HISTORIA MATHEMATICA 5 (1978), 379-409

THE ORIGINS OF CAUCHY'S THEORY OF THE DERIVATIVE

BY JUDITH V. GRABINER,
CALIFORNIA STATE UNIVERSITY, DOMINGUEZ HILLS

SUMMARIES

It is well known that Cauchy was the first to define the derivative of a function in terms of a rigorous definition of limit. Even more important, he used his definitions to prove theorems about the derivative. We trace the historical background of the property of the derivative which Cauchy used as his definition and of the proof techniques Cauchy used. We focus on Cauchy's theorem that, for f(x) continuous on $[x_0, X]$,

$$(1) \quad \min_{[x_0, X]} f'(x) \le \frac{f(X) - f(x_0)}{X - x_0} \le \max_{[x_0, X]} f'(x) .$$

(Cauchy's statement and delta-epsilon proof of this theorem are reproduced as an Appendix to this article). We show how J.-L. Lagrange used what later became Cauchy's defining property of the derivative, and the associated proof techniques--though differently conceived and inadequately justified--to prove facts about derivatives, including Cauchy's theorem (1). We show, looking at the work of Euler and Ampère, where Lagrange got these ideas, how he developed and used them, and by what means they reached Cauchy. Finally, we see how Cauchy, recognizing what was essential in earlier work, clarified and improved what had been done, and for the first time placed the theory of derivatives on a firm mathematical foundation.

Il est bien connu que Cauchy définit le premier la dérivée d'une fonction en termes d'une définition rigoureuse de limite. Fait encore plus important, il employa ses définitions pour démontrer des théorèmes sur la dérivée. Nous retraçons les antécédents de la propriété de la dérivée que Cauchy employa pour définition et des techniques de preuves qu'il utilisa. Nous concentrons notre attention sur le théorème de Cauchy qui dit que pour toute fonction

0315-0860/78/0054-0379$02.00/0

f(x) continue sur $[x_0, X]$,

$$(1) \quad \min_{[x_0,\, X]} f'(x) \le \frac{f(X) - f(x_0)}{X - x_0} \le \max_{[x_0,\, X]} f'(x) .$$

(L'énoncé de Cauchy et la preuve en epsilon-delta de ce théorème sont repris en Appendice au présent article). Nous montrons comment J.- L. Lagrange, pour prouver divers faits à propos de la dérivée, y compris le théorème de Cauchy (1), employa ce qui plus tard deviendra la propriété définissante de la dérivée de Cauchy, et les techniques de preuves associées quoique différemment conçues et insuffisamment justifiées. Jetant un coup d'oeil aux oeuvres d'Euler et d'Ampère, nous verrons où Lagrange a pris ses idées, comment il les développa et les utilisa et par quels chemins elles atteignèrent Cauchy. Enfin, nous voyons comment Cauchy, discernant l'essentiel dans les travaux antérieurs, clarifia et améliora ce qui avait été fait et fit reposer pour la première fois la théorie des dérivées sur des fondements mathématiques solides.

INTRODUCTION

It is a commonplace that Augustin-Louis Cauchy gave the first generally acceptable account of the basic concepts of the calculus. After Cauchy, the calculus was no longer just a set of problem-solving techniques, widely applied but only intuitively understood. Of course, Cauchy's rigor was far from perfect, but it nevertheless set a new standard for nineteenth century analysis. After Cauchy, in large part because of the example he set, the calculus became a set of theorems, based on rigorous definitions. The wealth of results obtained by eighteenth-century mathematicians were justified in the nineteenth century by careful definitions and precise proofs. The "revolutionary" nature of Cauchy's foundations of the calculus has often been noted [Abel 1826; Klein 1926, 82-87; Freudenthal 1971]. In the present paper, we will treat one specific topic: Cauchy's theory of the derivative, and, particularly, its historical roots. The derivative is of course central to the calculus; one could claim that it is the most important of the concepts of the calculus. And it was Cauchy who, in 1823, gave the first rigorous theory of derivatives.

By a "theory of derivatives" I mean more than just a correct definition and some simple proofs. There was an impressive body of eighteenth-century results about derivatives, ranging from the product rule to Taylor's theorem with Lagrange remainder;

a satisfactory theory of derivatives would have to deduce these results rigorously from the definition. Also, there was a set of applications of the derivative: to extrema, tangents, contacts between curves, and so on; a theory of derivatives would have to prove the validity of these applications.

Cauchy first presented his definition of the derivative to the mathematical world in his *Leçons sur le calcul infinitésimal* of 1823. He defined the derivative $f'(x)$ of a continuous function $f(x)$ as the limit, when it exists, of the ratio $f(x+i) - f(x)/i$ as i went to zero. But it is not the mere definition of the derivative as the limit of the quotient of differences which constitutes Cauchy's achievement. Newton, after all, had described some of his results in terms of limits [Newton 1934, Scholium to Lemma XI]. Jean-le-Rond D'Alembert, under Newton's influence, had explicitly defined the differential quotient as the limit of the quotient of differences [1789, article "Différentiel." The definition used by D'Alembert was fairly common by the end of the eighteenth century. See Boyer 1949, Chapter VI.] The difference between Cauchy's work and that of men like D'Alembert lies in the understanding and the use of the definition. D'Alembert did not have what we now call a delta-epsilon translation of the limit-concept; as we shall see, Cauchy did. The only real use D'Alembert made of his definition was to illustrate the finding of a tangent to a parabola as the limit of secants [*op. cit.*]. In contrast, Cauchy's definition was the beginning of his task, not the end; his achievement was to produce an extended body of proved results about derivatives.

Our task in the present paper, though it will begin by isolating the origins of Cauchy's definition of the derivative, will go far beyond that. We shall trace the history of Cauchy's crucial theorems about derivatives and the associated proof-techniques, and shall discuss the history of attempts to give a "theory of derivatives." In particular, through looking at the work of Cauchy's major predecessors in the theory of derivatives, Euler, Lagrange, and Ampère--most especially Lagrange--we shall find the origin both of the property of the derivative which Cauchy used as his definition, and of the tools his predecessors left him to construct his proofs of the major properties and applications of the derivative.

