By analogy with an enzyme, the rate of turnover, or the turnover number of a carrier, may be defined by the following fictitious experiment. Consider a lipid membrane which separates two aqueous phases and which contains a fixed number, N, of carrier molecules. The solution on the left side contains ions of concentration c_M, which are transported by the carrier; the ion concentration in the solution on the right side is assumed to be zero. If the aqueous phases are electrically short-circuited, a carrier-mediated ion flux of magnitude Φ occurs through the membrane. At low ion concentration c_M, the flux increases linearly with c_M. At high concentrations, however, the carrier becomes gradually saturated, and Φ finally approaches a maximal value Φ_{max} (in a similar manner an enzyme-catalyzed reaction approaches a maximal rate in the limit of high substrate concentration). The turnover number f is then defined as the maximum number of ions which may be transported per second by a single carrier molecule:

$$f = \frac{\Phi_{max}}{N} \qquad (12)$$

This theory leads to the following simple expression for the turnover number:

$$f = \left(\frac{1}{k_S} + \frac{1}{k_{MS}} + \frac{2}{k_D}\right)^{-1} \qquad (13)$$

With the above values of the rate constants, one finds $f \simeq 10^4$ sec^{-1}, which means that a single valinomycin molecule is able to transport 10^4 K^+ ions per second across the membrane. This number is much higher than the turnover number of most enzymes.

Thus, we may state that valinomycin has a surprisingly low affinity for K^+ in the heterogeneous system membrane-water. In spite of this low affinity, valinomycin is an effective ion carrier; obviously, the reason is the high turnover number of the molecule.

References and Notes

1. W. J. V. Osterhout, *Ergeb. Physiol.* **35**, 967 (1933).
2. W. Wilbrandt and T. Rosenberg, *Pharmacol. Rev.* **13**, 109 (1961); R. Blumenthal and A. Katchalsky, *Biochim. Biophys. Acta* **173**, 357 (1969); W. R. Lieb and W. D. Stein, *ibid.* **265**, 169 (1972).
3. M. M. Shemyakin *et al.*, *J. Membrane Biol.* **1**, 402 (1969).
4. W. Simon and W. E. Morf, in *Membranes—A Series of Advances*, G. Eisenman, Ed. (Dekker, New York, 1972), vol. 2.
5. Z. Stefanac and W. Simon, *Chimia* **20**, 436 (1966); L. A. R. Pioda, H. A. Wachter, R. E. Dohner, W. Simon, *Helv. Chim. Acta* **50**, 1373 (1967).
6. E. Grell, T. Funck, F. Eggers, *Proceedings of the Symposium on Molecular Mechanisms of Antibiotic Action on Protein in Biosynthesis and Membranes, Granada 1971*, D. Vasquez, Ed. (Elsevier, Amsterdam, 1972).
7. B. C. Pressman, E. I. Harris, W. S. Jagger, I. H. Johnson, *Proc. Nat. Acad. Sci. U.S.A.* **58**, 1949 (1967).
8. G. Eisenman, S. Ciani, G. Szabo, *J. Membrane Biol.* **1**, 294 (1969).
9. M. Onishi and D. W. Urry, *Biochem. Biophys. Res. Commun.* **36**, 194 (1969); *Science* **168**, 1091 (1970); M. Pinkerton, L. K. Steinrauf, P. Dawkins, *Biochem. Biophys. Res. Commun.* **35**, 512 (1969).
10. C. Moore and B. C. Pressman, *Biochem. Biophys. Res. Commun.* **15**, 562 (1964).
11. B. P. Pressman, *Proc. Nat. Acad. Sci. U.S.A.* **53**, 1079 (1965); S. N. Graven, H. A. Lardy, D. Johnson, A. Rutter, *Biochemistry* **5**, 1729 (1966).
12. P. Mueller, D. O. Rudin, H. Ti Tien, W. C. Wescott, *Nature* **194**, 979 (1962); *Circulation* **26**, 1167 (1962); *J. Phys. Chem.* **67**, 534 (1963).
13. T. Hanai, D. A. Haydon, J. Taylor, *Proc. Roy. Soc. Ser. A* **281**, 377 (1964).
14. C. Huang and T. E. Thompson, *J. Mol. Biol.* **13**, 183 (1965).
15. P. Mueller and D. O. Rudin, *Biochem. Biophys. Res. Commun.* **26**, 398 (1967); A. A. Lev and E. P. Buzhinsky, *Tsitologiya* **9**, 102 (1967); T. E. Andreoli, M. Tiefenberg, D. C. Tosteson, *J. Gen. Physiol.* **50**, 2527 (1967).
16. G. Eisenman, S. Ciani, G. Szabo, *Fed. Proc.* **27**, 1289 (1968); E. A. Liberman and V. P. Topaly, *Biochim. Biophys. Acta* **163**, 125 (1968).
17. G. Stark and R. Benz, *J. Membrane Biol.* **5**, 133 (1971).
18. G. Benz, thesis, Univ. of Konstanz (1972).
19. H. Diebler, M. Eigen, G. Ilgenfritz, G. Maass, R. Winkler, *Pure Appl. Chem.* **20**, 93 (1969); W. E. Morf and W. Simon, *Helv. Chim. Acta* **54**, 794 (1971).
20. G. Eisenman, in *Ion Selective Electrodes*, R. A. Durst, Ed. (National Bureau of Standards Special Publication 314, Washington, D.C., 1969); S. Krasne and G. Eisenman, in *Membranes—A Series of Advances*, G. Eisenman, Ed. (Dekker, New York, 1972), vol. 2.
21. D. C. Tosteson, *Fed. Proc.* **27**, 1269 (1968); S. Ciani, G. Eisenman, G. Szabo, *J. Membrane Biol.* **1**, 1 (1969); G. Szabo, G. Eisenman, S. Ciani, *ibid.*, p. 346.
22. S. Krasne, G. Eisenman, G. Szabo, *Science* **174**, 412 (1971).
23. B. Ketterer, B. Neumcke, P. Läuger, *J. Membrane Biol.* **5**, 225 (1971).
24. B. Neumcke and P. Läuger, *Biophys. J.* **9**, 1160 (1969).
25. P. Läuger and G. Stark, *Biochim. Biophys. Acta* **211**, 458 (1970).
26. G. Stark, B. Ketterer, R. Benz, P. Läuger, *Biophys. J.* **11**, 981 (1971).
27. W. Lesslauer, J. Richter, P. Läuger, *Nature* **213**, 1224 (1967); S. G. A. McLaughlin, G. Szabo, G. Eisenman, S. Ciani, *Proc. Nat. Acad. Sci. U.S.A.* **67**, 1268 (1970); B. Neumcke, *Biophysik* **6**, 231 (1970).

