Fermat's Methods of Maxima and Minima and of Tangents. A Reconstruction

PER STRØMHOLM

Communicated by C. B. BOYER

It is well known that FERMAT[1] was the first to use the characteristic behaviour of an algebraic expression near its extrema as a criterion for the determination of these extrema.[2] When it comes, however, to the delineation of what his method really consisted of, we are overlooking a field of scholarship of which the famous saying by a Greek philosopher (HERACLITOS, we are told) seems a singularly apt description. The confusion started already in FERMAT's own lifetime. His essays and letters, circulating only in manuscript, were mostly known in selection. The various conceptions of the Fermatian method were therefore determined by the particular part of the corpus that had been studied by the different persons. Add to this the influence of the rapid development of mathematics from about 1630, leading to each succeeding generation of mathematicians interpreting his method within their own conceptual frame, and we cannot wonder that when the *Varia Opera*[3] appeared in 1679, it was already too late. Tradition had by then grown strong and was not to be thwarted by mere texts.

Since then, historians have taken their cues from 17^{th} century conceptions of FERMAT's method. They have either represented it as the expansion of $f(x+h)-f(x)\approx 0$ in powers of h and leaving out second and higher order terms, h being 'infinitely little', or they have taken it to be based on the criterion that the equation $f(x)-y=0$ have a double root.[4] Algorithmically there is no difference between these two approaches. In both cases the mechanical procedure would be to expand $f(x+h)$ and to take the coefficient of the first order term set equal to zero as determining the extrema. This was certainly FERMAT'S way too. However, most historians went further and declared that his *method* (that is, his justification of this algorithm) was identical with one of the two above. As I shall show, there are considerable, if not insurmountable, difficulties in making such views fit with his own words. In my opinion, therefore, not a single one

[1] The main sources are: OF = *Oeuvres de Fermat*, eds. P. TANNERY & C. HENRY. 4 vols. Paris, 1891—1912; + *Supplément*, ed. C. DE WAARD. Paris, 1922. OD = *Oeuvres de Descartes*, eds. C. ADAM & P. TANNERY. 12 vols. Paris, 1897—1912. CM = *Correspondance du Mersenne*, ed. C. DE WAARD etc. Paris, 1932—1967. None of these sumptuous and much-applauded editions reproduces all the relevant material though there is considerable overlap.

[2] The best short account of the history of extremum procedures is still to be found in TROPFKE, *Geschichte der Elementarmathematik*, 2nd edn., vol. 6, pp. 84—91.

[3] *Varia opera mathematica D. Petri de Fermat*, Tolosæ, MDCLXXIX. A good edition.

[4] There also exist other variants, for instance one closely resembling the first, but with h actually being zero.

among the many historians who treated the method of FERMAT before 1929 gave a correct picture of it.[5]

1929 is an important date, because in that year WIELEITNER published a paper[6] which ought to have made a revolution in Fermatian scholarship. He was the first to maintain reason against superstition in this particular field — I believe him to have been a sane man and a sound historian, I do not know of a better thing to say. His crafty approach consisted quite simply in that he did not bring to his task a preconceived opinion of what he was going to find. He was content with reconstructing FERMAT'S conception of the "method of Fermat". This led him immediately to the discovery that to draw from FERMAT'S work a single procedure that could be called his method was impossible. The texts were purely and simply incompatible. He therefore saw what no one before him had seen: that there existed *two* separate and sharply distinct methods of extrema in the work of FERMAT. This pin-pointed the faults of earlier historians — they had always started with the assumptions that there was something that could be called the "method of Fermat", and that this something could be reconciled with his own words. It would seem a natural consequence that both his method and his words would be losers in such a competition.

WIELEITNER gave a good (though not profound) description of the two methods, and he also connected them with FERMAT'S method of tangents. Still, I do not think he pressed his advantage to the full. In particular, he was, in my opinion, wrong on one important account. He seemed to labour under the delusion that editors are a different breed from historians, presumably immunized against human infirmity through divine intervention. Though he was not influenced by the false Fermatian tradition and his judgement of the texts was sound, when he saw some editor's wildest conjecture in a footnote, he instantly accepted it as if coming from a superior being. He misinterpreted the fact that editors are much less often found out than historians to mean that they are much more often in the right. The truth is that editors are nothing but historians in disguise and are therefore disposed to make the same mistakes. This trust in editors led WIELEITNER to accept an impossible chronology of the various writings of FERMAT.

However, when a historian preserves such a great portion of truth to such a small part of falsehood as WIELEITNER, we should expect him to make a profound impression on the historiography of mathematics. Astonishingly, this is not the case. WIELEITNER'S discovery seems hardly to have been noticed at all. To be sure, he is always referred to when it is question of the method of FERMAT, but very few seem to have reaped all the fruits of his perspicacity.[7] I shall base the following analysis and reconstruction squarely on WIELEITNER'S work. When I polemize, it will usually be against him (or the editors of OF, OD and CM), I take the false Fermatian traditions to have been adequately refuted by him.

[5] I have not read everything that has been done in this very tiny grove of learning; the number of historians who have treated it must be close to fifty, but I have sampled enough to convince myself.

[6] "Bemerkungen zu Fermats Methode der Aufsuchung von Extremwerten und der Bestimmung von Kurventangenten," *Jahresber. dtsch. Math. Vereinig.*, **38**, 24—35 (1929).

[7] I suppose the article has now attained the state when it is considered part of that garnish one sprinkles as footnotes.

The study of the mathematics of FERMAT is seriously impeded by the lack of texts and letters from the years 1628—35, which was his first creative period. From what he says himself, he had found the methods of extrema and of tangents as early as 1629.[8] They were, however, not generally known among French mathematicians before the last months of 1637 when MERSENNE received (through the hands of CARCAVI) copies of the two short treatises *Methodus ad disquirendam maximam et minimam* and *De tangentibus linearum curvarum.*[9] FERMAT was immediately embroiled in a dispute with DESCARTES, a dispute that started with his criticism of the latter's *Dioptrique*, but which soon grew to include his own method of tangents.[10] To defend and explain his approach, he engaged in a work of application and clarification in a number of essays and letters, all dating from the years 1638—43.

To FERMAT, his way of finding tangents was only a particular case of a more general method which would also serve to determine extrema and centres of gravity and to solve problems in the theory of numbers.[11] To understand the basic features of this general method and, if possible, to uncover the history of its origins and development, we must study three of its fundamental applications in detail.

Method 1. If we are determined to use the designation "Fermat's method", this should be it. In all his known letters and essays it is, with one notable exception, the procedure he used to find the extrema of algebraical expressions.[12] First a few words about the notation he used and the significance of the expressions he treated. He adherred closely to VIETA's manner of writing equations, using vowels for the unknown (variable) quantities, mainly A and E. For the constant coefficients he used consonants, always beginning with B. Like VIETA, FERMAT never severed the link between geometry and algebra — the variations of the terms of an equation were always limited by the demand that they have a geometrical meaning. This is also noticeable in his insistence on retaining the homogeneity of an equation, for instance writing "B in A — A quad aequale Z plano" (which I shall represent as $BA - A^2 = Z$). For the same reason he never considered negative roots, and if $A = 0$ was a solution of an equation, he did not mention it as it was nearly always geometrically uninteresting.

Let us first see how FERMAT finds the maximum value of the expression above[13] — it is only fitting that, except for a constant sin α factor, it is equivalent to the famous *diorismos* in EUCLID VI, 27. For the unknown A he substituted

[8] Letter to ROBERVAL, 22 Sept. 1636. OF II: 71.

[9] OF I: 133—36.

[10] There does not exist an adequate account of *la querelle des tangentes*, but see MILHAUD, *Descartes savant*, Paris, 1921, esp. ch. VII.

[11] See the letter to ROBERVAL referred to in note 8 above.

[12] *Methodus ad disquirendam maximam et minimam* (1637), OF I: 133—34; *Ad eamdem methodum* (1638), OF I: 140—47; and *Ad methodum de maxima et minima appendix* (1644), OF I: 153—58. See also letters I, IX, XIII, XVIII, XXVI, XXX, LV and LVI in OF II; the essay *Touchant la mesme méthode*, OF Suppl. [V]: 74—83; the letter to MERSENNE, OF Suppl.: 84—86; and the letter to BRÛLART, OF Suppl.: 121—25.