CAUCHY'S DEFINITION AND HIS CRUCIAL THEOREM

We shall quote Cauchy's definition of derivative in full below, but first, in order to understand precisely what he meant by his definition, we must recall how he had defined the basic concepts of analysis on which his definition of derivative is based. Cauchy had defined "limit" in his celebrated *Cours d'analyse* of 1821. He wrote:

> When the successively attributed values of one variable indefinitely approach a fixed value, finishing by differing from that fixed value by as little as desired, that fixed value is called the *limit* of all the others. [Cauchy 1821, 19].

This definition is purely verbal, to be sure, but when Cauchy needed it for use in a proof, he often translated it into the language of inequalities. Sometimes, instead of so translating it, he left the job for the reader, but there are enough examples to demonstrate that Cauchy knew exactly how to make the translation. For instance, he interpreted the statement "the limit, as x goes to infinity, of $f(x+1) - f(x)$ is some finite number k" as follows:

> Designate by ε a number as small as desired. Since the increasing values of x will make the difference $f(x+1) - f(x)$ converge to the limit k, we can give to h a value sufficiently large so that, x being equal to or greater than h, the difference in question is included between $k-\varepsilon$ and $k+\varepsilon$. [1821, 54].

The epsilon notation was introduced into analysis by Cauchy. We will find a delta to go with the epsilon when we reach Cauchy's work on derivatives [1823]. (The theorem whose proof requires the passage just quoted is that, if as $x \to \infty$ lim $f(x+1)$ $f(x) = k$, then lim $f(x)/x = k$ also.)

Cauchy's definition of limit, with the delta-epsilon understanding that accompanied it, was the basis for the theory of convergent series he gave in the *Cours d'analyse* [1821, 114ff; still a good introduction to the subject]. The limit-concept was also the basis of Cauchy's theory of continuous functions [1821, 43ff; 378-80] and of the definite integral [1823, 122ff]. The "infinitely small quantity" so often discussed in eighteenth-century calculus was, for Cauchy, defined simply as a variable whose limit is zero [1821, 19]. And a function was continuous on an interval if, for all x on that interval, "the numerical [i.e., absolute] value of the difference $f(x+\alpha) - f(x)$ decreases indefinitely with that of α ... [That is,] an infinitely small increment in the variable produces always an infinitely small increment in the function itself" [1821, 43]. Note that Bolzano [1817] had independently given a similar definition. (In both cases, what was really being defined was uniform continuity.)

Now let us see precisely how Cauchy defined the derivative of a continuous function:

> If the function $y = f(x)$ is always continuous between two given bounds [his word is "*limites*"] of the variable x, and if we choose a value of the variable between these limits, than an infinitely

> small increment given to the variable will produce an infinitely small increment in the function itself. Therefore, if we set $\Delta x = i$, the two terms of the *ratio of the differences* $\Delta y/\Delta x = f(x+i) - f(x)/i$ will be infinitely small quantities. But, when the two terms indefinitely and simultaneously approach the limit zero, the ratio itself can converge toward another limit, which may be positive or negative. This limit, when it exists, has a determined value for each particular value of x; but it varies with x.... The form of the new function which serves as the limit of the ratio $f(x+i) - f(x)/i$ will depend only on the form of the proposed function $y = f(x)$. In order to indicate this dependence, we give the new function the name *derived function* [fonction derivée, our "derivative"], and we denote it, by means of an accent mark, by the notation y' or $f'(x)$. [1823, 22-23; his italics].

Both the name "fonction derivée" and the notation $f'(x)$ are due to Lagrange, whose influence on Cauchy will be discussed below. See [Lagrange 1797; 1813].

Cauchy's phrase "this limit, when it exists" exemplifies his attitude toward rigor. Perhaps his use of the phrase was motivated only by the behavior of known functions at isolated points, but the language was sufficiently general to open the whole question of the existence or non-existence of derivatives. And, though his definition of derivative, like that of limit, is verbal, we shall see immediately that he translated the definition into the algebra of inequalities for use in proofs.

Cauchy applied his definition to prove that the derivative had in fact the full range of properties that it was supposed to have. The crucial theorem he needed was this:

> (1) If $f(x)$ is continuous between $x = x_0$ and $x = X$, and if A is the minimum of $f'(x)$ on that interval while B is the maximum, then
>
> $$A \leq f(X) - f(x_0)/(X-x_0) \leq B.$$
>
> [1823, 44. Cauchy expressed "$\leq$" verbally, and consistently distinguished (verbally) between "$\leq$" and "$<$"].

We have reproduced Cauchy's statement and proof of this theorem as an appendix to this paper, since its history will be our major subject. For the central role of (1) in Cauchy's calculus, see, for instance [1823; 89, 123, 131, 151-2, 243].

In his proof, Cauchy for the first time translated his definition of derivative into the language of delta-epsilon inequalities. That is, given an ε, δ can be chosen in such a way that, for $|i| < \delta$, $f'(x) - \varepsilon < f(x+i) - f(x)/i < f'(x) + \varepsilon$. [For Cauchy's precise words, see the Appendix; the delta and epsilon are his.] To be sure, Cauchy assumed that his δ would work for all x on the given interval--an assumption equivalent to that of the uniform convergence of the derivative. But aside from this "oversight," Cauchy knew exactly what he meant by the statement "the derivative is the limit of the quotient of differences," and he was really the first mathematician in history to know this. Perhaps, one might want to except Bolzano, who knew in 1816 that $\phi(x+w) - \phi(x)/w = \phi'(x) + \Omega$, where Ω can be made as small as desired when w is small; however, this was a *property* of the derivative for Bolzano, not a definition. The property, as we shall see, can already be found in [Lagrange 1797]. For Bolzano, see [Stolz 1881, 264]. Bolzano's major work on derivatives, which may be found in [Bolzano 1930], was written in the 1830's and explicitly cites Cauchy's work. See, e.g., [Bolzano 1930, 94].