28. I thank my collaborators R. Benz, B. Ketterer, B. Neumcke, and G. Stark, on whose work this article is largely based.

Fermat's Mathematics: Proofs and Conjectures

Fermat's working habits as a mathematician shed new light on the mystery of his famous "last theorem."

Michael S. Mahoney

One of the comforts of investigating the work of Pierre de Fermat (1601–1665), one of the very few comforts, lies in not having to explain to non-historians of science who he was. Anyone who has studied mathematics remembers, at least vaguely, the legend connected with the name Fermat, the famous "last theorem." "Wasn't he the man who didn't have room in the margin of some book for a proof he had, a proof which no one since has been able to find?" "Yes," runs the answer, "that's the one." To be more precise, in the margin of his copy of Claude Bachet de Méziriac's 1621 edition of Diophantus of Alexandria's *Arithmetica*, next to Proposition II,8 (To split a square into two squares), Fermat wrote (*1*, p. 291):

> But one cannot split a cube into two cubes, nor a quadratoquadrate [that is, fourth power] into two quadratoquadrates, nor in general any power *in infinitum* beyond the square into two like powers. I have uncovered a marvelous demonstration indeed of this, but the narrowness of the margin will not contain it.

That is, Fermat claimed he had found a proof of the theorem that the equation $x^n + y^n = z^n$ has no rational solution for integer n greater than 2, but did not have room to set it down. Apparently he did not set it down elsewhere, nor has anyone since been able to prove the theorem; not Euler, nor Gauss, nor

Kummer, to mention just a few who have tried (*2, 3*).

On hearing this account, the dramatically inclined listener may think of the kingdom lost for want of a horseshoe nail: For want of a wider margin, a proof was lost, and so on. Despite the legend's dramatic appeal, however, more is wrong with it than merely that Bachet's edition of Diophantus had extremely wide margins, wide enough for Fermat to insert at other points marginalia up to 17 times as long as the one just cited, or that a professional lawyer and prominent member of the Parlement of Toulouse was unlikely to run out of writing paper. As in the case of Galileo, the simplicity of the legend masks the complexity of the real, historical character. Fermat was no novice blessed with a flash of insight while dabbling in Diophantus. He was a well-read number theorist and mathematician who thought long, hard, and carefully about his work. If, to the detriment of posterity, he kept much of his mathematics in his head, he nonetheless did commit some of it to paper.

Indeed, what Fermat wrote down in mathematics conforms to a distinct pattern, which reflects not only the external style of his work, but also its internal content. A study of that pattern, especially in the realm of number theory, yields a twofold reward: It reveals a very likely candidate for the proof Fermat had in mind when he made his tantalizing promise in the margin of Diophantus' *Arithmetica*, and it suggests why he did not set down the proof itself. To see that pattern, however, even in outline, one must make something of a survey of Fermat's entire career to discover what sort of mathematician he was (*4*).

Mathematical Career

Underlying Fermat's mathematical career as a whole is a fundamental tension that often led to paradoxical behavior. Fermat was a bold and ingenious problem-solver who at the same time strove toward, and prided himself on, the full generality of his methods of solution. Almost all of his major achievements had their roots in a concrete solution to a concrete problem. His analytic geometry, for example, which he invented simultaneously with, and independently of, Descartes, stemmed from a restoration of Apollonius' lost work, *Plane Loci*, in particular from Proposition II,5 of that work. His research on quadrature, which established some of the fundamental techniques of definite integration, had its origins in the specific problem of determining the area under each turn of the "Galilean" spiral, $\rho^2 = k\theta$. In each case, Fermat worked outward from the specific problem by articulating its solution into a general method of solution for a general class of problems. To justify the generalization, however, he tended to rely more on his intuitive sense of its efficacy than on theoretical demonstration of its validity. The development of his famous method of maxima and minima illustrates this tendency rather well.

Sometime in the late 1620's, Fermat was trying to find an algebraic derivation of Proposition VII,61 of Pappus of Alexandria's *Mathematical Collection*. That proposition, a lemma to Apollonius' lost work, ***On Determinate Section***, involved the determination of a minimum value, and Fermat was perplexed by Pappus' remark that that minimum value was "singular" (*5*). Analyzed algebraically, the original geometrical problem reduced to an intricate quadratic expression, and it occurred to Fermat that Pappus might have meant that, for the minimum value of the expression, the equation obtained by setting the expression equal to that value has only one root. Fermat knew, however, from his study of François Viète's algebraic theory of equations that quadratic equations always have two roots (*6*). The apparent contradiction prompted him to take a simpler problem and examine it closely.

Turning to Euclid's *Elements*, he investigated (*7*) a special form of the problem to which Proposition VI,27 provides the answer: to divide a given line segment into two parts such that their product is a maximum, or, in algebraic terms, to find the maximum value of $x(b-x)$, where b is the length of the given line segment and x the length of one of its parts. It was not the solution $x = b/2$ that interested him, but rather the equation $x(b-x) = b^2/4$ that resulted from setting the expression equal to its maximum value. By Fermat's interpretation of Pappus, that equation should have only one root, and the standard solution procedures of the time showed that it, in fact, did (*8*).

