1, \ldots, q_{s-1}, q_s] + [q_{s-2}, \ldots, q_1]\},$$

which represents the prime number p as the product of two factors that, using (6.3), are both greater than 1; this is a contradiction to $p \in \mathbb{P}$.

Thus in (6.7) the integer $n = 2t$ must be even and so (6.6) takes the form, for $s \geq 1$,

$$p = [q_1, \ldots, q_s, q_s, \ldots, q_1]$$

with $q_1 \geq 2$ from (6.3). Now apply property 6 of Lemma 1 to give

$$p = [q_1, \ldots, q_s][q_s, \ldots, q_1] + [q_1, \ldots, q_{s-1}][q_{s-1}, \ldots, q_1]$$

and then

$$p = [q_1, \ldots, q_s]^2 + [q_1, \ldots, q_{s-1}]^2.$$

From property 1 it follows that $[q_1, \ldots, q_{s-1}]$ and $[q_1, \ldots, q_s]$ are co-prime.

This completes Smith's proof of the Fermat part of Theorem 1. ■

Consider the case $p = 13$. Then $\mu \in \{2, 3, 4, 5, 6\}$ and the application of the Euclidean algorithm to each choice of μ gives

$$\begin{array}{llllll}
\mu = 2 & n = 2 & q_1 = 6 & q_2 = 2 & & \\
\mu = 3 & n = 2 & q_1 = 4 & q_2 = 3 & & \\
\mu = 4 & n = 2 & q_1 = 3 & q_2 = 4 & & \\
\mu = 5 & n = 4 & q_1 = 2 & q_2 = 1 & q_3 = 1 & q_4 = 2 \\
\mu = 6 & n = 2 & q_1 = 2 & q_2 = 6. & &
\end{array}$$

Thus in the Smith pairing, 2 pairs with 6, 3 pairs with 4, and 5 pairs with itself. The palindromic continuant given by 5 then yields the two squares result

$$13 = [2, 1, 1, 2] = [2, 1][1, 2] + [2][2] = [2, 1]^2 + [2]^2 = 3^2 + 2^2.$$

7. A COROLLARY. We have

Corollary 1. *Let $p \in \mathbb{P}$ with $p = 4r + 1$. Then there are exactly $2r$ distinct continuant representations of p*

$$p = [q_1, \ldots, q_n]$$

with $q_n \geq 2$.

Proof: Let $\mu \in \{1, 2, \ldots, 2r\}$; then from (6.1)

$$\frac{p}{\mu} = \frac{[q_1, q_2, \ldots, q_n]}{[q_2, \ldots, q_n]} \quad \text{and} \quad p = [q_1, q_2, \ldots, q_n] \tag{7.1}$$

with $q_n \geq 2$; these continuant representations of p are distinct since otherwise, from (7.1), $p/\mu = p/\mu'$ for $\mu \neq \mu'$.

Let p have a representation $p = [q_1, q_2, \ldots, q_n]$ with $q_n \geq 2$. If $n = 1$ then $q_1 = q_n = p$ and we can take $\mu = 1$ in (7.1). If $n \geq 2$ then, since $q_n \geq 2$, it follows from properties 1 and 3 of Lemma 1 that $[q_2, \ldots, q_n] \leq (p-1)/2$ so that in (7.1) we have $[q_2, \ldots, q_n] \in \{2, \ldots, 2r\}$. ■

For $p = 13$ we obtain six continuant representations

$$13 = [13] = [6, 2] = [4, 3] = [3, 4] = [2, 6]$$

and

$$13 = [2, 1, 1, 2] = \begin{vmatrix} 2 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & -1 & 2 \end{vmatrix}.$$