Since it is the property of the derivative expressed by Cauchy in delta-epsilon terms, and not his verbal definition, which is essential to his proofs and thus to his rigorous theory of derivatives, our principal task in this paper will be to trace the history of this property and its use in proofs about the derivative. We shall begin by stating the major result of our inquiry into the origins of Cauchy's theory of the derivative. By the end of the eighteenth century, Joseph-Louis Lagrange had given two crucial properties of $f'(x)$. First, he argued that

$$f(x+i) = f(x) + if'(x) + iV, \tag{2a}$$

where V goes to zero with i. Furthermore, for Lagrange, "V goes to zero with i" meant that, given any D [for "donnée"], i can be chosen sufficiently small so that V is between $-D$ and $+D$. Thus, given any D, Lagrange said, i can be found such that:

(2b) "$f(x+i) - f(x)$ lies between $i[f'(x)-D]$ and $i[f'(x)+D]$."
[Lagrange 1806, 87; cp. 1813, 77].

The kinship of (2b) with Cauchy's formulation is obvious.

We shall call (2a) the *Lagrange property of the derivative*, not only because Lagrange was the first to state it, but because he was the first to use properties (2a) and (2b), as they were later used by Cauchy and are still used today, to derive many of the results known in the eighteenth century about functions and their derivatives. For Lagrange, of course, (2a) was just one property of the derivative, not the definition as it was for Cauchy.

After translating the Lagrange property (2a) into the inequality property (2b), Lagrange himself became the first to apply the algebra of inequalities to proofs about derivatives. Lagrange developed a specific proof technique, designed to go with (2b), for establishing theorems about the behavior of derivatives on intervals. As we shall see, this technique is essentially that used by Cauchy in his proof of Theorem (1).

These resemblances between the work of Lagrange and the work of Cauchy are no coincidence. Cauchy knew Lagrange's books on the calculus well (see, e.g., [Cauchy 1829, 268; 1823, 9-10]). Moreover, [Valson 1868, 27] tells us that when Cauchy went on his first engineering job, [Lagrange 1797] was one of the four books he took along. In addition, in 1806, André-Marie Ampère used both (2a) and the associated Lagrangian proof technique, with copious references to Lagrange, in a paper about derivatives which proved a result analogous to (1) and to which Cauchy referred explicitly in stating Theorem (1) [Cauchy 1823, 44n].

Cauchy, then, got the Lagrange property from Lagrange. But where did Lagrange get it? Its origin lies in the history of Taylor series--in particular, in Euler's pioneering applications of the Taylor series to the calculus. We will discuss, in the next section of this paper, Euler's use of a property of the Taylor series from which the Lagrange property could be obtained. We will then describe how Lagrange developed and transformed Euler's suggestions to formulate the Lagrange property of the derivative. We will describe Lagrange's use of (2b) in proofs, and will show how Ampère adapted Lagrange's proof technique to prove a result closely akin to Theorem (1). Finally, we shall consider Cauchy's proof of Theorem (1), and evaluate his theory of derivatives in the light of the historical background we have described.

EULER'S CRITERION: INFINITE SERIES AND REMAINDERS

We shall begin our investigation by considering the work which probably inspired Lagrange not only to state the Lagrange property of the derivative, but also to exploit it at length. Leonhard Euler, in his *Institutiones calculi differentialis*, gave a criterion for when to use a finite number of terms of a power series, neglecting their remainder--that is, a criterion for the usefulness of power-series approximations. It was Euler's criterion which led Lagrange to the Lagrange property.

Euler explained his criterion in the following way. Given y, a function of x, and ω, a change in x. Then

$$\Delta y = P\omega + Q\omega^2 + R\omega^3 + \ldots .$$

"If the increment ω, which is added to the variable quantity, is very small, the terms $Q\omega^2$, $R\omega^3$ ..., also become very small, until $P\omega$ greatly exceeds the sum of *all the rest*." This is Euler's criterion. Essentially, $P\omega$ can be taken to stand for the whole series in all those computations "where the greatest accuracy is not needed." Euler added that "in many cases in which the calculus is applied in practice, this kind of considera tion is very fruitful" [1755, Section 122; my italics].

What Euler had in mind when he mentioned "cases to which the calculus is applied" is best illustrated by an application he made himself. He showed that if x is a relative maximum or minimum of y, then dy/dx is zero there. Suppose $y(x)$ is a relative maximum. Then

$$(3a)\qquad y(x) > y(x+\alpha) = y(x) + \alpha\, dy/dx + (1/2)\,\alpha^2 d^2y/dx^2 + \ldots$$

$$(3b)\qquad y(x) > y(x-\alpha) = y(x) - \alpha\, dy/dx + (1/2)\,\alpha^2 d^2y/dx^2 - \ldots$$

In each series, for α sufficiently small, the term in α will exceed all the rest; this means that the sign of the entire series of terms containing powers of α will have the sign of the term $\alpha\, dy/dx$ in (3a), and $-\alpha\, dy/dx$ in (3b). Thus the only way both inequalities (3a) and (3b) can be simultaneously satisfied is for dy/dx to be zero [1]; [Euler 1755, Secs. 253-4]. Of course, this argument requires that the function y is uniquely the sum of its Taylor series, a point which Euler took for granted.

A Taylor-series treatment of maxima and minima in terms of the signs of fluxions of various orders, with geometric justification, had been given by Maclaurin [Maclaurin 1742, Sections 261, 858-859]. However, Maclaurin did not base this on a general statement like what I have called "Euler's criterion." More important, Euler's derivation of the properties of maxima and minima was intended to be purely analytic, not geometric. He had in effect given an algebraic theory of maxima and minima, based on an approximation and thus based on the algebra of inequalities. This theory would have been an important innovation even if it had not influenced Lagrange. But it did influence Lagrange. It appealed to Lagrange because it was consistent with his general program of founding the calculus with no appeal to geometry or intuition, but based solely on "the algebraic analysis of finite quantities," since Lagrange viewed expanding $f(x+i)$ in a Taylor series as an algebraic process [1797, *passim*]. And Euler's work also fits in perfectly with the *specifics* of Lagrange's "algebraic" foundation for the calculus. Since Lagrange wanted to base his calculus on Taylor series, he would have been especially impressed by Euler's use of "Euler's criterion" in cases like that of extrema. Lagrange in fact seized on Euler's criterion and extended it far beyond Euler's few examples. He even tried to prove it [1813, 28-9]. Though Lagrange, in this proof, did not make a reference to Euler, the many similarities between the *Fonctions analytiques* and *Institutiones calculi differentialis* argue overwhelmingly for Euler's influence. See [Yushkevich 1954, *passim*]. Finally, Lagrange explicitly credits the *property* just described of maxima and minima (if not its derivation) to Euler [Lagrange 1813, Part 2, Chapter XI, § 51, p. 282]; he makes no reference to Maclaurin's treatment of this subject. As we shall explain, it was from Euler's criterion that Lagrange was led to what we have been calling the "Lagrange property" of the derivative.