Fermat was worried, however, about the missing second root. To find it, he altered the equation somewhat: What if $x(b-x)$ were set equal to some value c less than $b^2/4$? Finding that the resulting equation contained the expected two roots, Fermat then took values of c closer and closer to $b^2/4$ and noticed that the difference of the two roots steadily decreased until, at the maximum value of c, that is, at $b^2/4$, the difference was zero. Hence, he reasoned, the second root does not disappear; it is simply equal to the first root. And then came the bold step: to find the maximum or minimum value of an expression, one must find the value which, when set equal to the expression, yields an equation having a repeated root. That repeated root will be the value of the unknown for which the expression attains the extreme value.

Viète's theory of equations provided Fermat with the technical means to turn this general insight into a concrete procedure. Consider the equation $bx - x^2 = c$, where c is some value less than $b^2/4$. That equation has two roots, which we may call x and y, whence the following pair of relations obtains:

$$bx - x^2 = c$$
$$by - y^2 = c$$

Subtracting, we obtain $b(x-y) - (x^2 - y^2) = 0$; dividing through by the (nonzero) difference $x - y$ of the roots, we obtain the final relation $b = x + y$. At first glance, the last equation reveals no more than that the sum of the roots of the original equation is equal to b, the coefficient of the x term; indeed, the purpose of Viète's method was in part to derive what are now called the "elementary symmetric functions" of an algebraic equation. In Fermat's hands, however, the equation became the gateway to a method of maxima and minima. He reasoned that the final relation giving the sum of the roots holds generally for the given equation, whatever the particular value of c. Hence, it holds also for the maximum value of c. But for that value the difference of the roots is zero; that is, $x = y$, or $b = 2x$.

Next, to shorten computation, Fermat replaced x and y as the two roots by x and $x + y$, where y now became the difference of the roots. And, having succeeded with a few elementary quadratic problems, including the original problem from Pappus' *Mathematical Collection*, he extended the method generally (*9*): Given a polynomial

The author is associate professor in the Program in History and Philosophy of Science at Princeton University, Princeton, New Jersey 08540.

$P(x)$, imagine it equal to some value M, not an extreme value. If $P(x)$ is of the second degree or greater, there will be at least two roots, x and $x+y$, for which $P(x)=M$ and $P(x+y)=M$. Hence $P(x+y)-P(x)=0$. Each of the terms on the left-hand side will contain y or some power of y, so divide through by y, that is, compute

$$\frac{P(x+y)-P(x)}{y}=0$$

This equation expresses the relationship between the coefficients of the original equation $P(x)=M$ and two of its roots. That relationship, Fermat maintained, is independent of the value of M. Hence, it holds also for an extreme value of M, in which case the two roots are equal, or their difference is zero. To find, then, the value for x such that M is an extreme value, just set $y=0$.

That is Fermat's method of maxima and minima, in which one finds at least the operational origin of the modern method for differentiation of an algebraic polynomial. Two aspects of it deserve emphasis. First, Fermat's concept of the method involves neither infinitesimals nor limits. Second, and more important for present purposes, its final form is the result of a bold and largely unjustified leap from a special case to a fully general assertion. In the special case, if a is the argument for which a quadratic Ax^2+Bx+C is maximized or minimized, then the equation $Ax^2+Bx+C=Aa^2+Ba+C$ reduces to the form $(x-a)^2=0$. Fermat's final method rests on the far broader claim that, if $P(x)$ is any polynomial and $P(a)$ an extreme value of it, then the equation $P(x)-P(a)=0$ reduces to the form $(x-a)^2R(x)=0$ (*10*).

Problem Solving and Theorem Proving

The development of ideas exemplified here is typical for Fermat. He finds a procedure for solving a particular problem and then generalizes that procedure as far as he can, often farther than logic would dictate. This path from concrete problem to general method often produced a brilliant investigation in the style of Viète's theory of equations: determination of solution families through techniques of reduction to canonical form (that is, the concrete problem), modification of procedures to allow wider application, classification of problems according to means of solution, and so on. In the realm of maxima and minima, the investigation culminated in the successful extension of the derived method of tangents to nonalgebraic curves (*11*). In the realm of quadrature, it ultimately yielded the exquisite reduction analysis of part 2 of the "Treatise on Quadrature" (1658) where, among other things, Fermat established criteria for distinguishing algebraically quadrable curves from those of which the quadrature depends on the quadrature of the circle (*12*).

But that approach to mathematics had its drawbacks. It made Fermat insensitive to the subtle exigencies of rigorous theorem-proving. Success in problem-solving gave him a confidence that frequently blinded him to the pesty counterexample. It meant that he at times lost track of the conceptual foundations of his own work. Nothing illustrates this better than his elegant attempt to justify the method of maxima and minima in a letter to Pierre Brûlart de St.-Martin in 1643 (*13*). That proof, which (like all of Fermat's discussions of the method) employs only finite differences, falters on his failure to distinguish between local and global extreme values.

Fermat begins the proof by noting that the uniqueness of an extreme value, say a maximum, for a given argument implies that, for arguments either greater or less than the given one, the value of the polynomial will be less than the extreme. That is, if $P(a)$ represents the maximum value of a polynomial $P(x)$, then for any y, $P(a)$ is greater than $P(a \pm y)$. Expanding $P(a \pm y)$ into the form $P(a) \pm yP_1(a) + y^2P_2(a) \pm \ldots \pm (-y)^nP_n(a)$, where n is the degree of the polynomial, Fermat then reasons as follows: For the inequality to hold for any nonzero y, first $P_1(a)$ must be zero and second $P_2(a)$ must be less than zero (for a minimum, it would have to be greater than zero). The first condition corresponds to the method of maxima and minima, the second (also original with Fermat) to what is now called the second-derivative criterion of an extreme value. On that much, Fermat is clear. When, however, he begins to muse over the remaining terms, he becomes vague. He takes a few stabs at explaining them away, but ultimately Brûlart must take Fermat's word for it that everything will come out right; the remaining terms will take care of themselves.