8. PROOF OF THEOREM 2. We begin with

Lemma 3. *Given any $n \in \mathbb{N}$ with $n \geq 2$ and any set of positive integers $\{q_1, q_2, \ldots, q_n\}$ define*

$$\begin{aligned} I_n(q_1, q_2, \ldots, q_n) := {} & [q_1, q_2, \ldots, q_n][q_2, q_3, \ldots, q_{n-1}] \\ & - [q_1, q_2, \ldots, q_{n-1}][q_2, \ldots, q_n]. \end{aligned} \tag{8.1}$$

Then

$$I_n(q_1, q_2, \ldots, q_n) = (-1)^n. \tag{8.2}$$

Proof: We have, from (4.2),

$$I_2(q_1, q_2) = [q_1, q_2] - [q_1][q_2] = 1. \tag{8.3}$$

For the general case we have, from property 6 of Lemma 1,

$$[q_1, \ldots, q_n] = [q_1, \ldots, q_{n-1}]q_n + [q_1, \ldots, q_{n-2}] \tag{8.4}$$

$$[q_2, \ldots, q_n] = [q_2, \ldots, q_{n-1}]q_n + [q_2, \ldots, q_{n-2}]. \tag{8.5}$$

Multiply (8.4) by $[q_2, \ldots, q_{n-1}]$ and (8.5) by $[q_1, \ldots, q_{n-1}]$ to give, using (8.1),

$$\begin{aligned} I_n(q_1, q_2, \ldots, q_n) &= [q_1, \ldots, q_{n-2}][q_2, \ldots, q_{n-1}] - [q_2, \ldots, q_{n-2}][q_1, \ldots, q_{n-1}] \\ &= -I_{n-1}(q_1, q_2, \ldots, q_{n-1}). \end{aligned}$$

Repeated application of this last result yields

$$I_n(q_1, q_2, \ldots, q_n) = (-1)^r I_{n-r}(q_1, \ldots, q_{n-r}) \quad (r \in \{1, 2, \ldots, n-2\});$$

taking $r = n - 2$ and using (8.3) gives $I_n(q_1, q_2, \ldots, q_n) = (-1)^{n-2} I_2 = (-1)^n$, so (8.2) follows, as required. ■

Proof of Theorem 2, Part 1. Let $p \in \mathbb{P}$ with $p \equiv 1 \pmod 4$. Then the proof of Theorem 1 ensures that p has at least one palindromic continuant representation

$$p = [q_1, \ldots, q_s, q_s, \ldots, q_1] \tag{8.6}$$

with $s \geq 1$ and $q_1 \geq 2$.

Define $x_0 \in \mathbb{N}$ by

$$x_0 := [q_2, \ldots, q_s, q_s, \ldots, q_1]; \tag{8.7}$$

it follows that, from $q_1 \geq 2$ and property 1 of Lemma 1,

$$1 < x_0 < (p-1)/2. \tag{8.8}$$

Now apply the result of Lemma 3 to the right-hand side of (8.6), with $n = 2s$, to obtain

$$[q_1, \ldots, q_s, q_s, \ldots, q_1][q_2, \ldots, q_s, q_s, \ldots, q_2]$$
$$-[q_1, \ldots, q_s, q_s, \ldots, q_2][q_2, \ldots, q_s, q_s, \ldots, q_1] = (-1)^{2s} = 1.$$

From (8.6), (8.7), and property 3 of Lemma 1 the preceding result reduces to

$$p[q_2, \ldots, q_s, q_s, \ldots, q_2] - x_0^2 = 1 \tag{8.9}$$

and so

$$x_0^2 \equiv -1 \pmod{p}. \tag{8.10}$$

This result, together with (8.8) completes the proof of Part 1. ■

Remarks

1. We note that (8.9) was known to Henry Smith and is stated at the end of [**20**].
2. We note also, from the proof of this part of Theorem 2, that the number

$$\lambda = x_0 \tag{8.11}$$

is a member of the set $\{2, 3, \ldots, r\}$ and, for this choice of λ, the quotient p/λ yields the palindromic continuant representation of p; see (6.5).