LAGRANGE AND THE LAGRANGE PROPERTY OF THE DERIVATIVE

Obtaining what we have been calling the Lagrange property of the derivative was one of the first tasks Lagrange undertook in his *Fonctions analytiques*. Lagrange had begun this work on the foundations of the calculus by trying to "prove" that any function f(x) had a power series expansion of the form

$$f(x+i) = f(x) + ip + i^2q + \ldots,$$

except possibly, at some finite number of isolated points. Lagrange thought he had proved this [1813, 22-23]. By "fonction," Lagrange meant any "expression de calcul" into which the variable entered in any way. This definition is borrowed from Euler's *Introductio* [1748]. (Note that Euler used this definition in the *Introductio* because the work is solely the study of infinite analytic expressions; he elsewhere recognized and used a broader definition of function.) Given such an expansion, Lagrange also followed Euler in stating that there was some i small enough so that any term of the series, "abstraction being made of the sign," would exceed the sum of the remainder of the terms in the series. But Lagrange, unlike Euler, tried to prove this fact [1813, 28-9] [2].

Lagrange began his proof by treating f($x+i$) as the sum of two expressions, one depending on i, the other not:

$$f(x+i) = f(x) + iP$$

where P is a function of both x and i. Analogously defining Q by $P = p + iQ$, R by $Q = q + iR$, etc., Lagrange gave Euler's criterion in the following form:

(4) For i small enough, $f(x) > iP$, $iP > i^2Q$, etc.

Lagrange then appealed to the continuity of iP, iQ, ... to assert that i could be found sufficiently small for any *particular* one of the inequalities of (4) to hold. This, for Lagrange, proved Euler's criterion, since if "we can always give i a small enough value so any term of the series ... becomes greater than the sum of all the terms that follow," than "any value of i smaller than that one always satisfies the same condition" [1813, 29].

Euler himself had viewed his criterion as occasionally useful in justifying applications of the derivative. Lagrange, however, recognized the result (4) as fundamental, saying "[This result] is assumed in the differential and the fluxional calculus, and it is because of this that one can say that these calculuses are the most fruitful, especially in their application to problems of geometry and mechanics," [Lagrange 1813, 29]. This quotation is an extraordinarily important statement, both from Lagrange's point of view and from ours. Compare it with Euler's analogous remark in [1755, Section 122] quoted above, page 385. Compare also [Lagrange 1806, 101]. For Lagrange's own "applications," see page 404.

For Lagrange, (4) provided the answer to a major question he himself had raised in the Berlin Academy's prize competition of 1784 [Yushkevich 1971]: How could the differential and fluxional calculuses, with their somewhat shaky hypotheses, nevertheless obtain "so many true results?" I think the reasoning behind Lagrange's statement of 1797 was something like this: The differential and fluxional calculuses allow i to become "infinitely small" or to "vanish" or to "have zero as its limit." Whatever else these phrases may mean, they seem *at least* to require that i be a very small finite number--in particular, small enough so that $|iQ|$ can be made less than $|p|$. Thus any result of the differential or fluxional calculus which requires *no more than* making $|iQ|$ less than $|p|$ should, in Lagrange's view, follow from the truth of (4).

Whatever the quoted statement may mean, Lagrange in fact did justify his applications of the calculus by appealing to a Taylor-series form of (4). That is, after Lagrange had defined $f'(x)$ as the coefficient of i in the Taylor-series expansion of $f(x+i)$ [3], he translated the Euler criterion into the statement that, if $f(x+i) = f(x) + if'(x) + i^2/2\ f''(x) + \ldots$; then, $if'(x)$ exceeds, in absolute value, the remainder of the series $i^2/2\ f''(x) + i^3/6\ f'''(x) + \ldots.$ Lagrange said that this fact about the remainder is equivalent to the Lagrange property (2a) $f(x+i) = f(x) + if'(x) + iV$, where V is a function of x and i which goes to zero when i does [1806, 86-7]. (Compare [1813, 72, 77], where Lagrange gave this alternate form: $f(x+i) = f(x) + if'(x) + i^2Q$, where Q is finite.)

We, of course, recognize the "Lagrange property" as the crucial defining property of the derivative. Lagrange did not, but he did recognize the equivalence of Euler's criterion and the Lagrange property [4]. Thus, though Lagrange's Taylor-series approach to the calculus was incompatible with a *definition* of $f'(x)$ according to the Lagrange property, it did not prevent him from recognizing the fundamental importance of that property.

Furthermore, Lagrange quickly translated (2a) into inequalities, a step essential to Cauchy's rigorous inequality-proofs about derivatives. For V to go to zero when i does, said Lagrange, meant that some i could be found so that the corresponding value of V, "abstraction being made of the sign," would be less than any given quantity. And he followed this verbal statement with a beautiful treatment in terms of inequalities, deriving (2b).

> Let D be a given quantity which we can take as small as we please. We can then always give i a value small enough for the value of V to be included between the limits D and $-D$. Thus, since

$$f(x+i) - f(x) = i[f'(x) + V],$$

it follows that the quantity

$$f(x+i) - f(x)$$

will be included between these two quantities:

$$if'(x) \pm D . \qquad [1806, 87].$$

This is precisely the property Cauchy used in proving theorems about derivatives in his *Leçons sur le calcul infinitésimal* of 1823.

Since Cauchy knew Lagrange's works on the calculus, it seems quite likely that Cauchy's rigorous definition of the derivative of a function was based on Lagrange's use of (2a) and (2b). Lagrange's view of the facts expressed in (2a&b) was not, however, the same as Cauchy's. For Cauchy, (2b) is equivalent to the definition of $f'(x)$; for Lagrange, (2b) is merely *one* property of the derivative, and not the most essential one--though it is the most useful in applications.