The modern reader may spot the flaw in this proof and the reason why, without further conditions placed on y, those remaining terms cannot be ignored. Fermat has failed to distinguish between local and global extreme values. To vitiate the proof, one need only take a fourth-degree expression with two local maxima (or two local minima) and choose a value for y that makes $a \pm y$ the argument for the second extreme value. Also, y must be taken as arbitrarily small: One can always find an expression with two local extremes of the same sort within an arbitrarily small interval.

But Fermat cannot be expected to have distinguished what he did not see, and Fermat literally never saw multiple extreme values. He thought in terms not of general polynomials (to that extent, the symbolic presentation above is unfair to him), but of concrete examples. And all of the concrete problems to which he applied his method of maxima and minima had but one extreme value that made sense in terms of the problem. That is, he was applying his method to geometric constructions, and in each case the analysis resulted in a single value greater than zero (*14*). Thus, the vitiating counterexamples that spring out at the modern reader (used as he is to thinking about general polynomials) from Fermat's proof lay unnoticed by him. For they are theoretical, rather than practical, counterexamples; they undermine the rigor of Fermat's demonstration, not the validity of his procedures in obtaining correct answers. The algebraic polynomials include few, if any, monsters (*15*).

"I am content," Fermat wrote to Father Marin Mersenne in 1636 (*16*, p. 14), "to have discovered the truth and to know the means of proving it whenever I shall have the leisure to do so." Lack of leisure remained Fermat's standard excuse for failing to provide proofs of his results throughout his career. It says more about him than that he was a busy man. As a problem-solver, Fermat neither liked rigorous proofs nor was very good at establishing them. He knew when he had the right answer and an effective means of arriving at it. He measured the generality of his procedures by his success at solving problems with them, and the sense (gained from experience) that they would always work, or could always be made to work, supplanted any felt need on his part for detailed theoretical justification. One sees evidence of this at every turn in his work. Indeed, the "experienced analyst" was frequently the man to whom Fermat referred difficulties: "The experienced analyst will know what to do" (*17*).

Number Theory

One sees this especially in Fermat's number theory. There most notably, his persistent failure to make good his promise of providing proofs he claimed to possess exasperated his contemporaries as it has frustrated historians. There his problem-solving approach to mathematics is most pronounced. Specific problems, usually in the form of challenges to others (Brûlart, Bernard Frenicle, and John Wallis in particular), constituted Fermat's main vehicle of research and exposition. So, in 1640, in the midst of several stunning results regarding perfect and multiply perfect numbers (*18*), Frenicle learned that the basis of Fermat's method of solution is the following theorem (here paraphrased for clarity's sake) (*16*, p. 209): Given a prime p and a sequence of numbers $a^t - 1$ (t positive) based on any positive integer a, then p divides some least member of the sequence, say $a^T - 1$, and T divides $p - 1$; further, p then also divides all subsequent members $a^{kT} - 1$ of the sequence. There it is, a slightly stronger version of what is today known as Fermat's Theorem: For p prime and a prime to p, $a^{p-1} \equiv 1$ (modulus p). But there is no hint of a proof, no trace of a derivation; Fermat offered only a long series of multiply perfect numbers generated on the basis of the theorem.

Then the subject of Fermat's and Fernicle's correspondence changed to a complex of intricate problems involving right triangles in numbers (*19*). Only through the foresight of Father Jacques de Billy (*20*), who added to the 1670 edition of Diophantus' *Arithmetica* (published with Fermat's marginalia) the treatise "Doctrinae analyticae inventum novum," culled from letters sent him by Fermat, do we now have an advantage over Frenicle and Brûlart in knowing how Fermat was solving seemingly impossible problems, such as: Find a right triangle in numbers such that the square of the difference of the two smaller sides exceeds twice the square of the smallest side by a square number (the answer is 1525, 1517, 156). We share his correspondents' quandary, however, when it comes to theorems such as: Every prime number of the form $4k + 1$ is uniquely the sum of two squares. In 1657, Fermat issued two challenges to European mathematicians at large. The second of these asked for the solution in integers of the equation $x^2 = py^2 + 1$ for any nonsquare p (*21*). Fermat claimed to possess a general solution, but the reader of his works searches in vain for the slightest hint of it. One can guess that it looked something like Euler's (*22*, chapter 7), but it is only a guess. And so it goes. Those are only the outstanding examples, and the list could be lengthened almost at will. Fermat's writings on number theory abound with specific solutions to specific problems; his silence concerning his methods is deafening. The vaunted proofs are nowhere to be found.

Close investigation of Fermat indicates that this silence was the norm for him, that the contradiction between his behavior in analysis and his behavior in number theory is only apparent. If Fermat wrote down proofs of his results in analysis, it is because those proofs were dragged from him by his correspondents. He had, for example, to provide justification for the method of maxima and minima and of tangents, if only to defend himself against Descartes' charges of being nothing more than lucky (*23*). He did not put his system of quadrature down in writing until the appearance of John Wallis' *Arithmetica infinitorum* in 1656 threatened his priority and independence. Had he been able to maintain his status as a brilliant problem-solver without revealing his methods he would have done so. His number theory shows that, for there Fermat had no competitors to threaten his predominance; on the contrary, he could evoke little interest in the subject among most of his contemporaries (*24*). In that situation, he could operate as he chose, revealing nothing more than he wished. Hence, the main characteristic underlying many of his memoirs comes to the fore in his number-theoretical jottings. For they are just that: jottings to jog the memory. Fermat wrote down only what he needed to remind himself, not of what he did so much as how he did it.

