NNTDM 5 (1999) 2, 71-79

SPECULATION ON FERMAT'S PROOF OF HIS LAST THEOREM

J.V. LEYENDEKKERS

The University of Sydney, 2006, Australia

A.G. SHANNON

University of Technology, Sydney, 2007,
& KvB Institute of Technology, North Sydney, 2060, Australia

Abstract

It has been suspected that if Fermat did indeed have a simple proof for his famous 'last theorem', that he probably employed his method of infinite descent. In a renewed attempt to see how Fermat might have thought that he had proved that if $c^n = a^n + b^n$, $a, b, c, n \in \mathbf{Z}$, $n > 2$, then a, b, c cannot all be integers, we set $c = a + b + m$, $m \in \mathbf{Z}$, and then raised it to the nth power. The roots of the resulting polynomial in m appear to be $-(a+b)$ only when $n \neq 1, 2$, and the result might have seemed to Fermat to follow from this. The plausibility of the algebra developed here is considered in the context of the work of the sixteenth century mathematicians, particularly Cardano and Bombelli.

AMS Classification Numbers: 01A45, 11D41

1. Introduction

Now that Andrew Wiles and Richard Taylor have proved Fermat's Last Theorem, but with methods quite undreamt of in the seventeenth century (cf. van der Poorten, 1996), one might legitimately suggest a hypothetical search for Fermat's lost proof, if indeed a valid one ever existed. Velleman (1997) in fact goes further by questioning whether a correct proof guarantees that the theorem is true: he asks, for instance, whether we should be convinced by Wiles' proof that no counter-example will ever be found, or merely that if the Zermelo-Fraenkel set of axioms is consistent then no counter-example will be found.

This is not an attempt to prove the theorem, but to renew speculation about Fermat's possible attempts of which modern commentators have possibly erred on the side of harshness in their judgement of him. The historical approach adopted here is, as yet, relatively uncommon among historians of mathematics, though it has been used by some professional historians over the last quarter century. The motivation ascribed by Riemer (1998) to the historian Jill Ker Conway, for instance, was that she found the traditional approach in history "far too dry, far too unimaginative, scornful of the intellect, chained to the study of historical 'fact', unable to break free of the tyrrany of documents and data into a more speculative world".

Pierre de Fermat (1601-1665) was a lawyer by profession and served in the Parlement in Toulouse. He was an 'amateur' mathematician in the best sense of that term if we reflect that it comes from the Latin *amare* (to love). "Fermat published almost nothing but instead made known his results in letters to a French priest, Marin Mersenne, who then passed them on to others. Fermat's own edition of Diophantus' text was published posthumously in Toulouse in 1670. His own works did not appear until 1679 in his *Varia Opera Mathematica*" (Hillman and Alexanderson, 1978).

Fermat's failure to publish in the modern sense has meant that he has not received due recognition at times for his achievements, which include the discovery of the basic idea of analytic geometry at least a year before the publication of Descartes' *La Geométrie*. Furthermore, Laplace (1749-1827) described Fermat as "the true inventor of the differential calculus" since in the 1630s he invented methods for calculating maxima and minima for certain curves, and his treatment of the 'generalized parabola', $y = x^n$, is operationally identical with the procedure of modern differentiation.

It seems possible that Fermat, in considering his 'last' theorem, could have begun with the following polynomial, although not with the terminology or notation used here of course.

2. Polynomial in m

Suppose for $a, b, c, m \in \mathbf{Z}$ that

$$c = a + b + m.$$

Then

$$c^n = ((a+b) + m)^n,$$

which can be expanded:

$$c^n - a^n - b^n = P(m)$$

where $P(m)$ is a polynomial given by

$$P(m) = \sum_{j=0}^{n-1} \binom{n}{j} (a+b)^j m^{n-j} + \sum_{i=1}^{n-1} \binom{n}{i} a^{n-i} b^i.$$

If $c^n = a^n + b^n$, $n > 2$, then we have a polynomial equation $P(m) = 0$ with roots $\alpha_1, \alpha_2, \ldots, \alpha_n$, so that since $n > 2$

$$-n(a+b) = \sum_{i=1}^{n} \alpha_i \tag{2.1}$$

and

$$\binom{n}{2}(a+b)^2 = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \alpha_i \alpha_j. \tag{2.2}$$

We now square both sides of (2.1):

$$n^2(a+b)^2 = \sum_{i=1}^{n} \alpha_i^2 + 2\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \alpha_i \alpha_j,$$

which we multiply through by $(n-1)/2$ (since $n \neq 1$) to obtain

Latin *amare* (to love). "Fermat published almost nothing but instead made known his results in letters to a French priest, Marin Mersenne, who then passed them on to others. Fermat's own edition of Diophantus' text was published posthumously in Toulouse in 1670. His own works did not appear until 1679 in his *Varia Opera Mathematica*" (Hillman and Alexanderson, 1978).