Of course, as we have already said, there is more to the rigorous theory of the derivative than a mere definition. The mathematical value of Cauchy's definition stems not from its logical correctness alone, but from the proofs of the theorems to which Cauchy applied it. We shall now see that in proofs as in definition, Cauchy owed much to Lagrange's work on derivatives. Lagrange himself was the first to use the Lagrange property in proofs. Let us now document how Lagrange's use of (2a) and (2b), and of associated proof techniques, became the basis for Cauchy's rigorous theory of derivatives.

PROVING THEOREMS ABOUT DERIVATIVES

A. INTRODUCTION: CAUCHY'S THEOREMS AND THEIR SOURCES

As we have said, the crucial theorem for Cauchy's calculus is (1). The mean-value theorem for derivatives is, for Cauchy, an easy corollary of (1) and of the intermediate-value theorem for continuous functions: If $f'(x)$ is continuous between $x = x_0$ and $x = x_0 + h$, then there is a θ between 0 and 1 such that

(5) $$f(x_0 + h) - f(x_0)/h = f'(x_0 + \theta h).$$
[Cauchy 1823, 44-6].

We shall find the statement of both Theorem (1), which I shall call the "mean-value inequality," and its corollary, (5), the mean-value theorem, in the work of Lagrange and Ampère, who developed the techniques Cauchy needed to prove (1). In this section, we shall document this statement in detail, but first let us summarize what happened.

Lagrange used the "Lagrange property" of $f'(x)$ to prove a Lemma, which, in modern terms, states that a function with a positive (or negative) derivative on an interval is increasing (or decreasing) there. Lagrange then used this Lemma to derive the Lagrange remainder of the Taylor series, a result which yields (5) as a special case. In 1806, Ampère, following Lagrange's lead, used inequality-techniques to "prove" that $f'(x)$ satisfied the relation (1) in the following slightly different form: if f is always finite, and not always zero, between $x = a$ and $x = k$, and if $f(a) = A$, $f(k) = K$, then there is some value of x between a and k for which $f'(x) \leq K-A/k-a$, and another for which $K-A/k-a \leq f'(x)$. In proving this fact, which is essentially equivalent to (1), Ampère stated and used a specific identity in the algebra of inequalities. The same identity later appears in Cauchy's *Cours d'analyse* of 1821, and is used by Cauchy in his own proof of (1) in his *Calcul infinitésimal* of 1823. (The relationship between the problems treated in these papers has often been noted. See e.g., [Pringsheim 1909, 30ff]. But nobody has thoroughly analyzed the history of the proof techniques or the logical relationship between them). Thus Cauchy adapted the techniques of Lagrange and Ampère to build his own theory of derivatives.

B. LAGRANGE BRINGS INEQUALITY-PROOFS INTO THE CALCULUS

Lagrange's use of inequalities in the calculus was in the same spirit as his use of inequalities in algebraic approximations [Grabiner 1974]. For him, inequalities had nothing essential to do with the definition either of the derivative or of the sum of a series. Instead, they were used to get the finite approximations needed in applications and problem-solving. Even the remainder for a Taylor series was, for Lagrange, just a way of finding, in general terms, the error made when an infinite series is replaced by a finite approximation. And it was to find the Lagrange remainder that Lagrange first stated and proved the Lemma saying in effect that a function with a positive derivative on an interval is increasing there.

The method of proof Lagrange used to get this Lemma is one of the most important positive contributions he made to the rigorous basis of the calculus. In his *Leçons sur le calcul des fonctions*, Lagrange stated the Lemma thus:

> A function which is zero when the variable is zero will necessarily have, while the variable increases positively, finite values of the same sign as its derived function, or of opposed sign if the variable increases negatively, as long as the values of the derived function keep the same sign and do not become infinite. [1806, 86] [5]

This is the sort of theorem which is obvious on looking at a diagram. It is remarkable that Lagrange felt called upon to give a proof at all--but we must remember his strong preference for algebraic methods over geometric intuition. And the proof itself is even more remarkable. When we look at Lagrange's argument in detail, we see that it is the kind of proof later used by Cauchy. It is as close to a "delta-epsilon" proof as can be found in 18th-century calculus.

The proof begins with the assertion, for any function f, that (2a) holds: $f(x+i) = f(x) + i[f'(x) + V]$, where V is a function of x and i such that, when i becomes zero, so does V. As we have seen, Lagrange justified (2a) by his appeal to the position of $f'(x)$ in the Taylor-series expansion for $f(x)$. For us or for Cauchy, (2a) just states the defining property of $f'(x)$.

Lagrange then translated (2a) to say that, given D, i could be chosen sufficiently small so that (2b) held: that is, given D, i could be chosen sufficiently small so that $f(x+i) - f(x)$ would be included between $i[f'(x) \pm D]$ [1806, 87. In the version of the Lemma given in the *Fonctions analytiques*, Lagrange did not make entirely clear whether or not the quantity the *Calcul des fonctions* calls V had to be restricted to positive values; see esp. [1813, 76-77]. No confusion on this point remains in the *Calcul des fonctions*]. Lagrange clearly appreciated that what was important was the *absolute value* of the difference between $f(x+i) - f(x)$ and $if'(x)$. His realization in 1801 of the significance of absolute values was an important step forward in the use of inequalities in the calculus. While Cauchy was more explicit in his treatment of absolute values than was Lagrange, Lagrange in the *Calcul des fonctions* had already shown how to use absolute values correctly in proofs.

Now we are ready to give the details of Lagrange's proof of the Lemma. Applying the "sufficiently small" i of (2b) to various points in the interval over which f was defined, Lagrange said that

$$f(x+2i) - f(x+i) \text{ lies between } i[f'(x+i) \pm D];$$

$$f(x+3i) - f(x+2i) \text{ lies between } i[f'(x+2i) \pm D]; \text{ etc.}$$

The reader may have already noted that, once D was given, Lagrange assumed that the same i would always work, for any x in the given interval. We shall return to this point later.