Method of Infinite Descent

Thus one can explain the disappointing character of Fermat's "Rélation de nouvelles découvertes en la science des nombres," written to Pierre de Carcavi for transmission to Christian Huygens in 1659 (*16*, pp. 431–436). Despite the promise of the title, it relates very little about Fermat's number theory. For example, Fermat described what he called the "method of infinite descent" (addressed to the problem of a right triangle with a square area) (*16*, pp. 431–432):

> If there were any integral right triangle that had an area equal to a square, there would be another triangle less than that one which would have the same property. If there were a second less than the first which had the same property, there would by similar reasoning be a third less than the second which would have the same property, and then a fourth, a fifth, etc., descending *ad infinitum*. Now it is the case that, given a number, there are not infinitely many numbers less than that one in descending order (I mean always to speak of integers). Whence one concludes that it is therefore impossible that there be any right triangle of which the area is a square. . . . I do not add the reasoning by which I infer that, if there were a right triangle of that nature, there would be another of the same nature less than the first, because the argument would be too long and because that is the whole mystery of my method. I will be content if the Pascals and the Robervals and so many other learned men search for it according to my indications.

Although the "whole mystery" can be unraveled in this instance, Fermat went on to relate that he had long been perplexed by how to apply this proof technique to affirmative propositions such as: All primes of the form $4k + 1$ split uniquely into two squares. He was, in his own words (*16*, pp. 431–432), "in a pretty fix." But then, "oft-repeated meditation gave me the insight I lacked, and affirmative questions passed under the aegis of my method with the aid of some new principles it was necessary to add to it." As Fermat described it in general terms, the trick lay in showing that, if some prime of the form $4k + 1$ does not split uniquely into two squares, then some smaller prime of the same form also does not, and then one smaller than the second, and so on down to the smallest, 5, which is, however, composed of 2^2 and 1^2. Hence, by contradiction, the original assumption is disproved. The details of the trick, however, the precise steps by which one reduces a given prime of the form $4k + 1$ not composed of two squares to a smaller prime of the same form also not so constituted, remained Fermat's secret.

In Observation 45 of his "Observations on Diophantus," Fermat lifted the wraps on the "mystery" of his method far enough to permit a reconstruction of the full details of his proof concerning right triangles with square areas. He wrote (*1*, pp. 340–341):

If the area of a triangle were a square, there would be given two quadratoquadrates of which the difference were a square. Whence it follows that two squares would be given, of which the sum and the difference would be squares. And thus a number composed of a square and the double of a square would be given equal to a square, under the condition that the squares composing it make a square. But if a square number is composed of a square and the double of another square, its root is similarly composed of a square and the double of a square, as we can most easily demonstrate. Whence one concludes that this root is the sum of the sides about the right angle of a right triangle, and that one of the squares composing it constitutes the base and the double square is equal to the perpendicular.

Hence, this right triangle is composed of two squares of which the sum and difference are squares. But these two squares will be proved to be smaller than the first squares initially posited, of which the sum and difference also made squares. Therefore, if two squares are given, of which the sum and the difference are squares, there exists in integers the sum of two squares of the same nature, less than the former.

By the same argument there will be given in the prior manner another one less than this, and smaller numbers will be found indefinitely having the same property. Which is impossible, because, given any integer, one cannot give an infinite number of integers less than it.

Narrowness of space prevents inserting in the margin the whole demonstration explained in detail. By this argument we have recognized and confirmed by demonstration that no triangular number except unity is equal to a quadratoquadrate.

Expressed algebraically, and with some missing steps supplied, Fermat's argument reduces to the following: Let the pair p,q of mutually prime integers be the generator of a right triangle; that is, let p^2+q^2 be the hypotenuse, p^2-q^2 and $2pq$ the sides. Assume that the area $pq(p^2-q^2)$ of the triangle is some square a^2. Since p and q are mutually prime, so too are pq and p^2-q^2; if their product is a square, each of them must also be a square. By the same reasoning, if $pq = b^2$, then $p=d^2$ and $q=f^2$. Hence, $p^2-q^2=c^2=d^4-f^4=(d^2+f^2)(d^2-f^2)$. But the factors of this last product are mutually prime, whence $d^2+f^2=g^2$ and $d^2-f^2=h^2$, or $d^2=h^2+f^2$ and $g^2=h^2+2f^2$. By the theorem Fermat cited concerning quadratic forms (*25*), g must be of the form k^2+2m^2. On the one hand, then, $g^2=(k^2+2m^2)^2=(k^2-2m^2)^2+2(2km)^2=h^2+2f^2$, whence $f^2=4k^2m^2$. On the other hand, $g^2=(k^2+2m^2)^2=k^4+4m^4+4k^2m^2=d^2+f^2$, whence $d^2=k^4+4m^4$. But then d is the hypotenuse of a right triangle with sides k^2 and $2m^2$ and with the square area k^2m^2; like the first triangle, the new one must have a generator of the form u^2, v^2. But, since $d^2=p$, clearly $d<p<p^2<p^2+q^2$, the hypotenuse of the original triangle. Hence, by a fully general argument, which can be repeated indefinitely, the assumption of a right triangle in numbers which has a square area entails an infinitely descending sequence of integers (the hypotenuses of the triangles), which is impossible in the domain of integers. Hence, the assumption is disproved.

In his various descriptions of the method of infinite descent, Fermat did not call special attention to an important application he had made of it. He asked in Observation 33 (*1*, p. 327), "But why not seek two quadratoquadrates of which the sum is a square?," and then he answered his own query, "Because this problem is impossible, as our method of demonstration can establish without a doubt." Although he said no more, one can fill in the missing proof by analogy to the proof just presented.