Examples

1. For the case when the prime $p = 13$ we have the following explicit results

$$p = [q_1, q_2, \ldots, q_s, q_s, \ldots, q_2, q_1] = [2, 1, 1, 2]$$
$$x_0 = [q_2, \ldots, q_s, q_s, \ldots, q_2, q_1] = [1, 1, 2] = 5$$
$$[q_2, \ldots, q_s, q_s, \ldots, q_2] = [1, 1] = 2.$$

The general result $1 < x_0 < (p-1)/2$ becomes $1 < 5 < 6$ in this case. We can now confirm the results (8.9) and (8.10):

$$p[q_2, \ldots, q_s, q_s, \ldots, q_2] - x_0^2 = 13 \cdot 2 - 5^2 = 1$$

and

$$x_0^2 = 25 = 26 - 1 \equiv -1 \pmod{13}.$$

2. Let $p = 1913$; then $p \equiv 1 \pmod{4}$ and, see (6.5) and (8.11), $\lambda = x_0 = 712$ since $712^2 = 506944 = 265 \times 1913 - 1 \equiv -1 \pmod{1913}$, with $1 < 712 < 956 = (1913 - 1)/2$. The Euclidean algorithm yields

$$\frac{p}{\lambda} = \frac{1913}{712} = \frac{[2, 1, 2, 5, 5, 2, 1, 2]}{[1, 2, 5, 5, 2, 1, 2]}$$

and since $[2, 1, 2, 5] = 43$, $[2, 1, 2] = 8$ we have $43^2 + 8^2 = 1849 + 64 = 1913$.

3. For $p = 969433$, with $p \equiv 1 \pmod{4}$ the appropriate palindromic continuant is $p = [2, 3, 4, 5, 6, 6, 5, 4, 3, 2]$ with $[2, 3, 4, 5, 6] = 972$ and $[2, 3, 4, 5] = 157$; thus $969433 = 972^2 + 157^2$.

4. For the example (1.1) produced by the Brillhart method, i.e. $p = 10^{50} + 577$, it may be shown that

$$\lambda = x_0 = 24574739597286316058804545812463447369459349571921$$

and the relevant palindromic continuant is

$$\begin{aligned} p = [&4, 14, 2, 4, 4, 1, 5, 1, 3, 4, 2, 17, 1, 1, 1, 3, 1, 2, 1, 7, 1, 4, 1, 1, 2, 7, 11, 1, \\ &1, 3, 3, 3, 1, 14, 1, 9, 1, 2, 1, 1, 1, 1, 22, 1, 1, 1, 1, 3, 21, 1, 1, 3, 3, 1, 5, 1, \\ &1, 5, 1, 3, 3, 1, 1, 21, 3, 1, 1, 1, 1, 22, 1, 1, 1, 1, 2, 1, 9, 1, 14, 1, 3, 3, 3, 1, \\ &1, 11, 7, 2, 1, 1, 4, 1, 7, 1, 2, 1, 3, 1, 1, 1, 17, 2, 4, 3, 1, 5, 1, 4, 4, 2, 14, 4], \end{aligned}$$

which has length 112. Thus we have, for the two squares representation,

$$\begin{aligned} u = [&4, 14, 2, 4, 4, 1, 5, 1, 3, 4, 2, 17, 1, 1, 1, 3, 1, 2, 1, 7, 1, 4, 1, 1, 2, 7, 11, 1, \\ &1, 3, 3, 3, 1, 14, 1, 9, 1, 2, 1, 1, 1, 1, 22, 1, 1, 1, 1, 3, 21, 1, 1, 3, 3, 1, 5, 1] \\ = &\ 761106534380835424545040 1 \end{aligned}$$

and

$$\begin{aligned} v = [&4, 14, 2, 4, 4, 1, 5, 1, 3, 4, 2, 17, 1, 1, 1, 3, 1, 2, 1, 7, 1, 4, 1, 1, 2, 7, 11, 1, \\ &1, 3, 3, 3, 1, 14, 1, 9, 1, 2, 1, 1, 1, 1, 22, 1, 1, 1, 1, 3, 21, 1, 1, 3, 3, 1, 5] \\ = &\ 648626890687392164224542 4 \end{aligned}$$

with $p = u^2 + v^2$, as in (1.1).