[13] I reproduce the earliest example of his from *Methodus ad disquirendam maximam et minimam*.

$A+E$ and compared the resulting expression to the original one, involving only A, by *adaequation*.[14] This could be represented as:

$$BA - A^2 \sim BA + BE - A^2 - E^2 - 2AE.$$

After having cancelled equal terms, he divided by E, or some power of E, in such a way that the result had terms not containing E and others containing powers (>0) of E:

$$B \sim 2A + E.$$

He now suppressed the terms that still contained E, and the *adaequation* was transformed into a true equality which determined the value of A that made the original expression a maximum:

$$A = \frac{B}{2}$$

He used exactly the same approach to find the extrema of $A^2B - A^3$ [15],

$\dfrac{BZ - BA + ZA - A^2}{GA - A^2}$ [15], $BA - 2A^2 + 2CA$ [16], and $B^3A + B^2A^2 - BA^3 - A^4$ [16].

Nowhere in his early writings did FERMAT give any proof of this procedure. This, as I shall try to show, was for the simple reason that no such proof existed. A demonstration that he published later was logically insufficient, and I do not believe he was ever able to justify his method in a modern sense. We can still learn a great deal, however, from a study of the earlier texts and letters. First, the word '*adaequalitas*', which was XYLANDER'S translation of DIOPHANTOS' '*παϱισότης*' or '*παϱισότητος ἀγωγή*'. This was the Greek algebraist's method for finding a solution of an indeterminate equation as near as possible to a given number.[17] HEATH called it "the method of approximation to limits" [18]; a better rendering would perhaps be 'the method of varying by a small amount'. The basis of FERMAT'S approach was the comparison of two expressions which, though they had the same form, were not exactly equal. This part of the process he called "*comparare per adaequalitatem*" or "comparer *per adaequalitatem*", and it implied that the otherwise strict identity between the two sides of the "equation" was destroyed by the modification of the variable on one side by a *small* amount:

$$f(A) \sim f(A+E).$$

This, I believe, was the real significance of his use of DIOPHANTOS' *πάϱισον*, stressing the *smallness* of the variation. The ordinary translation of '*adaequalitas*'

[14] "et i'ay appellé en mon escrit latin cette sorte de comparaison *adæqualitatem* comme Diophante l'appelle, car le mot grec *παϱισότης* dont il se sert, peut estre ainsy traduit." *Touchant la mesme méthode*, OF Suppl.: 74.

[15] *Ad eamdem methodum* (1638).

[16] *Ad methodum de maxima et minima appendix*.

[17] See *Arithmetica*, V, probl. 11 and 14. For an analysis of FERMAT's dependence on DIOPHANTOS, see I. BACHMAKOVA, "Diophante et Fermat," *Rev. hist. sci.*, **19**, 289—306 (1966).

[18] T. L. HEATH, *Diophantus of Alexandria*, New York, 1964, p. 95.

seems to be 'approximate equality', but I much prefer 'pseudo-equality' to represent FERMAT'S thought at this point.[19]

Secondly, there was never in M1 (*Method 1*) any question of the variation E being put equal to zero. The words FERMAT used to express the process of suppressing terms containing E was "*elido*", "*deleo*", and "*expungo*", and in French "i'efface" and "i'ôte". We can hardly believe that a sane man wishing to express his meaning and searching for words, would constantly hit upon such tortuous ways of imparting the simple fact that the terms vanished because E was zero. Most historians were misled, I admit, by the undated essay *Methodus de maxima et minima*, where there is indeed a magnitude E which is put equal to zero, but this essay treats of quite another method of extrema.[20] After the discovery of FERMAT'S letter to BRÛLART in 1918 and WIELEITNER'S paper of 1929, there remains no excuse for representing M1 as depending on E being zero.

Thirdly, FERMAT told his readers that one was to divide by *some* power of E.[21] This, of course, was wrong as can be seen from

$$f(A+E)-f(A)=\sum_{n=1}^{\infty}\frac{E^n}{n!}f^{(n)}(A).$$

Still, his mistake was understandable. As he could not possibly foresee the peculiarities and future significance of the $f^{(n)}(A)$, he guarded himself against the possibility that $f'(A)$ be zero, a case which might conceivably (to him) turn up in some future problem.

Apart from two letters to MERSENNE[22] where he hinted that the method "tire son principal fondement de ce que $A+E$ fait la même chose que $A-E$" and that "si $A+E$ donne moins que A, $A-E$ doit aussi donner moins que A", this much was known up to 1918 when GIOVANOZZI[23] discovered a copy of the long lost letter to BRÛLART in Florence.[24] It proved to contain the only exposition FERMAT ever gave of what I have called M1 and is therefore of the utmost importance. Written in 1643, it contains his mature thought on extremum procedures, and we should be cautious not to ascribe to him a too great degree of sophistication in the earlier writings.

FERMAT starts by explaining to BRÛLART (and of course to MERSENNE) the foundations of his method:

Mon invention de *Maxima* et *minima* n'a que deux ou trois fondemens. Je suppose premierement que cette recherche aboutit à un point ou à un terme unique, comme, par example, quand on veut *diviser une ligne en sorte que le rectangle sous les segmens soit esgal à un space donné*. Nous avons deux points dans la ligne qui satisfont à la question, mais quand on cherche le plus grand de tous ces rectangles, nous n'avons qu'un seul point qui y puisse satisfaire, lequel, en l'éxample proposé, est celluy qui

[19] He called these pseudo-equalities "comparaisons feintes et *adæquales*." *Touchant la mesme méthode*, OF Suppl.: 75. The term 'pseudo-equality' was used by BOYER, *History of the calculus*, New York, 1959, p. 156.

[20] OF I: 147—53. I treat this as *Method 2*.

[21] *Methodus ad disquirendam* ..., OF I: 133—34; and *Ad eamdem methodum* (1638), OF I: 141.

[22] XXX, OF II: 152; and LVI, OF II: 254.

[23] At the instigation of DE WAARD.

[24] 31 March 1643, OF Suppl.: 121—25. The editor of OF Suppl., C. DE WAARD, questions this date for no good reason at all.

divise la ligne en deux parties esgales. Voylà pourquoy Pappus, dans le septième livre, appelle tousiours *maximam, unicam et singularem,* et tout de mesme *minimam;* le mot grec est *μοναχός*, qui avoit si fort estonné Commandin, qu'il avoue tout net ne point entendre ce que Pappus a voulu dire par ce terme. Il faut donc chercher un point unique, au delà et au deçà duquel tous les termes de la question soient ou tousiours plus grands ou tousiours plus petits que celuy qui sera produit par le point cherché.[25]

We see that FERMAT's conception of the problem of determining A in such a way that $f(A) = \text{extr.}$ was still that of an indeterminate equation. To him, his discovery was a method for transforming this equation into a solvable one. This he managed through use of the fact that if for example $f(A) = \text{max.}$, then $f(A \pm E) < f(A)$, implying that E was "sufficiently" small and $f(A)$ "well-behaved". He never mentioned these restrictions; we can even wonder if he was explicitly aware of their necessity. Still, the fact that $f(A \pm E) < f(A)$ would not have given him a solvable equation if the expansion of $f(A+E)$ had not contained the same terms as that of $f(A-E)$, except for the signs of the odd powers of E:

$$f(A \pm E) = \sum_{n=0}^{\infty} (\pm 1)^n \frac{E^n}{n!} f^{(n)}(A).$$

All the expressions $f(A)$ that FERMAT treated were polynomials, so he had no difficulty in forming the $f^{(n)}(A)$. We should nevertheless write

$$f(A \pm E) = \sum_{n=0}^{k} (\pm 1)^n a_n(A) E^n$$

to imitate the sort of insight he had gained into the expansion of a polynomial of degree k.

From this expansion he took his solvable equation: $\pm a_1(A) = 0$. Solving this: $A = A(B, C, \ldots)$[26], where $B, C, \ldots$ were the constants of the original expression, and substituting: $a_2(A) = a_2(B, C, \ldots) = \pm |c|$, would determine the nature of the extrema. A positive a_2 would mean a minimum, and a negative a maximum. All this FERMAT expressed clearly and forcefully in the letter to BRÛLART — that is, he treated the particular expression $BA^2 - A^3$, but with the understanding that the method was generally applicable to polynomials: "Et la méthode et les raisons que i'ay alléguées, seront générales". Still, we are a long way from having demonstrated the correctness of the procedure. The modern reader might expect FERMAT to assume, or show, that all the $a_n(A)$ were bounded and to take E to be less than a suitable number, which would complete the proof. This is definitely not FERMAT's way — his proof is brilliant, sketchy, difficult, and — unfortunately — wrong.