Fermat's failure to publish in the modern sense has meant that he has not received due recognition at times for his achievements, which include the discovery of the basic idea of analytic geometry at least a year before the publication of Descartes' *La Geométrie*. Furthermore, Laplace (1749-1827) described Fermat as "the true inventor of the differential calculus" since in the 1630s he invented methods for calculating maxima and minima for certain curves, and his treatment of the 'generalized parabola', $y = x^n$, is operationally identical with the procedure of mdoern differentiation.

2. Polynomial in m

Suppose for $a, b, c, m \in \mathbf{Z}$ that

$$c = a + b + m.$$

Then

$$c^n = ((a+b)+m)^n,$$

which can be expanded:

$$c^n - a^n - b^n = P(m)$$

where $P(m)$ is a polynomial given by

$$P(m) = \sum_{j=0}^{n-1} \binom{n}{j} (a+b)^j m^{n-j} + \sum_{i=1}^{n-1} \binom{n}{i} a^{n-i} b^i.$$

If $c^n = a^n + b^n$, $n > 2$, then we have a polynomial equation $P(m) = 0$ with roots $\alpha_1, \alpha_2, \ldots, \alpha_n$, so that since $n > 2$

$$-n(a+b) = \sum_{i=1}^{n} \alpha_i \tag{2.1}$$

and

$$\binom{n}{2} (a+b)^2 = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \alpha_i \alpha_j. \tag{2.2}$$

We now square both sides of (2.1):

$$n^2 (a+b)^2 = \sum_{i=1}^{n} \alpha_i^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \alpha_i \alpha_j,$$

which we multiply through by $(n-1)/2$ (since $n \neq 1$) to obtain

$$\tfrac{1}{2}n^2(n-1)(a+b)^2 = \tfrac{1}{2}(n-1)\sum_{i=1}^{n}\alpha_i^2 + (n-1)\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\alpha_i\alpha_j. \tag{2.3}$$

Multiply (2.2) through by n to yield

$$\tfrac{1}{2}n^2(n-1)(a+b)^2 = n\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\alpha_i\alpha_j. \tag{2.4}$$

We then equate the right hand sides of (2.3) and (2.4):

$$(n-1)\sum_{i=1}^{n}\alpha_i^2 = 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\alpha_i\alpha_j. \tag{2.5}$$

3. Speculation

We now speculate on how Fermat might have proved that $\alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_n$. Suppose on the contrary that

$$\alpha_1 = \alpha_n + \beta_1,\ \alpha_2 = \alpha_n + \beta_2, \alpha_3 = \alpha_n + \beta_3, \ldots, \alpha_{n-1} = \alpha_n + \beta_{n-1}.$$

Substitution of these values into (2.5) gives

$$(n-1)\sum_{i=1}^{n-1}\beta_i^2 = 2\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}\beta_i\beta_j. \tag{3.1}$$

We now re-write (3.1) in the form

$$\sum_{i=1}^{n-2}\sum_{j=i+1}^{n-1}(\beta_i-\beta_j)^2 = -\sum_{i=1}^{n-1}\beta_i^2. \tag{3.2}$$

Since the squared terms cannot be negative, all the β_i must be zero if we do not know of the existence of imaginary numbers, and so the $\alpha_1 = \alpha_2 = \ldots = \alpha_n = \alpha$ in (2.5). We now substitute into (2.1) and find that

$$-n(a+b) = n\alpha = nm,$$

and thus

$$m = -(a+b),$$

which in (1.1) means that $c = 0$. In the case of $n = 3$, the foregoing is particularly plausible since (2.5) reduces to

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3.$$

In considering

$$\alpha_1\alpha_2\alpha_3 = -(a+b)^3 = -3ab(a+b)$$

it is clear that a and b cannot both be integers.