Proof of Theorem 2, Part 2. Suppose that r is another solution of the quadratic equation (2.1) with $r \neq x_0$ and $r \neq p - x_0$; without loss of generality we may suppose that r is a least, positive residue (mod p). Then $r^2 \equiv x_0^2 \equiv -1 \pmod{p}$ and hence p divides $r^2 - x_0^2 = (r - x_0)(r + x_0)$; since p is prime it divides $r - x_0$ or $r + x_0$. The former case implies $r \equiv x_0 \pmod{p}$, but since both r and x_0 are least, positive residues it follows that $r = x_0$. In the latter case $r \equiv -x_0 \equiv p - x_0 \pmod{p}$ and since r and $p - x_0$ are least, positive residues it follows that $r = p - x_0$. This contradiction completes the proof of Part 2. ■

9. THE "SMITH" PROOF OF THE GAUSS THEOREM. We are now in a position to give a proof, using the methods of Henry Smith, of the Gauss uniqueness result for the Fermat theorem, as presented in Theorem 1.

Proof: Let $p \in \mathbb{P}$ with $p \equiv 1 \pmod{4}$ and suppose that there are two, co-prime two squares representations: $p = u^2 + v^2$ and $p = s^2 + r^2$, with $u < v$, $s < r$.

Apply Algorithm 1 to the rational numbers v/u and r/s to obtain

$$1 < \frac{v}{u} = \frac{[q_1, q_2, \ldots, q_n]}{[q_2, \ldots, q_n]} \quad \text{and} \quad 1 < \frac{r}{s} = \frac{[t_1, t_2, \ldots, t_m]}{[t_2, \ldots, t_m]};$$

then

$$\begin{aligned} u &= [q_2, \ldots, q_n] \quad \text{and} \quad v = [q_1, q_2, \ldots, q_n] \\ s &= [t_2, \ldots, t_m] \quad \text{and} \quad r = [t_1, t_2, \ldots, t_m]. \end{aligned}$$

Hence, property 6 of Lemma 1 ensures that

$$\begin{aligned} p = u^2 + v^2 &= [q_2, \ldots, q_n]^2 + [q_1, q_2, \ldots, q_n]^2 \\ &= [q_n, \ldots, q_2, q_1, q_1, q_2, \ldots, q_n] \end{aligned} \tag{9.1}$$

and

$$p = s^2 + r^2 = [t_2, \ldots, t_m]^2 + [t_1, t_2, \ldots, t_m]^2$$
$$= [t_m, \ldots, t_2, t_1, t_1, t_2, \ldots, t_m]. \tag{9.2}$$

Part 1 of Theorem 2 guarantees that the continuants $[q_{n-1}, \ldots, q_1, q_1, \ldots, q_{n-1}, q_n]$ and $[t_{m-1}, \ldots, t_1, t_1, \ldots, t_{m-1}, t_m]$ are both solutions of the quadratic equation $x^2 \equiv -1 \pmod{p}$ and satisfy $1 < x < (p-1)/2$. From the uniqueness of this solution we have

$$[q_{n-1}, \ldots, q_1, q_1, \ldots, q_{n-1}, q_n] = [t_{m-1}, \ldots, t_1, t_1, \ldots, t_{m-1}, t_m] = \rho \text{ (say)}. \tag{9.3}$$