As the $f(A)$ always were polynomials, the $a_n(A)$ all have the form

$$\sum_{k=0}^{p} \alpha_{nk} A^k - \sum_{l=0}^{q} \beta_{nl} A^l \quad \text{where} \quad \alpha_{nk}, \beta_{nl}, A > 0.$$

After solving $a_1(A) = 0$ and substituting the solution in $a_2(A)$, we shall have one of these two cases: $\sum \alpha_{2k} A^k \lesseqgtr \sum \beta_{2l} A^l$. If, for example, the upper sign applies,

[25] OF Suppl.: 121—22.

[26] FERMAT, of course, used only solutions $A > 0$.

then, says FERMAT, we shall have the following:

$$1=\frac{\sum \alpha_{1k}A^k}{\sum \beta_{1l}A^l}>\frac{\sum \alpha_{2k}A^k}{\sum \beta_{2l}A^l}>\cdots>\frac{\sum \alpha_{nk}A^k}{\sum \beta_{nl}A^l}>\cdots.$$

Now, this would certainly (if it is true)[27] take care of $f(A+E)$, but what about $f(A-E)=\sum(-1)^n a_n(A)E^n$? In fact he did not consider this at all — he seemed to assume that if $a_2(A)<0$ guaranteed $f(A+E)<f(A)$, it would also lead to $f(A-E)<f(A)$. But proving this is impossible unless one also considers the magnitude of the E^n factors, and this he never did.

Let me be quite clear about this. It is obviously not possible to justify M1 without imposing some sort of limitation on the E^n factors, at least in the mild form $E<1$.[28] One will also search in vain in the writings of FERMAT for any mention of such a limitation. Most historians took this in their stride and reasoned that as he had been in possession of a "correct" method, he must also have proved it in a "correct" way, that is, by letting E become infinitely small or actually zero. When confronted with the lack of support for this assumption in his work, they cheerfully postulated that even if he left the condition that E be limited out of his writings, he had *really* thought of it that way all the time. I shall be fair and admit that a superficial reading of the texts could indicate some support for this. We have, for instance, the process of *adaequation*, which I have interpreted to imply a small variation. But this was purely conjectural, and my conjecture was concerned with insuring that in the case of an extremum E would not be so large as to destroy the inequality $f(A\pm E)\lesseqgtr f(A)$. What we are after could also be implied in FERMAT's words when he spoke of the substitution of $A-E$ in BA^2-A^3, resulting in $BA^2-A^3+(3A^2-2BA)E+(B-3A)E^2+E^3$: "Et la dernière puissance de E, qui se trouve tousiours seule, et qui est icy E^3, ne changera point l'ordre de l'équation de quelque signe qu'elle soit marquée, ce qui nous paroistra clairement à la seule inspection." But this hope is immediately destroyed, because he continues: "La raison principalle de cecy est que les deux termes marquez par E^2, estans en plus grande raison que ceux qui sont mesurez par les plus hautes puissances au dessus de E^2"[29]. That is, he was still thinking in terms of the variation of the $a_n(A)$ factors and did not consider the magnitude of E^n at all. If ours is to be an empirical science, we can only conclude that there is nothing to indicate that FERMAT ever thought it necessary to impose any limitation on E to ensure that $a_2(A)$ would determine the nature of the extrema.

This would seem to leave us with an impossible situation on our hands, because here we have FERMAT with a beautifully advanced method of great power, and he did not know why it worked, nor did he know how to prove its correctness. And in this case it seems that the discovery of the method must depend on an understanding of its workings because of its very intricacy. We can hardly believe that he hit upon it by pure chance, but it looks as if there exists no natural or

---

[27] He summed it up in the following strong statement: "Et partant tous les termes qui seront de mesme marqués du signe + seront moindres que ceux qui seront marqués du signe —." OF Suppl.: 125.

[28] This would lead to $a_nE^n>a_{n+1}E^{n+1}$, and if the interval $[A-1, A+1]$ did not contain more than one extremum, one could say that he had at least sketched a proof.

[29] Letter to BRÛLART, OF Suppl.: 125.

rational road to the discovery. Still, in my opinion there does exist such a road, but to attain this we shall have to throw overboard the commonly accepted chronology of some of FERMAT's writings. I leave this for a later part of my paper.

Method 2. This method, which is based on the criterion of the double root, was used in only one work, the undated essay *Methodus de maxima et minima.*[30] It is much simpler than M1 and can therefore be treated less thoroughly. Happily, for once, FERMAT does not leave us in ignorance of his road to discovery:

> While I was working on Vieta's method of *syncrisis* and *anastrophe,* and was carefully investigating its use in the discovery of the nature of constitutive equations, it occurred to me to derive from it a new method for the determination of maxima and minima that will easily resolve all difficulties pertaining to *diorismon,* which have caused so much trouble both in ancient and modern geometry.

The real foundation was, however, PAPPOS' observation which he had also referred to elsewhere[31]: "*Maximae quippe et minimae sunt unicae et singulares.*"

The *syncrisis* of VIETA was a method for expressing the coefficient B of the equation $B^n A^p - A^{n+p} = Z^{\mathrm{hom}}$ in terms of the roots of the equation.[32] If A and E are two of the roots, we have $B^n A^p - B^n E^p = A^{n+p} - E^{n+p}$ and $B^n = \frac{A^{n+p} - E^{n+p}}{A^p - E^p}$. The *anastrophe,* which is not used in M2, was a technique for lowering the degree of an equation when a root of the "negative transform" was known.[33]

FERMAT's first example in *Methodus de maxima et minima* was the expression $BA - A^2$ which he often used. By *syncrisis* we have $B = A + E$, and according to PAPPOS the expression will be a maximum when $A = E$, that is, when $A = B/2$. His second example was $BA^2 - A^3$. Using the method, we obtain $BA^2 - BE^2 = A^3 - E^3$. We should now perform a division by $A^2 - E^2$, but FERMAT had already transformed the *syncrisis:* he divided by $A - E$. I believe that the reason was quite simply that this division always worked out. Setting E equal to A resulted in the equation $2BA = 3A^2$ where he used only the root $A = \frac{2}{3}B$, suppressing as usual $A = 0$. His next step was to name the two roots A and $A + E$ instead of A and E. This would lead, as he observed, to the easier division by E in place of $A - E$. He then proceeded to solve a problem from the 7th book of PAPPOS' *Collectio,* which he had also treated by M1.[34] I shall not analyse his solution — this has been adequately done by HOFMANN.[35] He ended the essay by proposing the following problem "*qui hanc methodum non probaverit*": given three points, find a fourth in such a way that the sum of its distances from the three given ones is a minimum.

[30] OF I: 147—53.

[31] *Ad eamdem methodum* (1638), OF I: 142; and letter to BRÛLART, OF Suppl.: 122.

[32] "Syncrisis est duarum æquationum correlatarum mutua inter se ad deprehendendum constitutionum collatio." *De recognitione æquationum = Opera mathematica,* ed. SCHOOTEN, p. 104.

[33] That is, when one knows a root of the equation which results when one substitutes $-A$ for A. See *De emendatione æquationum, Op. math.,* p. 134—35.

[34] *Ad eamdem methodum* (1638), OF I: 142—44.

[35] "Über ein Extremwertproblem des Apollonios und seine Behandlung bei Fermat," *Nova Acta Leopoldina,* N. F., **27**, 105—13 (1963). HOFMANN represents M2 as $f(x) \approx f(y)$ followed by a division by $x - y$ and putting $x = y$. This is an inaccurate rendering; M2 is based on an *exact* equality.

We can then summarize M2 in the following way: in the equation $P(A)=Z$, where $P(A)$ is a polynomial, we shall determine the values of A which lead to the unspecified constant Z being an extremum. We take A and $A+E$ to be two of the roots of $P(A)=Z$, which gives us $P(A)=P(A+E)$. Division by E and putting E equal to zero, leads to the solvable equation $Q(A)=0$.