4. Complex Roots

Given Fermat's strong interest in the theory of equations and the availability of the books by Bombelli and Cardano on algebra, it is conceivable that he might have taken the above argument further; (see Section 6). When $n = 3$, Equations (2.1) and (2.2) become in turn

$$-3(a+b) = 3\alpha + \beta_i + \beta_j \tag{4.1}$$

and

$$3(a+b)^2 = 3\alpha^2 + 2\alpha(\beta_i + \beta_j) + \beta_i\beta_j \tag{4.2}$$

in which α is any root and the subscripts refer to the other two roots. The product of the three roots is given by

$$-3ab(a+b) = \alpha^3 + \alpha^2(\beta_i + \beta_j) + \alpha\beta_i\beta_j. \tag{4.3}$$

On eliminating the term in α^2 we get

$$\alpha^3 - \alpha(\beta_i\beta_j + 3(a+b)^2) - 6ab(a+b) = 0 \tag{4.4}$$

which has the form

$$x^3 + px + q = 0 \tag{4.5}$$

where

$$p = -(\beta_i\beta_j + 3(a+b)^2) \tag{4.6}$$

and

$$q = -6ab(a+b). \tag{4.7}$$

Cardano's solution for such a cubic shows that if $27q^2 + 4p^3 < 0$, then the roots rmust always be real (Hollingdale, 1989). Now in the Fermat case

$$27q^2 = 27 \times 36a^2b^2(a+b)^2$$

and

$$4p^3 = -4(\beta_i\beta_j + 3(a+b)^2)^3,$$

so that if

$$27q^2 + 4p^3 \geq 0,$$

then

$$27 \times 9a^2b^2(a+b)^2 \geq (\beta_i\beta_j + 3(a+b)^2)^3. \tag{4.8}$$

If the βs are zero, then $9a^2b^2 > (a+b)^4$ which cannot be true. Hence $(27q^2 + 4p^3)$must be negative and all roots are real, which is consistent with the above results.

If β_i and β_j are a conjugate pair, then the product is the sum of squares and positive, so that the right hand side of (4.8) increases; hence $27q^2 + 4p^3 < 0$, so that all the roots must be real. When the roots are real, they must all be equal as demonstrated in Section 3, and so a, b, c cannot all be integers. From (2.1) the roots all equal $-(a+b)$. As noted above this implies that

$$-3ab(a+b) = -(a+b)^3$$

so that

$$a^2 + b^2 = ab$$

and a and b cannot both be non-zero integers.

Cardano published his stately book on algebra, *Ars magna*, in 1545. Tartaglia's methods for solving cubics were in this book and were the source of a dispute from which emerged some interesting documents. Among these were the *Quaesiti* of Tartaglia in 1546 and the *Cartelli* of Ferrari in 1547-48, from which the whole history of the spectacular discovery concerning the cubic equation $x^3 + px + q = 0$ became public knowledge (Struik, 1965).

The solution is now known as the Cardano solution and has the form

$$x = ((p^3/27 + q^2/4)^{\frac{1}{2}} + q/2)^{1/3} - ((p^3/27 + q^2/4)^{\frac{1}{2}} - q/2)^{1/3},$$

as utilized in this paper. We observe that this solution introduced quantities of the form $(a + b^{\frac{1}{2}})^{1/3}$which are different from the Euclidean $(a + b^{\frac{1}{2}})^{\frac{1}{2}}$.

5. General Case

Finally, for the general case, $n > 3$, we equate the coefficients of m^{n-2} and m^{n-3}in $P(m)$ to the products of the roots, so that with α any root and $\beta \neq 0$,

$$\tfrac{1}{2}n(n-1)(a+b)^2 = \tfrac{1}{2}n(n-1)\alpha^2 + (n-1)A\alpha + B, \tag{5.1}$$

and

$$\begin{aligned} -n(n-1)(n-2)(a+b)^3/6 &= n(n-1)(n-2)\alpha^3/6 \\ &+ \tfrac{1}{2}(n-1)(n-2)A\alpha^2 + (n-2)B\alpha + C, \end{aligned} \tag{5.2}$$

where

$$\begin{aligned} A &= \sum_{i=1}^{n-1} \beta_i, \\ B &= \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \beta_i\beta_j = \tfrac{1}{2}(n-1)\sum_{i=1}^{n-1}\beta_i^2, \\ C &= \sum_{i=1}^{n-3} \sum_{j=i+1}^{n-2} \sum_{k=j+1}^{n-1} \beta_i\beta_j\beta_k. \end{aligned}$$