From (9.1), (9.2), and (9.3) it follows that

$$1 < \frac{p}{\rho} = \frac{[q_n, \ldots, q_2, q_1, q_1, q_2, \ldots, q_n]}{[q_{n-1}, \ldots, q_1, q_1, \ldots, q_{n-1}, q_n]} = \frac{[t_m, \ldots, t_2, t_1, t_1, t_2, \ldots, t_m]}{[t_{m-1}, \ldots, t_1, t_1, \ldots, t_{m-1}, t_m]}. \tag{9.4}$$

Applying Algorithm 1 to both the continuant terms in (9.4) shows that $m = n$ and $q_i = t_i$ $(i = 1, 2, \cdots, n)$; thus $u = s, v = t$ and the uniqueness result is established. ■

10. APPENDIX. In this appendix we reproduce the original 1855 paper [**20**] of Henry Smith.

DE COMPOSITIONE NUMERORUM PRIMORUM FORMAE $4\lambda + 1$ EX DUOBUS QUADRATIS

Sit

$$q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cfrac{}{\ddots \; + \cfrac{1}{q_n}}}}$$

fractio continua, cujus numerator, qui determinanti

$$\begin{vmatrix} q_1 & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ -1 & q_2 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & -1 & q_3 & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & -1 & q_4 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & -1 & q_n \end{vmatrix}$$

aequalis est, per hujusmodi formulam $[q_1\, q_2\, q_3 \ldots q_{n-1} q_n]$ exprimatur. Erit ergo

$$[q_1\, q_2 \ldots q_{i-1} q_i] = [q_i\, q_{i-1} \ldots q_2\, q_1]$$

et

$$[q_1 \ldots q_n] = [q_1\, q_2 \ldots q_i] \cdot [q_{i+1} \ldots q_n] + [q_1\, q_2 \ldots q_{i-1}] \cdot [q_{i+2} \ldots q_n];$$

quae aequationes pendent ab illa forma determinantali, ambae autem L. Eulero debentur.

Itaque, si quantitatum q *par* sumatur numerus, ipsaeque ita serie symmetrica disponantur, ut binae inter se aequales fiant, elucet, quantitatem $[q_1\, q_2 \ldots q_i\, q_i \ldots q_2\, q_1]$ summam fore duorum quadratorum inter se primorum; fit enim

$$[q_1\, q_2 \ldots q_i\, q_i \ldots q_2\, q_1] = [q_1\, q_2 \ldots q_i]^2 + [q_1\, q_2 \ldots q_{i-1}]^2.$$

Contra in numero quotientium *impari*, erit

$$[q_1 \ldots q_{i-1}\, q_i\, q_{i-1} \ldots q_2\, q_1] = [q_1 \ldots q_{i-1}] \cdot \{[q_1 \ldots q_i] + [q_1 \ldots q_{i-2}]\},$$

unde colligis, numerum $[q_1 \ldots q_i \ldots q_1]$ primum esse non posse, nec duplicem numeri primi; si quidem casus excipis, in quibus, aut i unitati aequatur, aut i binario, q_1 unitati.

Sit p numerus integer datus; $\mu_1, \mu_2, \ldots \mu_s$ series numerorum, qui ad p primi sunt, ipsiusque p dimidio minores.

Formentur fractiones continuae $\frac{p}{\mu_1}, \frac{p}{\mu_2}, \ldots, \frac{p}{\mu_s}$; quae omnes ita terminentur, ut is quotiens qui in extremo loco ponatur, unitatem superet. Hinc patet, quanta fuerit numerorum $\mu_1, \mu_2, \ldots \mu_s$ multitudo, tantum fore numerum determinantium $[q_1 \ldots q_n]$, qui dato numero p aequales erunt, neque praeter illos ullum dare ejusdem formae determinantem, cujus et primus et extremus quotiens unitate major sit, quique numero p aequalis esse possit.