Suppose $a^2=b^4+c^4=(b^2)^2+(c^2)^2$. Then a, b^2, c^2 constitute a right triangle generated by some mutually prime pair of integers p,q, where $b^2=2pq$ and $c^2=p^2-q^2$. Clearly, $q=2r$ (*26*), whence $b^2=4pr$, or $(b/2)^2=pr$. Since p and q are mutually prime, so too are p and r, whence $p=d^2$ and $r=f^2$. Hence, $c^2=d^4-4f^4=(d^2+2f^2)(d^2-2f^2)$, and it can quickly be shown that the factors of the last product are mutually prime. Therefore, $d^2+2f^2=g^2$ and $d^2-2f^2=h^2$, or $d^2=h^2+2f^2$ and $g^2=h^2+4f^2=h^2+(2f)^2$. Now it follows from Fermat's theorem concerning primes of the form $4k+1$ that, if g^2 is the sum of two squares, then g must also be of the form k^2+m^2. On the one hand, then, $g^2=(k^2+m^2)^2=(k^2-m^2)^2+(2km)^2=h^2+(2f)^2$, whence $f=km$. On the other hand, $g^2=(k^2+m^2)^2=k^4+m^4+2k^2m^2=d^2+2f^2$, whence $d^2=k^4+m^4$. But then d^2 represents another square expressible as the sum of two fourth powers, and, since $d^2=p$ is less than $p^2+q^2=a$, clearly d^2 is less than a^2. Again the reduction procedure is fully general and hence entails an infinitely descending sequence of integers. Therefore, there exists no pair of fourth powers of which the sum is a square. A fortiori, there exists no pair of fourth powers of which the sum is a fourth power, since all fourth powers are squares. But the last statement is Fermat's "last theorem" for the case $n=4$.

Last Theorem

Is the method of infinite descent what Fermat had in mind when he spoke in Observation 2 (*1*), the famous "last theorem," of his "marvelous demonstration" that "one cannot split a cube into two cubes . . . "? The close similarity of the two proofs just presented suggests strongly that this is the case. In his "Rélation" to Carcavi and Huygens, Fermat counted the impossibility of a rational solution of $x^3+y^3=z^3$ among the theorems proved by infinite descent. Indeed, whenever Fermat mentioned the two least cases of his "last theorem" (that is, cubes and fourth powers) he did so in connection with the theorem concerning triangles with square areas; the proofs given above show why. Moreover, the same "narrowness of the margin" (*marginis exiguitas*) that prevented him in Observation 2 from carrying out his marvelous demonstration also prevented him in Observation 45 from filling in all the details of the proof outlined there, a proof (to repeat) intimately linked to the proof of the "last theorem" for the case $n=4$.

Fermat stated the "last theorem" in full generality only once, in Observation 2, but he cited the cases of cubes and fourth powers repeatedly in his correspondence. It would seem, therefore, most probable that his success in proving these two cases led him to assume that the method of infinite descent would work for all cases, and to make the assumption without carrying out all the details. It would not have been the first time that he made a (possibly mistaken) conjecture on the basis of not attending to details. For, in the "Rélation," he informed Carcavi that the method of infinite descent had led to the proof of a theorem that he had been struggling with for years, to wit, that all square powers of 2 increased by 1 are prime, that is, $2^{2^n}+1$ is prime for all n. He added, apropos of the proof by infinite descent (*24*):

This last problem results from a very subtle and very ingenious research and, even though it is conceived affirmatively, it is negative, since to say that a number is prime is to say that it cannot be divided by any number.

Clearly, the subtlety and ingenuity of Fermat's proof lay more in his faith in the applicability of his method than in having actually carried out the application in full. As Euler showed in 1732 (*27*), $2^{2^5}+1$ is divisible by 641. And, strikingly, just as the prime number conjecture breaks down for $n=5$, so too the demonstration by infinite descent of the "last theorem" makes a quantum jump in difficulty for $n \geqq 5$; for $n \geqq 23$, the method fails altogether (*28*).

The quantum jump in difficulty just referred to applies to the technical and conceptual apparatus one finds, on the one hand, in Fermat's proof for $n=4$ and Euler's proof for $n=3$ and, on the other, in Legendre's proof for $n=5$. The crux of Fermat's proof is the theorem that, if a square number is the sum of two squares, so too is its root; that of Euler's, a similar assertion about squares of the form p^2+3q^2 (*29*). But, by the same token, Fermat's proof concerning triangles with square areas turns on the assertion (*26*) that "if a square number is composed of a square and the double of another square, its root is similarly composed of a square and the double of a square." That assertion, Fermat claimed, "we can most easily demonstrate." But how? All three of these theorems can be demonstrated in any rigorous way only by the use of complex quadratic fields (even Euler ran afoul of this), and Fermat never wrote down or conceived of a complex number in his life. Rather, it seems clear that he had something else in mind.

The theorem about squares that are the sum of two squares must also have underlain any proof Fermat had for the uniqueness of the decomposition of a prime of the form $4k+1$ into two squares. Investigations he carried out regarding the decomposition of powers and products of such primes also involved claims of uniqueness, and they rested on Fermat's belief that the identity $(k^2+m^2)^2=(k^2-m^2)^2+(2km)^2$ represented the only way to decompose $(k^2+m^2)^2$ into two squares (*30*). Hence, it seems more than likely that the "easy demonstration" of the theorem about squares and double squares derived from no more than the unproved uniqueness of the identity $(p^2+2q^2)^2=(p^2-2q^2)^2+2(2pq)^2$. That is, Fermat avoided quadratic fields by the unjustified use of algebraic identities. This is possible in cases $n=3$ and $n=4$ of the "last theorem" but would not seem to be true for cases $n \geqq 5$, for which resort to quadratic fields is (practically at least) unavoidable.