Method 3. This is FERMAT's method of tangents.[36] He used it to find the tangents of the parabola[37], the ellipse[38], the cissoid[39], the conchoid[39], and the *folium* of DESCARTES[40]. The solution of the last problem was contained in an important paper where he described the origins of his method of tangents.[41] Here he mentioned that the methods of tangents and of maxima and minima "sont parfaites depuis huit ou dix ans et que plusieurs personnes qui les ont vues depuis cinq ou six ans le peuvent témoigner." In the following description I have for once slightly changed FERMAT's terminology. To him, the unknown (variable) quantity was the length of the subtangent and therefore usually labelled A. I shall use A and $f(A)$ for the abscissa and the ordinate of the curve.

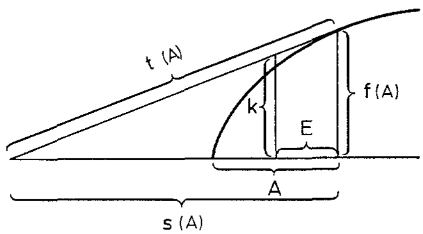

Fig. 1

The problem is to find the tangent at a given point on the curve $f(A)$ — that is, to find the point where this tangent cuts the axis, which again amounts to the determination of the length of the subtangent $s(A)$. FERMAT did this with the help of a comparison by *adaequation* in the following way:

$$\frac{k}{f(A)}=\frac{s(A)-E}{s(A)}$$

and as $k \sim f(A-E)$, we have:

$$s(A) \sim \frac{E \cdot f(A)}{f(A)-f(A-E)}.$$

[36] It has been treated by WIELEITNER ("Bemerkungen ...") and by J. ITARD, "Fermat, précurseur du calcul différentiel," *Arch. int. hist. sci.*, **1**, 589—610 (1947); so I shall confine myself to generalities.

[37] *De tangentibus linearum curvarum* (1637), OF I: 134—36.

[38] *Ad eamdem methodum* (1638), OF I: 140—47.

[39] *Ad eamdem methodum* (1640?), OF I: 158—67. This should not be confused with the piece in note 38. Other and better titles are *De tangentibus linearum curvarum* (see OF Suppl.: xvii) and *Doctrinam tangentium* (its *incipit*). I shall later question the date 1640 given to it.

[40] In a piece known as *Méthode de maximis et minimis expliquée et envoyée par M. Fermat à M. Descartes*, sent to MERSENNE in June 1638. OF II: 156—58.

[41] OF II: 154—62.

As $f(A)$ was a polynomial he could expand it:

$$s(A) \sim \frac{f(A)}{\sum_{n=1}^{\infty} (-1)^{n-1} \frac{E^{n-1}}{n!} f^{(n)}(A)}$$

and by his ordinary procedure of suppressing terms containing E:

$$s(A) = \frac{f(A)}{f'(A)}.$$

I have given a modern rendering of FERMAT'S method of tangents — he was never aware that his results had all this general form. Still, he well understood the usefulness of the method:

> Et ce qu'il y a de merveilleux, est que l'opération nous indique si la ligne courbe est convexe ou concave, si la tangente est parallèle à l'axe ou diamètre, et de quel côté elle fait son concours lorsqu'elle n'est pas parallèle;[42]

All this he took from the form and variation of $s(A)$ and not from any understanding of the significance of $f'(A)$. This can be seen from his determination of the points of inflection of a curve in *Ad eamdem methodum*[43], which he did by finding the extrema of $\frac{f(A)}{s(A)}$.

In the piece mentioned in note 40 FERMAT explained his insistence on the unity of the methods of extrema and of tangents by the fact that he had originally treated the latter as an extremum problem. DESCARTES had interpreted this insistence to mean that the tangent was found as the maximum line from a point on the axis to the curve. But FERMAT now told him that it amounted to the determination of the normal as a line of minimum length. That is, given

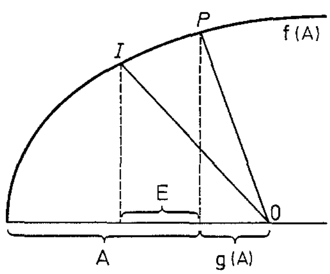

Fig. 2

the point P, if one determined O in such a way that OP was the shortest line from O to the curve, OP would be the normal to the curve in P. The problem was solved by writing

$$g(A)^2 + f(A)^2 \sim (g(A) + E)^2 + f(A - E)^2$$

and using the ordinary routine.

[42] Letter to MERSENNE, 20 April 1638. OF II: 137—38.
[43] OF I: 166—67.

FERMAT was unequivocal when speaking of the chronology of the two approaches to the tangent problem. This is what he said about the method of normals:

> C'est ainsi que j'appliquois ma méthode pour trouver les tangentes, mais je reconnus qu'elle avoit son manquement, à cause que la ligne OI ou son quarré sont d'ordinaire malaisés à trouver par cette voie;

This, he ended, had led him to the standard method which was perfectly general and easier to use.

We search in vain in the written work of FERMAT for any justification of his standard method of tangents. In fact, such a proof would not even have been possible for the form he gave it. This could only have been done through use of the valid *adaequation:*

$$\frac{f(A)}{s(A)} \sim \frac{f(A-E)}{s(A)-E}$$

but this he never did. The nearest he ever came was to notice that $\frac{t(A)}{f(A)}$, $t(A)$ being the length of the tangent, was a minimum.[44]

We now come to the very difficult question about the sequence and chronology of the various methods of FERMAT. The main problem, as I see it, is the interdependence of what I have called M1 and M2. Let me make one thing clear: at the present time, and unless further texts or letters are found, there can be no hope of arriving at a final settlement of the question. Still, no historian up to now has presented an adequate discussion of it. This deficiency I intend to remedy.

As there are two possible relative arrangements of M1 and M2 in time, one could predict that historians treating the question would form in two parties. WIELEITNER, whose account of FERMAT's methods is easily the best, took M1 to be the earliest. According to him, it was discovered at some time during the years 1628—30, and it was straightforwardly applied to the problems of tangents and centres of gravity. The *Methodus de maxima et minima* which contained M2, he alleged, was probably the latest among all of FERMAT's writings on extremum problems. The editor of the supplementary volume of OF, C. DE WAARD, had conjecturally dated it as of 1643—44, so WIELEITNER placed both the treatise and the invention of M2 in 1644. Apart from the dating of IV[45], he gave the following reasons. First, he found no reference to M2 in any other of FERMAT's letters or essays. This would have been very strange, he opined, if it had existed during the period 1638—40 when FERMAT composed nearly everything he ever wrote on extremum procedures and problems. Secondly, he put great weight upon FERMAT's saying that he had invented "a new method for the determination of maxima and minima"[46], thereby implying, WIELEITNER thought, that he had already invented an older one. HOFMANN, who contributed a mathematical note to WIELEITNER's article, has recently reaffirmed this opinion.[47]

---

[44] Letter to MERSENNE, June 1638. OF II: 153.

[45] I shall refer to *Methodus de maxima et minima* as IV, its number in OF I.

[46] "nova ad inventionem maximae et minimae ... methodus." OF I: 147. See my translation of the relevant passage on p. 54 above.

[47] "In einer späteren Darstellung [IV] schlägt Fermat einen rechnerisch viel zweckmäßigeren Weg ein." "Über ein Extremwertproblem ...": 107. The implication is that there was no reason to invent and use M1 if the perfectly good M2 existed.

The other possible chronological arrangement of M1 and M2 was suggested by TANNERY.[48] But as he, of course, had never seen the letter to BRÛLART, he did not distinguish between the two methods. What he tried to do, was to reconstruct the most natural development of the hybrid "method of Fermat". His version has been accepted by ITARD[49], but then ITARD still concurred with TANNERY's conception of the "method".

Before discussing the advantages and disadvantages of the two possible arrangements of M1 and M2, let me make clear my own position. If one in reconstructing any process of scientific discovery or invention wishes to see it as embodying a logical succession of ideas, then WIELEITNER's chronology becomes impossible. The only arrangement which presents the development of FERMAT's thoughts as progressing from the simpler to the more complex, from that which was known to him towards that which is known to us, and which makes it possible to assign psychological causes to the steps he took, is to place M2 before M1 in time. But the desire to find intellectual history a rational process must often yield to the perversity of the texts, which, I suppose, mirrors that of the human mind. If in our case reason is left in the lurch by facts, it will not be for the first time, nor for the last.