Jam vero, quum duo determinantes $[q_1 \ldots q_n]$ et $[q_n \ldots q_1]$ aequales sint, quumque ipsum q_n unitate majus sit, apparet $[q_n \ldots q_1]$ ex una aliqua fractionum $\frac{p}{\mu}$ oriri. Unde sequitur, data quavis fractione $\frac{p}{\mu}$, inveniri posse aliam in eadem serie, quae quotientes eosdem, ordine inverso, repraesentet.

Sit p primus, formae $4\lambda + 1$; ut numerus determinantium ipsi p aequalium par existat. Quum ipse p unus e determinantium serie fiat, unus certo alius inveniri poterit in quo quotientium ordo invertendo non mutatur. Cum sit ergo

$$p = [q_1\, q_2 \ldots q_i\, q_i \ldots q_2\, q_1]$$

erit denique

$$p = [q_1\, q_2 \ldots q_i]^2 + [q_1\, q_2 \ldots q_{i-1}]^2.$$

Quam theorematis Fermatiani demonstrationem maxime elementarem esse patet, quum pendeat a conversione fractionum vulgarium in fractiones continuas.

Singulos autem formae $1 + x^2$ divisores ex duobus quadratis conflari, eodem modo demonstrare in promptu est. Sit enim

$$\mu\nu = 1 + x^2,$$

apparet fore

$$\mu = [q_1\, q_2 \ldots q_i\, q_i \ldots q_2\, q_1]$$

$$\nu = [q_2\, q_3 \ldots q_i\, q_i \ldots q_3\, q_2]$$

$$x = [q_1\, q_2 \ldots q_i\, q_i \ldots q_2].$$

Oxford, *Maio* 1854.

ACKNOWLEDGMENTS. Norrie Everitt thanks his three co-authors for their agreement to dedicate this paper to Paul Halmos who, from afar, has been his guide and mentor in mathematics. This paper should have been completed some years ago for a volume dedicated to Paul Halmos; apologies for the delay but I hope the paper is now the better for subsequent collaboration and extension.

All four authors thank Keith Hannabuss, Fellow and Tutor in Mathematics of Balliol College in Oxford, for his contribution to the Section on the life of Henry Smith; we have been guided by and quoted from his papers [**10**] and [**11**]; additionally we have had access to, and quoted from a yet unpublished account of the life of Henry Smith.

REFERENCES

1. C. W. Barnes, The representation of primes of the form $4n + 1$ as the sum of two squares, *Enseign. Math.* (2) **18** (1972) 289–299.
2. E. T. Bell, *Men of Mathematics*, Victor Gollancz Ltd., London, 1937.
3. J. Brillhart, Note on representing a prime as a sum of two squares, *Math. Comp.* **26** (1972), 1011–1013.
4. D. M. Burton, *Elementary Number Theory*, The McGraw- Hill Companies, Inc., New York, 1998.
5. G. E. Chrystal, *Algebra*: **I** and **II**, Adam and Charles Black, Edinburgh, 1889. The 6th. edition reprinted by Chelsea Publishing Co., New York, 1959.
6. H. A. Davenport, *The Higher Arithmetic*, Hutchinson House, London, 1952.
7. L. E. Dickson, *History of the Theory of Numbers* **11**, Chelsea Publishing Co., New York, 1966.
8. U. Dudley, *Elementary Number Theory*, 2nd. edition, W. H. Freeman and Company, New York, 1978.
9. J. A. Ewell, A simple proof of Fermat's two-square theorem, *Amer. Math. Monthly* **90** (1983) 635–637.
10. K. Hannabuss, The mathematician the world forgot, *New Scientist* **97** (1983) 901–903.
11. K. Hannabuss, Forgotten fractals, *Math. Intelligencer* **18** (1996) 28–31.
12. G. H. Hardy, *A Mathematician's Apology*, Cambridge University Press, 1969.
13. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, 1952.
14. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th edition, Oxford University Press, 1979.
15. T. Hawkins, *Lebesgue's Theory of Integration; Its Origins and Development*, Chelsea Publishing Co., New York, 1975.
16. D. R. Heath-Brown, Fermat's two-squares theorem, *Invariant.* (1984) 3–5.
17. C. Hermite, Note au sujet de l'article précedent, *J. Math. Pures Appl.* **13** (1848) 15.
18. T. Muir and W. H. Metzler, *A Treatise on the Theory of Determinants*, Dover Publications Inc., New York, 1960.
19. J.-A. Serret, Sur un théorème relatif aux nombres entiers, *J. Math. Pures Appl.* **13** (1848) 12–14.
20. H. J. S. Smith, De Compositione Numerorum Primorum $4\lambda + 1$ Ex Duobus Quadratis, *Crelle's Journal* **L** (1855) 91–92.
21. H. J. S. Smith, *The Collected Mathematical Papers of Henry John Stephen Smith*: **I** and **II**, (Edited by J.W.L. Glaisher), The Clarendon Press, Oxford, 1894. Reprinted by Chelsea Publishing Co., New York, 1965.
22. S. Wagon, The Euclidean algorithm strikes again, *Amer. Math. Monthly* **97** (1990) 125–129.
23. D. Zagier, A one-sentence proof that every prime $p \equiv 1 \pmod 4$ is a sum of two squares, *Amer. Math. Monthly* **197** (1990) 144.