Hence, I am now convinced that, having proved his theorem for $n=3$ and $n=4$ (the only two cases he mentions more than once), Fermat left off, confident that the same method of proof, combined with the ingenuity he knew was his own, could be applied in all cases. In short, there was no proof that would not fit into the margin; rather, there was a proof schema that had not been thoroughly tested and in fact had only limited application.

Although the "last theorem" almost certainly was not in fact Fermat's last theorem in number theory, it may well serve that purpose here. It sums up Fermat's career, both in number theory and in mathematics as a whole. It is shrouded in mystery because Fermat could not or would not find the time to record his "proof" for posterity, or even for himself. The "proof" probably was no proof, because Fermat could not be bothered with detailed demonstrations of theorems his superb mathematical intuition told him were true. The theorem probably is true because that intuition seldom erred. And Fermat's contributions to number theory, unlike his work in other fields, never slipped into obscurity, because the "last theorem," together with many others, was a true theorem lacking a proof. No mathematician can resist that!

References and Notes

1. C. Henry and P. Tannery, Eds., *Oeuvres de Fermat* (Gauthier-Villars, Paris, 1891–1912), vol. 1.
2. R. Noguès, *Théorème de Fermat, son histoire* (Hermann, Paris, 1932; reprinted in 1966).
3. O. Ore, *Number Theory and Its History* (McGraw-Hill, New York, 1948).
4. For a full account of Fermat's mathematical work, see the author's forthcoming book, from which this article is taken: M. S. Mahoney, *The Mathematical Career of Pierre de Fermat, 1601–1665* (Princeton Univ. Press, Princeton, N.J., in press).
5. Federigo Commandino (1509–1575), who translated the *Mathematical Collection* from Greek into Latin (Pesaro, 1588), could make no mathematical sense of the term and admitted so in his commentary on the proposition.
6. On Viète (1540–1603), the inventor of symbolic algebra and the theory of equations, see: F. Ritter, "François Viète (1540–1603), inventeur de l'algèbre moderne. Essai sur sa vie et son oeuvre," *Rev. occident.* (1905), pp. 234–274, 354–415; J. E. Hofmann, introduction to the reprint edition of Viète's *Opera mathematica* (Elzevier, Leiden, Netherlands, 1646; reprinted by Olms, Hildesheim, West Germany, 1969).
7. In a treatise entitled "Analytica eiusdem methodi investigatio" ["Analytic investigation of the same method (of maxima and minima)"] (*1*, pp. 147–153). For greater detail of the account to follow, see Mahoney (*4*).
8. The solution, that is, would follow operationally the quadratic formula, arriving at a zero characteristic and thus at a single root, $b/2$.
9. In his "Methodus ad disquirendam maximam et minimam" ("Method for determining maxima and minima") (*1*, pp. 133–136). By presenting just the bare algorithm of Fermat's method, this treatise has allowed a great deal of misinterpretation in the past, especially with regard to the conceptual foundations of the method. Moreover, the symbolism employed above for the sake of both clarity and economy is not found in Fermat's treatise, where he proceeds by example and verbal description.
10. In the modern theory of equations, this theorem includes the further condition that $R(a) \neq 0$. Fermat's various discussions of his method leave unclear, however, whether he was aware of this condition.
11. The extension of the method involved, at least implicity, the use of infinitesimals and limits, and it is interesting that Fermat nonetheless made no changes in the finitistic foundations of his original derivation. His problem-solving orientation was such that he may well not have recognized the presence of the fundamentally new elements in his reasoning, especially in the realm of quadrature (*4*).
12. The "Treatise on quadrature" (its short-title form) is printed in *Oeuvres de Fermat* (*1*, pp. 255–285) and discussed in full in Mahoney (*4*). By "algebraically quadrable curves" I mean those algebraic curves of which the area, expressed as an indefinite integral, can be evaluated in closed algebraic form.
13. C. de Waard, Ed., *Oeuvres de Fermat, Supplément* (Gauthier-Villars, Paris, 1922), pp. 120–125.
14. In practice, Fermat dealt only with quadratic and cubic problems involving the division of a line segment into sections satisfying various conditions. In only one case [see (*17*)] did he encounter more than one nonzero extreme value. Of course, for geometric constructions, extreme values produced by negative values of the unknown are discounted. Fermat never tackled a quartic problem, or at least his extant writings give no indication that he did.
15. I might note in this regard that in 1648 Fermat successfully determined the point of inflection of a curve by maximizing the angle made by the tangent and one of the axes. He gave no indication, however, that the general method he proposed corresponded to deriving $P_2(a)$ and setting it equal to zero; nor in turn, having noted that $P_1(a)$ had to be zero for an extreme value, did he ever wonder what would happen if $P_2(a)$ [or, for that matter, any $P_i(a)$] were zero. His mind did not work that way.
16. C. Henry and P. Tannery, Eds., *Oeuvres de Fermat* (Gauthier-Villars, Paris, 1891–1912), vol. 2.
17. Fermat made this remark in the context of an application of his method of maxima and minima to a cubic expression, for which the determination of an extreme value required the solution of a quadratic equation having two positive roots. He left it to the "learned analyst" to decide which of the roots corresponded to the desired extreme value (in this case, a minimum) and did not stop to consider the status of the other root; see (*14*).
18. A perfect number is a number equal to the sum of its proper divisors, including 1, as $6=3+2+1$. A multiply perfect number is one that is equal to an integral submultiple of its proper divisors, as 672 is one-half the sum of its proper divisors, including 1.
19. A right triangle in numbers is a triple of numbers, a,b,c, satisfying the relationship $a^2=b^2+c^2$.
20. Billy was a Jesuit priest and close friend of Fermat's in Toulouse. As a teacher of mathematics at several Jesuit colleges in France, Billy was among the first to introduce the new developments in analysis into the classroom.
21. For the details of this famous challenge, see: Mahoney (*4*) and J. E. Hofmann, "Neues über Fermat's zahlentheoretische Herausforderungen von 1657," *Abh. Preuss. Akad. Wiss. Math. Naturwiss. Kl.* (No. 9) (1944).
22. L. Euler, *Vollständige Anleitung zur Algebra* (St. Petersburg, 1770); revised edition, J. E. Hofmann, Ed. (Stuttgart, 1959); in *Leonhardi Euleri opera omnia* (Leipzig-Berlin, 1911), vol. 1, ser. 1. See part 2, section 2.
23. For about a year between the summers of 1637 and 1638, Fermat and Descartes clashed bitterly over Descartes' theory of refraction and Fermat's method of tangents. Fermat ultimately won the mathematical point, and Descartes afterward retaliated by trying to ruin Fermat's reputation among the Parisian mathematicians. Fermat had to compose the the proof for Brûlart in 1643 because Descartes seemed to be succeeding all too well (*4*).
24. Fermat's inability to interest others in his new results in number theory was the greatest dis-