Since $f'(x)$, $f'(x+i)$, ..., $f'(x+[n-1]i)$ all have the same sign by the hypothesis of the Lemma, $f(x+ni) - f(x)$ must lie between the quantities

$$\{if'(x) + i\,f'(x+i) + \ldots + if'(x+[n-1]i)\} \pm niD.$$

Lagrange expressed this conclusion by saying that the telescoping sum

$$f(x+i) - f(x) + f(x+2i) - f(x+i) + \ldots + f(x+ni) - f(x+[n-1]i)$$

"will have for limit the sum of the limits." (By "limit," Lagrange here meant "bound" [1806, 88]. Recall that the derivatives $f'(x+ki)$ are all assumed finite.) That is, $if'(x) + if'(x+i) + \ldots + if'(x + (n-1)i) - niD$ and $if'(x) + if'(x+i) +\ldots+if'(x + (n-1)i)+niD$ bound the sum.

Since D is arbitrary, said Lagrange, it can be taken as less than the value of $[f'(x) + f'(x+i) + \ldots + f'(x + [n-1]i)]/n$, "abstraction made of the sign." Lagrange gave no reason for being able to choose such a D. He probably had in mind the fact that D could be taken as less than the minimum value of $|f'(x)|$ between $x=0$ and $x=ni$. If D is so chosen, it will certainly be less in absolute value than $[f'(x) + f'(x+i) + \ldots + f'(x + [n-1]i)]/n$ [6]. However, the existence of a non-zero minimum for $|f'(x)|$ requires not only that $|f'(x)| > 0$, but that it be bounded away from zero. Lagrange's hypothesis would thus have to be strengthened for this choice of D to be possible. (The confusion between "greater than 0" and "bounded away from zero" is frequent in the eighteenth century; as we shall see, the distinction was first correctly made, in practice though not in words, by Cauchy.)

Suppose that D has been so chosen. From this, Lagrange concluded that $f(x+ni) - f(x)$ will lie between 0 and $2i[f'(x) + \ldots + f'(x + [n-1]i)]$. This is true, though he does not explain because for such D, $0 < [if'(x) + \ldots + f'[x + (n-1)i]] - niD$ and $[if'(x) + \ldots + if'[x + (n-1)i]] + niD < 2[if'(x) + \ldots + if'(x + [n-1]i)]$.

Lagrange then defined P to be the greatest positive or negative value of the n quantities $f'(x)$, $f'(x+i)$, ..., $f'(x + [n-1]i)$. For such P, then, Lagrange has proved that $f(x+ni) - f(x)$ lies between 0 and $2inP$.

Lagrange then explained what this last inclusion meant. Suppose we represent any function of z by $f(x+z) - f(x)$, and let $z=ni$. z has the same sign as i. If i is taken as small as desired, n can become as large as desired. Lagrange's proof then shows that $f(x+z) - f(x)$ lies between 0 and $2zP$. Thus the Lemma is proved to Lagrange's satisfaction.

Since Lagrange's proof of this Lemma is the best proof about derivatives by an eighteenth-century mathematician, and since, as we shall show, it influenced Cauchy, it will be worth while to specify in detail both its virtues and its deficiencies. There are impressive features distinguishing this proof from almost all previous proofs about properties of the derivative. Small positive quantities were treated by means of the algebra of inequalities, and a "delta-epsilon" calculation was undertaken. Lagrange did not finish his proof by saying "the sum of a finite number of positive infinitesimals is positive," nor by an appeal to a geometrical diagram, nor by building an impressive curtain of words; the proof is algebraic. Furthermore, he supplied a respectable amount of detail. He developed an extremely useful technique for going from a property of $f'(x)$ on the interval $[x, x+i]$ to a property of $f'(x)$ on the larger interval $[x, x+ni]$, by treating $f(x+ni) - f(x)$ as the telescoping sum $f(x+ni) - f(x + [n-1]i) + \ldots + f(x+i) - f(x)$; we shall see this procedure again in the work of Cauchy.

His proof has, however, several weaknesses. It assumed implicitly that $f'(x)$ was both bounded and bounded away from zero; Lagrange seems to have thought it enough for $f'(x)$ to be finite and never equal to zero. And there are even more serious objections. First, the proof is based on the Lagrange property of the derivative, (2a) $f(x+i) = f(x) + if'(x) + iV$, V vanishing with i, which Lagrange could prove only (and even then incorrectly) by using the full Taylor-series expansion of $f(x+i)$ in powers of i, which requires that *all* the derivatives be bounded. Cauchy overcame this objection by defining $f'(x)$ precisely so it would satisfy (2a). Second, Lagrange, assuming that one choice of i would make V small for all values of x in the given interval, confused convergence with uniform convergence. Cauchy reproduced this error.

C. THE LAGRANGE REMAINDER OF THE TAYLOR SERIES

Lagrange had obtained his Lemma so that he could find the "limits" (that is, bounds) of the remainder term of the Taylor series. Let us now look at his investigation of the Lagrange remainder itself. This result is of great interest in its own right. Furthermore, the Lagrange remainder for the case $n=1$ is important for us because it later became Cauchy's mean-value theorem for derivatives.

Lagrange derived the remainder from his Lemma in the following way. Let the maximum of $f'(x)$ on a given interval be $f'(q)$, the minimum, $f'(p)$. Define two auxiliary functions g and h, according to the equations

$$g'(i) = f'(x+i) - f'(p)$$

$$h'(i) = f'(q) - f'(x+i).$$

The definitions of g' and h' make $g'(i)$ and $h'(i)$ positive for x on the given interval, so that the Lemma can be applied. Going from these derivatives g' and h' to their "primitive functions," and assuming that $g(0) = h(0) = 0$, he obtained

$$g(i) = f(x+i) - f(x) - if'(p)$$

$$h(i) = if'(q) - f(x+i) + f(x),$$

which by the Lemma must be positive, as long as f' remains finite. (For Lagrange, "primitive functions" were the functions whose derivatives were being considered. Finding a primitive function, for Lagrange, was what most eighteenth century mathematicians would call "integration.") If $g'(i)$ and $h'(i)$ are positive, $f(x+i) - f(x) - if'(p) \geq 0$ and $f(x) - f(x+i) + if'(q) \geq 0$. Thus,

(6) $f(x) + if'(p) \leq f(x+i) \leq f(x) + if'(q)$,

which sets "limits"--that is, bounds--on the value of $f(x+i)$. This is Cauchy's theorem (1) [Lagrange 1806, 91; compare 1813, 80-81]. (Lagrange does not write "≤," but "<"; however, the context indicates that he means "≤.")