Let me first discuss the dating of IV, which is of decisive importance. As I have said, DE WAARD placed it in 1643—44, though he gave no reasons for this dating.[50] I suspect he based it upon the fact that the Groningen MS, which contains a collection of FERMAT's writings made before the summer of 1643, did not have a copy of IV. No more does ITARD, who dates IV as of 1640—42[51], stoop to reveal his grounds. In this case I must confess that I am utterly at a loss — it may be that Professor ITARD has spiritist connections and thus access to sources of information denied to the rest of us.

The fault with these datings is that they do not spring from a study of the contents of IV. All attempted datings of undated scientific writings should proceed from a consideration of the probable relations of their ideas to those of other datable pieces. Only after this should one's conclusions be tested by the circumstantial evidence. Even a superficial study of IV shows it to be a polemical writing, *vide* its ending:

> Confidenter itaque sicut olim, ita et nunc pronuntiamus semper et legitimam, non autem fortuitam (ut quibusdam visum), maximae et minimae disquisitionem hoc unico et generali contineri epitagmate:
>
> Statuatur etc. [referring to the opening part of *Methodus ad disquirendam* where M1 is treated] ... innotescet.
>
> Si qui adhuc supersunt qui methodum hanc nostram debitam sorti pronuntiant,
> Hos cupiam similes tentando exudere sortes
>
> Qui hanc methodum non probaverit, ei proponatur: *Datis tribus punctis, quartum reperire, a quo si ducantur tres rectae ad data puncta, summa trium harum rectarum sit minima quantitas.*[52]

[48] Review of VIVANTI, *Il concetto d'infinitesimo e la sua applicazione alla matematica; Bull. sci. math.*, 2nd ser., **18**, 1894: 230—33.

[49] "Fermat, précurseur ...": 590.

[50] OF Suppl.: xvi.

[51] Op. cit.: 591, n. 8.

[52] OF I: 153.

FERMAT was only involved in one controversy which could have inspired these digs, namely the famous *querelle des tangentes* between DESCARTES and himself in 1638. If we accept this plus the polemical character of IV, it would seem that this must have been written before the end of August 1638. At that time FERMAT received a handsome letter of approval[53] from DESCARTES which ended at least the public part of the controversy. On the other hand, FERMAT's choice of words in IV was clearly determined by his having read DESCARTES' unfair allegation:

> Car premiérement la sienne [méthode] (c'est-à-dire celle qu'il a eu envie de trouver) est telle que, sans industrie et par hasard, on peut aisément tomber dans le chemin qu'il faut tenir pour le rencontrer,[54]

MERSENNE had not originally transmitted these harsh words, but FERMAT knew of their existence from ROBERVAL. In a letter of April we find him asking MERSENNE for a copy of DESCARTES' comments:

> J'attends aussi par votre faveur les réponses que M. Descartes a faites aux difficultés que je vous ai proposées sur sa *Dioptrique*, et ses remarques sur mon Traité *de maximis et minimis et de tangentibus*.[55]

If MERSENNE sent the comments asked for, and we may safely assume that he did so, they would be in FERMAT's hands say a little before 1 June 1638.

Let us try to reconstruct the psychological situation of FERMAT at this date and see what can be deduced from it. He had just received from MERSENNE a copy of DESCARTES' letter of 18 January where the latter insinuated that the method had probably been found by pure luck and not as the result of an industrious search for the solution of a general problem. In the same package from MERSENNE he probably also found a copy of DESCARTES' letter to MYDORGE[56] answering the objections that FERMAT had raised in his second letter on the *Dioptrique*.[57] Thirdly, shortly after 1 June he received a letter from ROBERVAL[58] who had twice (in partial collaboration with E. PASCAL) written against DESCARTES in defence of FERMAT's method of tangents. It appears from what ROBERVAL says that he had not undertaken this defence out of any deep understanding of the method, but rather from a general dislike of DESCARTES. He expressly asked FERMAT to explain the foundations of the method of which he had only seen a particular case:

> Mandez-moi quel est votre sentiment, car, n'ayant pas encore le loisir de considérer bien partculièrement le fonds de votre méthode et sa démonstration, il se peut ètre qu'elle ne contienne des mystères qui me sont encores cachés.[59]

If the behaviour of a rational human being is predictable at all, it would seem that about 10 June FERMAT would be engaged in three separate tasks. First, he was writing (or had just written) an account of the discovery of his basic

---

[53] XXXII in OF II, or CXXXII in OD II. Most of the letters I cite can be found in both OF, OD, and CM.

[54] Letter to MERSENNE, 18 Jan. 1638. OF II: 129.

[55] OF II: 136:

[56] OD II: CXI.

[57] OF II: XXIV.

[58] OF II: XXIX.

[59] OF II: 149.

extremum procedure to defend himself against DESCARTES' allegation.[60] This account, I submit, was the *Methodus de maxima et minima* (IV) which contains M2, and which up to now has been variously dated as of 1640—44. Secondly, he was writing a general description of his method of tangents, primarily meant for ROBERVAL (and PASCAL) as an answer to the latter's request in the letter of 1 June, but also for perusal by DESCARTES. This description was the piece XXXI in OF II with the title *Méthode de maximis et minimis. Expliquée et envoyée par M. Fermat a M. Descartes.* If the perspicacious reader protests that the title precludes XXXI being written for ROBERVAL, I shall counter that particular objection with the time-honoured expedient of historians — I shall explain it away. The title of XXXI was clearly MERSENNE'S, as can be seen from the ARBOGAST-BONCOMPAGNI MS[61], and not FERMAT'S own. Moreover, the latter consistently referred to DESCARTES in the third person in the body of the piece. Still another indication of the inappropriateness of the title is that XXXI treats of the method of tangents and not of extrema as the title implies. I unblushingly suggest that MERSENNE attached it at some confused moment when he had not yet established the true correspondence between the various writings of FERMAT which he received about 1 July and the recipients as specified in FERMAT'S accompanying letter. Thirdly, it seems that FERMAT wrote his third piece of criticism against DESCARTES' *Dioptrique*. This would have been a rejoinder to DESCARTES' letter to MYDORGE mentioned above.[62] But this third piece is of no particular interest in our context, so I leave it quietly aside.

It is, then, my conjecture that these three pieces (or at least the first two) were sent to MERSENNE with an accompanying letter on 15 June 1638. The first part of this letter is well known — it is found as letter XXX in OF II.[63] Another part — not so well known — is found as document VI in the supplementary volume of OF, beginning with the words "I'ay receu un mot de vostre part[64] de M^{r} Roberval ...", and expressly dated 15 June. I strongly suspect that a completely unknown fragment of the same letter is to be found as section 6 of FERMAT'S letter of 20 April to MERSENNE[65], beginning with these words: "Outre le papier[66] envoyé à R[OBERVAL] et P[ASCAL] ...".[67]

[60] It is possible that FERMAT had written a (lost) letter to MERSENNE, say about 20 May, complaining of DESCARTES' insinuations. See the letter DESCARTES-MERSENNE of 29 June, OD II: CXXVI, esp. note *a* on p. 175 and note to p. 263, l. 9 on p. 278 in OD II. If there did exist such a letter from FERMAT, it could have been occasioned by ROBERVAL'S account of DESCARTES' remarks, or it could mean that FERMAT had received the copy of DESCARTES' letter of 18 Jan. from MERSENNE in the middle of May.

[61] See OF I: xxiv.

[62] March 1638. OD II: CXI. FERMAT'S piece is lost but it seems a safe conjecture that he wrote a third letter on the *Dioptrique*. It was in DESCARTES' hands before 27 July as can be seen from the latter's letter to MERSENNE of that date (OD II: 263—65).

[63] Conjecturally dated June 1638 in OF. See OF II: 152—54.

[64] Referring, probably, to ROBERVAL'S letter of 1 June.

[65] OF II: 137—38.

[66] If my conjecture is true, this would refer to XXXI.