appointment of his career; it explains in part the fervor behind the challenge problems of 1657 (*4*).
25. Quadratic forms are numbers of the form $a^2 + mb^2$, where m is a nonsquare integer, and a and b integers.
26. Since we assume the relation $a^2 = b^4 + c^4$ to be in reduced form (that is, that the terms contain no common factor), b^2 and c^2 cannot be both even. Since, in turn, $b^2 = 2pq$, one of p or q must be even, the other odd. Were q odd, $c^2 = p^2 - q^2 \equiv -1$ (modulus 4), which is impossible. Therefore, q is even.
27. L. Euler, "Observationes de theoremate quodam Fermatiano, aliisque ad numeros primos spectantibus," *Comment. Acad. Sci. Imp. Petropolitanae* **6**, 103 (1733).
28. See J. Itard's introduction to the 1966 reprint of Noguès' *Théorème de Fermat* (*2*, p. iv).
29. For Euler's proof, see (*22*, paragraph 243); for his theory of quadratic forms, see (*22*, chapter 12).
30. A form of this identity is the basis for Diophantus' solution of Proposition II,8 of the *Arithmetica*, the proposition that prompted Fermat's statement of the "last theorem."

NEWS AND COMMENT

Medicine at Michigan State (II): The Architecture of Accountability

After World War II, the availability of federal funds for biomedical research and hospital construction encouraged a particular growth pattern for the American medical school. First came a basic sciences building; next, a clinical sciences teaching and research facility, often combined with or closely followed by a big university hospital. Frequently, the cycle would be repeated, with the university medical center expanding to monumental proportions.

In the late 1960's, however, medical schools began to suffer the financial pinch caused by the Vietnam war and inflation. Keynesian logic led federal agencies to make their deepest cuts in construction funds, and the effects, consequently, were felt most sharply by new medical schools or those in the midst of building programs.

The new syndrome was nowhere more inopportune than at Michigan State University (MSU), which operates two new medical schools, one of them the College of Osteopathic Medicine. While the MSU schools have the advantage over private schools of receiving state support, they have had to contend with a restive state legislature. The state is committed to supporting four state medical schools, but the legislature has developed something of an immune reaction to the costly prospect of supporting major university medical centers at three campuses (*Science*, 22 September). The legislators' constituents have been complaining about the shortage of physicians in many areas and the high cost of medical care, and the legislators have grown increasingly skeptical about the likelihood that the medical centers will turn out the recruits for general practice and community medicine they think are needed.

The appropriations committees of the legislature have used their considerable practical influence to increase class size at all state medical schools and to encourage an emphasis on family and community medicine. At MSU, the legislators had more leverage than at the University of Michigan or Wayne State, simply because the new schools had to be built literally from the ground up.

It would be a distortion to say that the Michigan legislature dictated the terms under which the two schools would operate. The medical schools themselves initiated the innovations at MSU which resulted in the extensive use of community facilities for clinical teaching and the unusual departmental arrangements for teaching in the basic sciences. And the university at large has a special adaptability that made it possible for innovations to take root. (The organization of medical education at MSU will be discussed in another article.) To ignore the influence of the legislature, however, would be highly unrealistic, and that influence is reflected most clearly in the medical school buildings constructed, and particularly those not constructed, on the MSU campus.

The first building specifically designed and built for medical education at MSU is Life Sciences I, which was completed in 1971. The building cost about \$9.5 million, with some \$4 million of that provided by the state. Life Sciences I was built on a site designated for a medical complex. The site is located near the edge of MSU's expansive main campus, with easy access to interstate roads and virtually unlimited room to build.

Life Sciences I was planned at a time when MSU had only a 2-year medical school. Like many a medical school building elsewhere, it reflects the patchwork financing by which medical schools put facilities for a number of programs under one federally subsidized roof. The building accommodates the school of nursing, the pharmacology department, the offices of the dean of the College of Veterinary Medicine, and laboratories and animal facilities. Architecturally, it is not viewed as a very flexible building.

Actually, the prospects of medical education on the campus were taken into account when building plans were made for the past 10 years. A biochemistry building, completed in the mid-1960's, includes facilities for instruction of medical students, and a new veterinary clinic incorporates space suitable for instruction and research in both human and animal surgery. Both facilities are near—by the standards of MSU's wide-open spaces—Life Sciences I.

But authorization by the legislature of a 4-year program in human medicine in 1969 and the decision a year later to move the College of Osteopathic Medicine to the East Lansing campus presented the university with a greatly increased demand for space. Two buildings have undergone major renovation. Giltner Hall, an academic building, now houses the departments of anatomy, microbiology, pathology, and physiology, as well as animal facilities. MSU's gifts for improvisation were more clearly demonstrated in the renovation of a residence hall for the use of both medical schools. In recent years, student life-styles have changed in such a way that many prefer sharing apartments off campus to traditional dormitory living. As a result, the university was left with empty rooms in the residence halls and a problem in paying off dormitory mortgages. One reaction to this situation at MSU was