We must point out that the application of the Lemma to find (6) requires that the Lemma hold for the *weak* inequality $f(x+z) - f(x) \geq 0$, since all that can be claimed here is $g'(i) = f'(x+i) - f'(p) \geq 0$. Lagrange could have avoided this difficulty had he instead considered the functions

$$g'(i) = f'(x+i) - f'(p) + \varepsilon \text{ and}$$

$$h'(i) = f'(q) - f'(x+i) + \varepsilon,$$

which would yield

$$-\varepsilon i + f(x) + if'(p) < f(x+i) < f(x) + if'(q) + \varepsilon i.$$

Realizing that ε is arbitrary then would yield Lagrange's result, (6). As we shall see, Cauchy used precisely this procedure. Unfortunately, Lagrange used no special notation to distinguish between strict and weak inequalities, and did not appear to appreciate the full significance of this distinction.

Lagrange repeated the procedure we have exhibited in the case of (6) to obtain the nth order Lagrange remainder. In general, $f(x) + if'(x) + \ldots + i^u/u!\ f^{(u)}(p)$

$$\leq f(x+i)$$

$$\leq f(x) + if'(x) + \ldots + i^u/u!\ f^{(u)}(q),$$

where p and q are, respectively, the minimum and maximum points of the u^{th} derivative of f on the interval $[x, x+i]$. Lagrange concluded from this that there is a quantity X in the interval such that

$$f(x+i) = f(x) + if'(x) + \ldots + i^u/u!\ f^{(u)}(X).$$

[1806, 91-5; compare 1813, 80-5. Again, Lagrange did not use the notation "≤," but "<"]. This is what is now called Taylor's series with Lagrange remainder. The intermediate-value theorem for continuous functions is necessary to find X; Lagrange here stated that theorem without proof, as something obvious. The first rigorous proofs of this theorem are [Bolzano 1817] and, independently by quite a different technique, [Cauchy 1821, 378-80].

Lagrange's interest in approximations made something like the remainder term seem essential to him, if he was to use the Taylor series in the calculus. So, instead of neglecting this vital point, Lagrange used the inequality methods he had already exploited in algebraic approximations to derive the Lagrange remainder. Although he made errors in applying these methods, he deserves credit for introducing and helping develop a method of proof which was eventually to establish rigor in analysis.

What Lagrange did for the Taylor series, Cauchy was to do for the derivative; to reduce the question of its value to that of a sequence of inequalities which include it. What for Lagrange was just a stepping-stone to a first-order error estimate in the Taylor series became a defining property in the hands of Cauchy. Even before, and in a different way, it became a defining property in the hands of Ampère, who helped pass it on to Cauchy.

D. AMPÈRE'S PROOFS ABOUT THE DERIVATIVE

André-Marie Ampère is best known for his work in electricity. Nevertheless, he wrote a paper which was important in the history of the foundations of the calculus [Ampère 1806]. As a matter of fact, most of Ampère's early work was in mathematics, and it was on the basis of his mathematical work that he was made a member of the Institut de France in 1814, six years before his epoch-making work on electricity began. Ampère's 1806 paper on derivatives and Taylor series, entitled "Recherches sur quelques points de la théorie des fonctions derivées ...," was one reason for his mathematical eminence.

Ampère's 1806 paper is important for the historian of the calculus for several reasons. First, Ampère's paper includes features which historians have previously associated with nobody before Cauchy. One of Ampère's goals was to free the calculus, not only from the eighteenth-century concepts of limits, fluxions, and infinitesimals, but also from Lagrange's Taylor-series foundation. In particular, Ampère's paper gave inequality "proofs" about the basic properties of the derivative of a function. It also contained an inequality-definition for the derivative--unfortunately, not a satisfactory one.

Second, Ampère's paper relied heavily on the work of Lagrange. For explicit references to the *Fonctions analytiques*, see [Ampère 1806, 160, 169]; to the *Calcul des fonctions* [Ampère 1806, 163]; for references to the common doctrine of both Lagrange's books, see [Ampère 1806, 149, 165 *et passim*]. This reliance was not only in matters of notation, in the term "fonction derivée," and in the concern with the remainder term of the Taylor series, but also in Ampère's refinement of the proof technique Lagrange had used to prove that a function with a positive derivative on an interval was increasing there. Ampère's first major order of business in this paper was to prove Cauchy's "mean-value inequality" (1). Like Lagrange, Ampère had no way of proving that $f(x+i) = f(x) + if'(x) + iV$, where V vanishes with i. This inability did not stop Ampère from using this property, and the inequalities based on the property, to prove theorems about $f'(x)$.

Even if we know nothing else about the relationship between the work of Ampère and Lagrange, and the later work of Cauchy, the resemblances we have cited would be enough to suggest that Ampère's paper might have served as a link between Lagrange and Cauchy. But we have more evidence than just these resemblances. Cauchy knew Ampère personally, and had once been his student. Cauchy said in the "Introduction" to his *Cours d'analyse* that he had "profited several times from the observations of M. Ampère, as well as from the methods that he has developed in his lectures on analysis" [1821, vii-viii]. He more than once acknowledged Ampère's assistance in a general way [1821, vii; 1826, 10]. And Cauchy explicitly referred to Ampère's 1806 paper on the occasion of his own proof of (1) [1823, 44n; 1829, 268].

We have already observed that Cauchy knew Lagrange's *Fonctions analytiques*. It is therefore worth adding that Lagrange himself had called attention to Ampère's 1806 paper in the second edition of the *Fonctions analytiques* [1813], and acknowledged the kinship between Ampère's proof method and the one he had already given in his *Calcul des fonctions*, first published in 1801 [1813, 85]. The reference, made without any comment, is to Leçon IX of the *Calcul des fonctions*, which is in *Oeuvres* X, pp. 85-105. (The present author wishes to thank Lagrange for this reference to Ampère.)