FRANCIS CLARKE was educated at Birmingham University (B.Sc. 1968) and at the University of Warwick (M.Sc. 1969, Ph.D. 1971). Since 1971 he has taught in the Mathematics Department at the University of Wales Swansea. He has research interests both in algebraic topology and in number theory. The Bernoulli numbers are an enduring fascination that provide a link between the two fields. Other interests include playing the clarinet, hill walking, skiing, and sailing.
University of Wales Swansea, Singleton Park, Swansea SA2 8PP, Wales, UK.
f.clarke@swansea.ac.uk

W. NORRIE EVERITT was educated at the University of Birmingham, and then at Balliol College, Oxford in Great Britain. He obtained his D.Phil. from the University of Oxford in 1955 with thesis advisor Edward Charles Titchmarsh. He first heard of the Smith proof of the Fermat two squares theorem at a lecture given by the late Harold Davenport, in 1950, to the Invariant Society, University of Oxford. He is Professor Emeritus of the University of Birmingham.
School of Mathematics and Statistics, University of Birmingham, Edgbaston, Birmingham B15 2TT, England, UK
w.n.everitt@bham.ac.uk

LANCE L. LITTLEJOHN received his B.Sc. in 1975, his M.A. in 1976 (both in mathematics from the University of Western Ontario), and his Ph.D. in mathematics from Penn State University in 1981. His first academic position was at the University of Texas at San Antonio; since 1983, he has been at Utah State University. He enjoys all types of rigorous mathematics; in particular, his research interests in differential equations, operator theory, and special functions. Most of his non-mathematical time is spent with his wife, Wendy, and their two children, Alex and Mary (although he does manage to sneak some time to follow his beloved, but hapless, Detroit Tigers).
Department of Mathematics and Statistics, Utah State University, Logan, UT 84322-3900, USA
lance@math.usu.edu

ROELOF VORSTER received a B.Sc. from the University of Pretoria, South Africa in 1962 and started his career with a big bang as an explosives chemist at the world's largest dynamite factory. Realizing that mathematics is his first love, he joined the staff of the distance teaching University of South Africa (UNISA) where he obtained a Ph.D. in topology and category theory in 1972, and where he has been head of the mathematics department since 1990. It was during a visit to UNISA by Norrie Everitt and Lance Littlejohn in 1990 that the former delivered a fascinating talk that aroused great interest in the work of HJS Smith and the Fermat Two Squares Theorem.
Department of Mathematics, Applied Mathematics and Astronomy, University of South Africa, P.O. Box 392, 0001 Pretoria, Republic of South Africa
vorstsjr@alpha.unisa.ac.za