On multiplying (5.1) by $\frac{1}{2}\alpha(n-2)$ and eliminating the terms in α^2 as before to get the Cardano form of the equation we find

$$p = -(3(a+b)^2 + 6(n-2)B/n(n-1)(n-2)) \tag{5.3}$$

and

$$q = -(2(a+b)^3 + 12C/n(n-1)(n-2)). \tag{5.4}$$

In order to determine the sign of $(27q^2 + 4p^3)$ more simply, we now find the relationship among A, B, C.

In the same manner as Equation (2.5) was obtained we have

$$\sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} \alpha_i \alpha_j \alpha_k = ((n-2)/6) \sum_{i=1}^{n} \alpha_i^2 (\sum_{j \neq i}^{n} \alpha_j).$$

Substitution of $\alpha_i = \alpha + \beta_i$, $\alpha_j = \alpha + \beta_j$, ..., and the use of Equation (3.1) yields

$$nC = BA$$

and

$$A^2 = \sum_{i=1}^{n-1} \beta_i^2 + 2 \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \beta_i \beta_j.$$

On using (3.1) again we obtain

$$B = (n-1)A^2/2n$$

and

$$C = (n-1)A^3/2n^2.$$

Replacing B and C in terms of A in Equations (5.3) and (5.4) we find that

$$p = -(3(a+b)^2 + 3(A/n)^2),$$

and

$$q = -(2(a+b)^3 + (6/(n-2))(A/n)^3).$$

Thus

$$27q^2 + 4p^3 = 4 \times 27(9/(n-2)^2 - 1)(A/n)^6 +$$
$$27 \times 4(a+b)^2(A/n)^2((6(a+b)(A/n)/(n-2)) - 3(a+b)^2 - 3(A/n)^2),$$

or

$$27 \times 4(a+b)^2(A/n)^2(-3(a+b-A/n)^2 - 6(n-3)(a+b)(A/n)/(n-2)).$$

The second term is negative no matter what the parity of A. The first term is less than the second term when $n = 4$, and the first term is zero when $n = 5$ and negative for $n > 5$. This means that with this line of argument the roots will all be real and hence equal, so that we have a contradiction if a and b are assumed to be integers. This would mean that Fermat had proved his theorem.

The consistency of the foregoing can be assessed as follows. From Cardano's equation for $q \in \mathbf{Z}$:

$$a = q + i$$

and

$$b = q - i,$$

so that

$$ab = q^2 + 1$$

and

$$a + b = 2q.$$

n	3	4
$Last\,term\,of\,P(m)$	$3ab(a+b)$	$2ab(2(a+b)^2 - ab)$
q	$\surd 3$	$(6q^2 - 1)^{1/4}$
$a^n + b^n$	$2q^3 - 6q$	$2q^4 - 12q^2 + 2$
	$= 2\surd 3(3 - 3)$	$= (12q^2 - 2) - 12q^2 + 2$

Geronimo Cardano (1501-1576) and Rafael Bombelli (1526-1573) had started to develop what we would now call imaginary numbers and so Fermat (1601-1665) may have been conceptually ready for them.

6. Discussion

In order to accept that Fermat might well have reasoned as we have speculated above, one needs to understand the 'mathematical climate' of Fermat's times. According to Struik (1965), Cardano's *Ars magna* contained another brilliant discovery, namely negative numbers which Cardano called "fictitious". He was, however, unable to do anything with the so-called "irreducible case" of the cubic equation in which there are three real solutions which appear as the sum or difference of what we now call imaginary numbers.

This difficulty was solved by the last of the great sixteenth century Bolognese mathematicians, Raffael Bombelli, whose *Algebra* appeared in 1572. In this book, and in a geometry written about 1550 which remained in manuscript, he introduced a consistent theory of imaginary complex numbers. He wrote $3i$ as $\surd(0 - 9)$ (literally: $R[0\,m, 9], R$ for *radix*, m for *meno*). This allowed Bombelli to solve the irreducible case by showing, for instance, that

$$(52 + (0 - 2209)^{\frac{1}{2}})^{1/3} = 4 + (0 - 1)^{\frac{1}{2}}.$$

Bombelli's book was widely read; Leibniz selected it for the study of cubic equations, and Euler quotes Bombelli in his own *Algebra* in the chapter on biquadratic equations.