[67] I shall not argue this at length, though the editorial work in OF was not particularly distinguished, to put it mildly. Apart from the sense of section 6, which bears no relation to the letter of 20 April at all but which links very well with that

I suggest that the reader piece the letter of 15 June together the way I have reconstructed it and read it under the view that it was accompanied by IV and XXXI. I believe it will be found that the whole makes excellent sense, and that it fits nicely with a number of remarks in the correspondence of DESCARTES and FERMAT with MERSENNE from the last half of 1638. I shall only point out one of these corroborating details. The letter XXX in OF II begins with the following words:

J'avois déjà fait un mot d'écrit pour m'expliquer plus clairement à M. Descartes, sur le sujet de ma méthode *de maximis et minimis et de inventione tangentium*[68], lorsque votre dernière m'a été rendue, qui contient copie de la réplique de M. Descartes.[69] Je ne reste pas de lui envoyer ce que j'avois déjà fait, où il trouvera sans doute de quoi se désabuser de la croyance qu'il semble avoir, que j'ai trouvé cette méthode par hasard et que je n'en connois pas les vrais principes.[70]

The last period refers, in my opinion, to IV, and not as the editor[71] of that part of OF maintained, to XXXI. IV does in fact constitute a refutation of DESCARTES' imputation, while this could hardly be said about XXXI. All three fragments of the letter of 15 June will stand up equally well to tests of this kind.

When MERSENNE received FERMAT'S package of writings, probably towards 1 July, he did not immediately transmit all of it to DESCARTES. It seems that he withheld XXXI for a time[72], but enclosed IV in a letter which reached the latter on 13 July. This appears to be a safe deduction from DESCARTES' answer of the same date:

I'en estois paruenu iusques icy lors que i'ay receu vostre derniere auec l'enclose de M. F[ermat], à laquelle ie ne manqueray de répondre à la premiere occasion; & ie serois plus marry qu'il m'eust passé en courtoisie qu'en science. Mais pour ce que vous me mandez qu'il m'a encore écrit vne autre lettre pour la deffence de sa regle, & que vous ne me l'auez point enuoyée, i'attendray que ie l'ay receuē, afin de pouuoir répondre tout ensemble à l'vne & à l'autre. Et entre nous, ie suis bien aise de luy donner cependent le loisir de chercher cette Tangente, qu'il a promis de vous enuoyer au cas que ie continuasse à croire qu'elle ne se peut trouuer par sa regle.[73]

Then, on 20 July, MERSENNE sent XXXI, which was received by DESCARTES on the 27th. The answer that the philosopher had promised to both of FERMAT'S writings on extremum procedures took the following characteristic form:

I'en estois paruenu iusques icy, lorsque i'ay receu vostre dernier pacquet du 20 de ce mois, lequel ne contient que des escrits de Mr Fermat, ausquels ie n'ay pas besoin de faire grande response; car pour celuy ou il explique sa methode *ad maximas*, il me donne assez gaigné, puisqu'il en vse tout autrement qu'il n'auoit fait la premiere fois, affin de la pouuoir accomoder a l'inuention de la tangente que ie luy auois proposée; & selon ce dernier biais qu'il la prend, il est certain qu'elle est tres bonne, a

of 15 June, the position and the title of the fragment in the ARBOGAST-BONCOMPAGNI MS seems conclusive. The editor of CM, C. DE WAARD, could not help seeing that section 6 did not belong to the letter of April, but the dating he suggested — January 1638 (CM VII: 6—7) — is worse than wrong; it is impossible.

68 That is, a defence against DESCARTES' letter of 18 Jan.

69 This *réplique* was probably DESCARTES' letter of 3 May.

70 OF II: 152.

71 TANNERY.

72 This would confirm my conjecture that XXXI was primarily meant for ROBERVAL.

73 OD II: 250. The piece DESCARTES received could not have been XXXI because this contained the solution to the problem he mentions. I beg the reader to note that DESCARTES mentions *two* pieces by FERMAT on the method of extrema and tangents.

cause qu'elle reuient a celuy duquel i'ay mandé cy-deuant qu'il la faloit prendre. En sorte que, pour en dire entre nous la verité, ie croy que s'il n'auoit point vû ce que i'ay mandé y deuoir estre corrigé, il n'eust pas sceu s'en demesler. Ie croy aussi que toute cete chiquanerie de la ligne EB, sçauoir si elle deuoit estre nommée la plus grande, que ses amis de Paris ont fait durer vn demi-an, n'a esté inuentée par eux que pour luy donner du tems a chercher quelque chose de mieux pour me respondre Et ce n'est pas grande merueille qu'il ait trouué en six mois vn nouueau biais pour se seruir de sa regle; mais on n'auroit pas de grace de leur parler de cela, car il n'importe pas en combien de tems ny en quelle façon il l'a trouué, puisqu'il l'a trouué.[74]

Let us then sum up what can be learnt from the study of the letters of DESCARTES and FERMAT from 1638. There is no decisive proof of my conjecture that IV was written in June 1638, that is, if one considers the logical connotations of 'proof'. On the other hand, I believe that the number of indications, which I have given samples of above, that point to my conjecture being true, is so large that it warrants the tentative acceptance of that conjecture. Finally, I have found nothing that is incompatible with the dating I have suggested.

If we turn to the MSS for further illumination, we are seriously hindered by the fact that there exist very few holographs of FERMAT. However, two of the MSS single themselves out as worthy of a closer study than the rest. The one is the ARBOGAST-BONCOMPAGNI MS, a copy made by ARBOGAST of the pieces relating to FERMAT in one of MERSENNE's letter-books.[75] In this letter-book IV was contained on two separate leaves in the handwriting of MERSENNE, while the fragments of the letter of 15 June, partly in FERMAT'S own hand, were found in various places. I have already mentioned that the position of section 6 of the letter of April was suggestive — it is to be found as item 6 of the MS with the title *De maximis et minimis* and is immediately followed by XXXI (*Méthode de maximis expliquée et envoyée par M. F. à M. des C.*).

The other important source, the Firenze MS[76], contains in folios 75*r* to 117*r* copies of various letters and treatises of FERMAT. These copies stem from transcriptions made by MERSENNE and communicated to Italian mathematicians. The chronological aspect of their arrangement is very interesting. Folios 75*r* to 88*v* contain five treatises from 1637—38 in chronological order. Then follows (89*r*—92*r*) *Ad eamdem methodum* conjecturally dated 1640.[77] In folios 93*v* to 97*r* we find

[74] OD II: 272—73. For a still more violent tirade see the letter of 23 Aug. to MERSENNE, esp. pp. 320—26. DESCARTES' part in *la querelle des tangentes* is not usually too flatteringly depicted by historians. He mostly comes out as a blockhead who was too dense to appreciate the finer points of FERMAT's method, or conversely as maliciously misrepresenting it to score cheaply. As I have said above, the method of tangents was an unjustified extension of the extremum method, and this DESCARTES was quick to seize upon. It was only after receiving IV, which implied that there was a sounder basis for it, that he pulled back.

[75] See OF I: xxii—xxvii for a description.

[76] It is in fact vol. 45 of the collected MSS of VIVIANI and bears the title *Scritti di diversi autori sopra varj soggetti matematici*. See OF Suppl.: Intr., for a description.

[77] This date has now crept into the secondary literature. The treatise was undated in OF, and it seems that the source of the 1640 date was DE WAARD (OF Suppl.: xvi). Again he gave no reasons for his dating; the only thing I can think of is that he made an unwarranted inference from ROBERVAL's letter of 4 August 1640. Personally, I am convinced that *Ad eamdem methodum* (OF I: 158—67) was written in the autumn of 1638.

IV immediately followed by an extract from the letter of 15 June 1638.[78] Then follows, still in chronological sequence, a suite of letters from the years 1636—38 (97*r*—110*r*). The whole collection of FERMAT items ends (110*v*—117*r*) with a new suite of five essays and letters from 1635—44 in chronological order. That the whole arrangement is not accidental can be seen from the Groningen MS.[79] This lacks a few of the items of the Firenze MS (among these IV), but it preserves the exact relative positions of the remaining. My argument is that as the Firenze MS has preserved, with the possible exception of *Ad eamdem methodum*, the correct intrinsic chronological arrangement of the various suites, the position of IV points to its being written in 1638 and associates it closely with the letter of 15 June.