It is a curious fact that the first introduction of imaginary numbers occurred in the theory of cubic equations, in the case where it was clear that real solutions exist though in an unrecognisable form, and not in the theory of quadratic equations, where our present textbooks introduce them. Perhaps it was the inherent challenge at the end of *Summa de arithmetica* (1494) by Luca Pacioli who asserted that the solution of the equations

$$x^3 + mx - n = 0 \text{ and } x^3 - mx + n = 0$$

was as impossible at the present state of science as the quadrature of the circle (Struik, 1965: 110). Clearly the study of cubic polynomials was important at the time, though Boyer notes that by the time of Leibniz (1646-1716), complex numbers were almost forgotten: he factored $x^4 + a^4$ into complex parts and he showed that $\surd 6 = \surd(1 + \surd - 3) + \surd(1 - \surd - 3)$, though he did not write the square roots of complex numbers in standard complex form. "The ambivalent status of complex numbers is well illustrated by the remark of Leibniz ... that imaginary numbers are a sort of amphibian, halfway between existence and nonexistence" (Boyer, 1968: 444). Complex numbers started to lose their 'supernatural' character, though full acceptance came only in the nineteenth century. We also observe that the notation of this paper is not chronological as we have merely tried to toy with a line of thinking. Notation as a tool of mathematical thought was to come later (Cajori, 1928), though the manipulation of polynomials in the way we have suggested had reached textbooks by the seventeenth century (cf. Cohen and Shannon, 1981).

By way of conclusion we observe that Fermat's method of infinite descent is often seen as the forerunner of the principle of mathematical induction. To what extent this is true, or whether he really used 'ascent' or 'descent' are debatable issues (cf. Hunter, 1964; van der Poorten; 1996, Stillwell 1998). When one compares (2.5) and (3.1) it is tempting to keep decreasing the upper limits of summation with some form of Fermat's method of infinite descent (Franklin and Daoud, 1988). In any case we are concerned here primarily with the validity of the logic. As Singh and Ribet (1997) said: "Either Fermat was mistaken, and his proof, if it existed, was flawed, or a simple and cunning proof awaits discovery". Or was Petsinis (1997) correct with his surmise that "sometimes I envisage Fermat peeping from behind his theorem, smiling mischievously in the knowledge that he has perpetrated a hoax that will tease and torment countless minds"?

References

Boyer Carl B. 1968. *A History of Mathematics*. Princeton: Princeton University Press.

Cajori F. 1928. *A History of Mathematical Notations*. Volumes 1,2. La Salle: Open Court.

Cohen G L & Shannon A G. 1981. John Ward's method for the calculation of pi. *Historia Mathematica*. **8**: 133-144.

Franklin, James & Daoud, Albert. 1988. *Introduction to Proofs in Mathematics*. Sydney: Prentice Hall.

Hillman, Abraham P. & Alexanderson Gerald P. 1978. *A First Undergraduate Course in Abstract Algebra*. Second Edition. Belmont, CA: Wadsworth.

Hollingdale, Stewart. 1989. *Makers of Mathematics*. London: Penguin.

Hunter, J. 1964. *Number Theory*. Edinburgh: Oliver and Boyd.

Petsinis, Tom. 1997. *The French Mathematician*. Ringwood, Vic.: Penguin.

Poorten, Alf van der. 1996. *Notes on Fermat's Last Theorem*. New York: John Wiley.

Riemer, Andrew. 1998. *Sandstone Gothic*. Sydney: Allen and Unwin.

Singh, Simon & Ribet, Kenneth A. 1997. Fermat's Last Stand. *Scientific American*. **277.5**: 36-41.

Stillwell, John. 1998. *Numbers and Geometry*. New York: Springer.

Struik, Dirk J. 1965. *A Concise History of Mathematics*. London: George Bell.

Velleman, Daniel J. 1997. Fermat's Last Theorem and Hilbert's Program. *Mathematical Intelligencer*. **19.1**: 64-67.