Now for the rest of the props supporting WIELEITNER's dating of IV. FERMAT, when speaking of M2 as *methodus nova*, was clearly contrasting it with the *ancient* method, that is, the one used by Greek mathematicians. If not, there would be no point in mentioning the difficulties of *diorismon* etc., because then M1 would already exist to take care of these. If FERMAT had wished to indicate that M2 was his second method, he might well have spoken of it as *alia methodus*. Another of WIELEITNER's "indications"[80] was that in the Firenze MS IV bore the title *Analytica ejusdem methodi investigatio*. If this was the title FERMAT gave it, and there is no reason to doubt this, it was surely meant as referring to *Methodus ad disquirendam maximam et minimam* of 1637, thus linking IV still more closely to the period 1637—38.[81]

But what about the WIELEITNER-HOFMANN argument that if M2, being as good a method as M1 or even better, was invented first, why was it not used exclusively, and why was it not mentioned in any other place than IV? This argument fails for two reasons. First, M1 is undoubtedly a better method than M2. It will determine the nature of the extrema, and it was, at least potentially, a powerful tool for attacking some of the most important problems of 17$^{\text{th}}$ and 18$^{\text{th}}$ century mathematics. M2, of course, was comparatively sterile. Secondly, even admitting that the two methods were objectively of comparable importance, this would not necessarily influence their relative frequencies of occurence in FERMAT's work. This would surely have been determined by the value attributed to each of them by him. As I shall argue later, he might well have been of the belief that M2 suffered from logical faults. This would explain why he never mentioned it except for the single occasion when he described the genesis of all his methods.

Up to now my arguments have only tended to show that IV was written in 1638. But the wary reader will by now suspect what I am leading up to. If we concede that IV was probably composed in June 1638, it seems that we must accept that M2 was FERMAT's earliest method of extrema. Because, granting the first part of my argument, if we assume that M2 was also invented in 1638, then it would be very strange if it was not mentioned in any of FERMAT's numerous

[78] The extract is document VI in OF Suppl.

[79] See OF Suppl.: Intr. for a description.

[80] See "Bemerkungen ...", p. 30, n. 1.

[81] See the ending of IV (OF I: 153) for such a direct reference. I have given the passage on p. 58 above.

writings on extremum problems from that year. The only way to explain his silence would be to relegate the invention of M2 to the early period of his scientific life from which we have no documentary evidence. Later, when he invented M1, he considered M2 to be insufficient, or superseded, or both.

But these deductions from FERMAT'S silence are at best only circumstantial. My identification of M2 with FERMAT'S earliest extremum method will be supported by three arguments. The first will be his own words in the first part of IV which I have translated above (p. 54). As I have said, this is the only instance we know of where he described the invention of any method of extrema or of tangents. There would be no point in his obvious pride in a method that could handle problems in connection with *diorismon* etc., if this method were not the first. Also, if IV was intended as a refutation of DESCARTES' insinuations, there would be little point in dragging in a second and later method which had never been used for anything. I claim that the only chronological arrangement that has at least some foundation in FERMAT'S own words, is the one which takes M2 to have been his earliest extremum procedure.

My second and third arguments do not have the same force as the first. We know that several of FERMAT'S friends had been shown examples of the application of his methods during the period 1630—35. He writes, for example, in a letter to ROBERVAL: "Sur le sujet de la méthode *de maximis et minimis*, vous savez que, puisque vous avez vu celle que M. Despagnet vous a donnée, vous avez vu la mienne que je lui baillai, il y a environ sept ans, étant à Bordeaux."[82] Now, a possibility of verifying (or falsifying) my chronology would be to find out what DESPAGNET'S conception of FERMAT'S method was: M2 or M1? I am not in a position to do this, but let me draw attention to the following passage in another of FERMAT'S letters to ROBERVAL: "Toutes ces propositions ... dépendent de la méthode dont M. Despagnet ne vous a pu faire voir qu'un seul cas, parce que, depuis que je n'ai eu l'honneur de le voir, je l'ai beaucoup étendue et changée."[83] He had extended the method, yes, but changed it? Surely, this must refer to the replacement of M2 by M1 at some time between 1629 and 1636.

Another of FERMAT'S early correspondents was BEAUGRAND, but their connections were severed around 1635. From various remarks in FERMAT'S letters we know that his friend had been shown examples of the workings of some of the methods. It is possible that BEAUGRAND'S conceptions of these methods would refer to the early phase of their development. We have in fact an exposition by BEAUGRAND of *a* method of tangents, probably dating from the autumn of 1638.[84] This method can be represented as follows. From Fig. 3 we have the equality

$$\frac{f(A+E)}{f(A)}=\frac{E+s(A)}{s(A)}$$

If $f(A)$ is a polynomial, we can expand and divide by E. If we now put E equal to zero, the secant is transformed into a tangent, and we have solved the problem.[85]

[82] 22 Sept. 1636. OF II: 71.

[83] 16 Dec. 1636. OF II: 94.

[84] Printed in OF Suppl.: 102—13. BEAUGRAND implicitly represented this method as his own, and FERMAT was barely mentioned in the text. He was later severely castigated by DESARGUES and PASCAL for not making clear his dependency on FERMAT.

[85] BEAUGRAND used 0 (zero) instead of E.

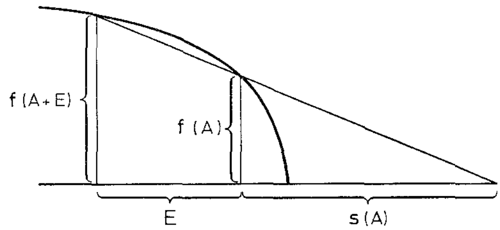

Fig. 3

It should be perfectly clear that BEAUGRAND'S method consists of nothing but the application of M2 to the tangent problem. I can think of at least three possible ways by which this could have come about. First, BEAUGRAND had originally been shown M1 by FERMAT and later, as his friend, he hit upon M2. This would then be a case of simultaneous discovery and should be brought to the notice of Professor MERTON. Secondly, BEAUGRAND was not inspired by FERMAT at all — the method was solely and wholly his own. Thirdly, BEAUGRAND received M2 from FERMAT in the early 1630's at a time when this was the only method in existence. I shall not try to squeeze more out of this slender bit of evidence, but leave the reader to assign his personal relative probability to each of the three explanations suggested.

It is now time to attempt a reconstruction of the development of FERMAT'S methods of maxima and minima and of tangents, a reconstruction which incorporates the chronology I have maintained in the earlier parts of this article. I do not pretend that this reconstruction is the only one possible which embraces that particular chronology — on the contrary, it is my contention that the falsity of the former does not imply that of the latter — but I feel that it is, perhaps, the neatest and simplest conceivable.

Sometime between 1628 and 1630 FERMAT was in possession of the necessary tools for the solution of the general extremum problem. He knew that in Greek mathematics it was essentially connected with the *diorismos*. He had also closely studied the *Conics* of APOLLONIOS, especially the fifth book where extremum problems were treated in their own right. The weak point in the ancient method, he saw, was that one had to know the solution beforehand; only then could one prove that this solution really was the correct one. Thus, what was needed was a procedure that would not only guarantee the correctness of the solution, but which would directly furnish that solution itself. He had one thing which gave him an enormous advantage over the ancients: the algebra of VIETA, which he was gradually transforming into a true analytic geometry. However, the two pillars on which he based his method were, on the one hand, PAPPOS' observation that extrema were unique and singular and, on the other, VIETA'S method of *syncrisis*. Up to this point I do not believe it possible to doubt this reconstruction — we have FERMAT'S express words to rely on. And now I ask the reader, was the method that issued the relatively simple M2, which is in fact the exact embodiment of the two pillars of PAPPOS and VIETA, or was it conversely the complex M1 which was built entirely on other foundations?

If we provisionally accept that this method must have been M2, let us proceed and see in what way FERMAT could have been led to the method of tangents. If we believe his words — and I have based all my arguments on this — he began by tackling the problem of normals by treating it as an ordinary minimum problem. This was in fact a natural thing to do, because the main part of book five of the *Conics* treated of such problems. But now he ran up against the first serious difficulty. M2, which worked well enough with ordinary extremum problems, certainly furnished the solutions for problems of normals too. It was not that it was insufficient, but rather that it was intellectually unsatisfactory.

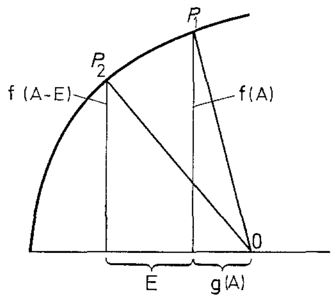

Fig. 4

In Fig. 4 FERMAT had $OP_1 = OP_2$, and from this he easily formed the equality

$$g(A)^2 + f(A)^2 = \left(g(A) + E\right)^2 + f(A-E)^2$$

which gave him the solution. But somehow this seemed too easy. If he imagined some specific point O, this must mean that there already existed a point P between P_1 and P_2 in such a position that OP was a line of minimum length. But this again meant that $g(A)$, which depended on the positions of both O and P, was not a free quantity and thus that the equation he formed could not be an exact equality but only an approximation. FERMAT'S thoughts were not, of course, as explicit and clear as this. The ambiguity in the notation of VIETA, where A represented both the unknown and one of the roots of the equation, would cloud the issue.[86] Still, he could not help feeling that though the method gave correct results, there was something wrong with it, at least when used for the determination of normals. This uneasy feeling, combined with the cumbersomeness of the method, made him continue his search for other ways of solving the tangent problem. I believe he took his cue from a study of the graphical representation

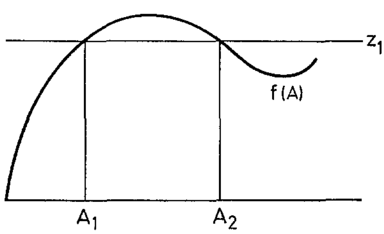

Fig. 5

[86] This ambiguity does not matter in the ordinary extremum problem and is therefore not so easily discovered there.

of the solution of the equation $f(A)=Z_1$ (see Fig. 5). His method of extrema treated the problem where $Z_1=Z_{\max}$, in that case the secant was transformed into a tangent. But this was a very special tangent, namely the one parallel to the axis. However, a short study would perhaps have led him to draw something

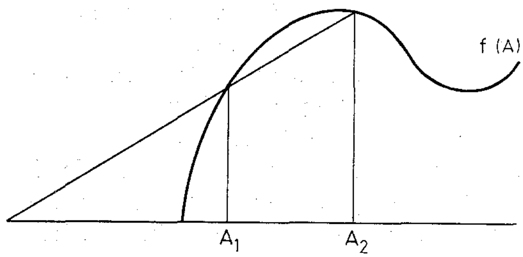

Fig. 6

like Fig. 6. M2 rested on the fact that $f(A_1)=f(A_2)$, but this was clearly not applicable in this case. Still, I believe it must have been an easy thing for FERMAT to begin thinking in terms of the subtangent $s(A)$, which would immediately have led him to the sufficient equality

$$\frac{f(A_1)}{s(A_1)}=\frac{f(A_2)}{s(A_2)}$$

and the solution of his problem.

We have now come to the point when we have FERMAT in possession of a method of extrema and the corresponding method of tangents. The question is now: if he had these methods, which according to historians were perfectly adequate, why and how did he come to search for and invent yet another set of methods? The why, I believe, is easily answered. First, there existed a certain class of problems — the determination of normals was one of them — for which M2 did not seem satisfactory. Secondly, I believe that the double process of first dividing by E and then setting E equal to zero did not appeal to him. It will not do here to drag forth the time-honoured "limiting process" of historians of mathematics — M2 depended on E being actually zero. FERMAT was still thinking in terms of equations; I agree that he stood on the verge of a period where mathematicians came to accept that sort of process, but he himself was in this particular case rather the last of the ancients than the first of the moderns. He just could not stomach the fact that one should be allowed to divide by a zero quantity when treating equations. His qualms led him back to the problem of normals,

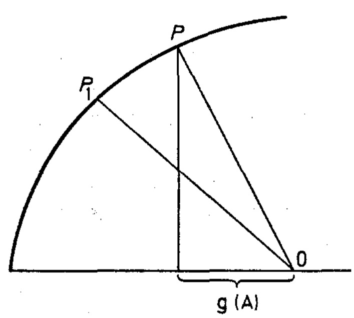

Fig. 7

and by now he was ready to rout out the implicit assumptions that had caused his earlier uneasiness (see Fig. 7). He started by drawing the line OP which was to be a normal and therefore of minimum length. If he now drew OP_1, where P_1 was a point close to P, it was clear that $OP_1 > OP$ and that the original equality could only be approximate: $OP_1^2 > OP^2$. Though he only dimly understood the difficulty caused by the double role that the quantity A had been made to play, it was easy to transpose the improvement into the terms of the original extremum procedure. This is a very fascinating thing about the development of FERMAT's thoughts. When confronted with the choice between the two roles of A, he chose that of the *root*, and thus took a decisive step towards the delimitation of a function-concept. That is, he applied what he had learnt from the problem of normals, where this step was necessary, to the ordinary extremum problem, where

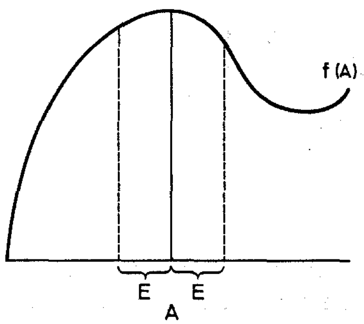

Fig. 8

it was not necessary. He now began by considering this problem as solved (see Fig. 8), with A as an undetermined constant. The criterion for a maximum would then be

$$f(A \pm E) < f(A)$$

and the converse for a minimum. He had now taken care of one of the difficulties connected with M2 and, inadvertently as it were, taken the first step on the way from the concept of an indeterminate equation towards that of a function.

The inequality he arrived at could of course be treated in exactly the same way as the old equality: he could divide by E and then put E equal to zero. But, as I have said, this he did not want to do. He kept the first part of the process, the division, but E would have to remain finite in some way or other.[87] From M2, which certainly furnished correct answers, he knew that it was the terms that did not contain E which gave the final solution. As a temporary move he instituted the mechanical process of suppressing the terms containing E. This, of course, had the same effect as setting E equal to zero, but now he had eliminated that particular obnoxious process. The only thing left was the justification of M1. That is, he had to show that the second order term in the expansion of $f(A \pm E)$ dominated the terms of the higher orders.

[87] This is a strong argument in support of my chronology: why was there a division by E involved in M1? It is quite unnecessary there. The answer is, of course, that it was a relic from the older M2.

We have now reached the period from which we have the earliest documentary evidence, about 1636. FERMAT had resolved the two difficulties connected with M2: the ambiguity caused by the double role of A, and the division by a zero quantity. He possessed a method whose correctness he was convinced of, but somehow he was unable to justify it. He generalized the criteria for maxima and minima in the process of *adaequation*, and it was an easy feat to apply the whole to the tangent problem, though in the way he did this it was an unjustified extension. Still, I believe that most of his work in the period 1635—40 went into the search for a proof of M1. Through this he reached astonishing heights in his understanding of the properties of the expansion of expressions $f(A \pm E)$ in powers of E. Not only did he find that the coefficient of E^2 determined the nature of the extrema, but I also believe that he at least partially understood the rules by which the coefficients a_n could be formed from a_{n-1}. He also found a brilliant method for determining the extrema of expressions of the form $f(A) \pm \sqrt{g(A)}$[88], and even $\sqrt{f(A)} \pm \sqrt{g(A)}$[89], though in the latter case the resulting equation was of the 6^{th} degree and therefore unsolvable. But still, even his most ambitious attempt to prove M1, as it was outlined in the letter to BRÛLART, was inadequate. Had he in fact come to see the necessity of imposing some limitation on the magnitude of E, it would have meant an amazing short cut through the bewildering field of conceptions about zero divisions and infinitely small quantities of the 17^{th} century. Just the same, the power and depth of FERMAT's mathematical thought is astonishing. He did not rest content, as mathematicians of lesser stature would have done, with an "adequate" method, but incisively diagnosed its deficiencies and thereby created one of calculus' most powerful tools.

D. T. WHITESIDE read an early version of this paper and offered certain suggestions for improvement, some of which have been incorporated. I have also received helpful comments from Professor CARL B. BOYER.

[88] *Ad methodum de maxima et minima appendix*, OF I: 153—58.

[89] He never expressly mentions this, but it was obviously the way he attacked the general anaclastic problem in 1661.

Department of the History of Science
Harvard University
Cambridge, Massachusetts

(Received April 23, 1968)