Thus it is natural for us, in our search for Cauchy's sources, to examine Ampère's paper. The paper is, unfortunately, confusing to read and not very well organized. It has on occasion been misinterpreted as an attempt to prove that every continuous function is differentiable. The major source of this interpretation seems to be [Pringsheim 1909, 44]. Ampère himself is partly responsible for this error; in his paper he used the term "exist" to describe a derivative when he meant that it was finite and non-zero [1806, 149]. Given the prevailing view on the part of many historians about the lack of rigor and sophistication in analysis before Cauchy, together with Ampère's unusual use of the term "exist," the misinterpretation of Ampère's words is not surprising. Nevertheless, this error has led to an unfortunate neglect of Ampère's paper by historians of the calculus. Since Ampère's innovations helped Cauchy learn to prove theorems about derivatives, they deserve to be looked at again with care.

Let us begin our discussion by noting Ampère's purpose in this paper. Lagrange, in the preface to [Lagrange 1797] had, he thought, effectively discredited all the eighteenth-century definitions of the derivative; he then had given one of his own, basing the calculus on Taylor's theorem. In effect, Ampère asked himself, could not the derivative $f'(x)$ be defined indepedently not only of the concepts of limit and infinitesimal that Lagrange had supposedly refuted, but of Taylor's series as well? The definition of $f'(x)$ that Ampère wanted would have to specify $f'(x)$ uniquely, and define it in unexceptionable terms. Ampère found what he thought was a suitable property of $f'(x)$ in Lagrange's work on the Taylor series remainder; the property in question was (6), which Ampère adopted as the defining property of $f'(x)$:

(7) (Ampère's definition)
The derived function of $f(x)$ is a function of x such that $f(x+i) = f(x)/i$ is always included between two of the values that this derived function takes between x and $x+i$ whatever x and i may be.

Ampère introduced this by saying it was "a definition of the derived function $f'(x)$ which seems to me the most general and the most rigorous one possible" [1806, 156]. Cauchy's theorem (1) proves that $f'(x)$ satisfies Ampère's definition.

All the rigorous definitions of $f'(x)$ in the nineteenth century define it in terms of the ratio $f(x+i) - f(x)/i$ and the inequalities which that ratio must satisfy; Ampère was thus the first to give such a definition. However, his definition has some major deficiencies. First, it defines $f'(x)$ at the point x in terms of its values on the whole interval; thus $f'(x)$ must exist on an entire interval to be defined at one single point. This is much too restrictive (though not as restrictive as assuming $f(x)$ to have an entire Taylor series). Second, there is no reason to believe that any such $f'(x)$ exists at all. Third, it is not clear that $f'(x)$ is the only function which satisfies Ampère's defining criterion (7), though he did try to prove that $f'(x)$ was unique. In fact, if the derivative $f'(x)$ is continuous, no other continuous function has this property. Ampère's uniqueness proof, with a bit of modification, can be adapted to show this. Ampère himself, however, did not explicitly restrict himself to continuous functions; before Cauchy, the distinction between continuous functions and non-continuous functions would not have seemed important in this context.

Ampère believed that he had shown both the "existence"--in his sense of being finite and not always zero--and the uniqueness of $f'(x)$, and that his definition was therefore justified. We, however, are less interested in his definition than in the method of proof he used in showing that the derivative of a function had the property expressed in (7), for the proof method, based on the work of Lagrange, was adopted by Cauchy.

Ampère began his work confronted with a problem in logic. To prove that his definition of $f'(x)$ made sense, he first had to prove some facts about $f'(x)$. To do this, he had to have ways of characterizing $f'(x)$ other than his defining property. The way he *introduced* $f'(x)$ was as the value of the ratio $f(x+i) - f(x)/i$ "when $i=0$" [1806, 149]. The properties he actually *used* in his proofs were the properties Lagrange had used: that is, the inequalities satisfied by the ratio $f(x+i) - f(x)/i$ for arbitrarily small i. In particular, he assumed that $f'(x)$ had what we have called the Lagrange property: $f(x+i) - f(x)/i = f'(x) + iI$, where I vanishes with i [1806, 154-55]. (This is Ampère's notation. He did not explicitly mention Lagrange as the source of this particular property). Once he had proved that the function $f'(x)$ so characterized had the property expressed by (7), he turned around and used the property (7) to define $f'(x)$.

Let us look at how Ampère proved that a function with the Lagrange property also satisfied (7). Let $f(x)$ be defined on an interval from $x=a$ to $x=k$. Let $f(a)=A$, $f(k)=K$, $a \neq k$, $A \neq K$, and let $f(x)$ be finite [1806,151]. (These conditions mean that the function is well-behaved, the interval is not a point, and the function is not a constant). Now Ampère, using a proof technique like Lagrange's, undertook to show that there was some value of x on the interval such that $f'(x) \leq K-A/k-a$, and some other value of x such that $K-A/k-a \leq f'(x)$, which gives (7). (This result also implies, among other things, that the derivative cannot be zero or infinite on the whole interval, and thus, in Ampère's language, "exists").

Ampère's proof of (7) required an algebraic lemma about inequalities. Ampère's lemma was later stated and proved by Cauchy in his *Cours d'analyse*. It is therefore worth our while to give it a formal statement:

(8) (Ampère's lemma)
In a given interval $[a,k]$, define $b,c,d,e...$ such that $a < b < c \ldots < e < h < k$, and define $B, C \ldots H$ such that $f(b) = B$, $f(c) = C$, ... $f(h) = H$. Now consider the fractions $B-A/b-a$, $C-B/c-b$, ... $K-H/k-h$. Among these fractions, we can always find a pair such that one of the pair will be greater than $K-A/k-a$, while the other will be less.

This is not a quotation; Ampère has intertwined the statement of this lemma so closely with his proof of (7) that it is impossible to disentangle the lemma enough to quote Ampère's statement of it. See [1806, 151-154]. Let us indicate how Ampère proved it. If, for instance, $a < e < k$, $K-A/k-a$ lies between $K-E/k-e$ and $E-A/e-a$. For since $K-E/k-e - K-A/k-a = aE-Ae+Ak-ak+eK-Ek/(k-e)(k-a)$, and $K-A/k-a - E-A/e-a = aE-Ae+Ak-aK+eK-Ek/(k-e)(e-a)$, we have two fractions with the same numerator and with positive denominators; both fractions must have the same sign. If both are positive together, $K-E/k-e > K-A/k-a > E-A/e-a$. If both are negative together, we have $K-E/k-e < K-A/k-a < E-A/e-a$. We might go on to formalize this proof by induction on the number of fractions; Ampère examined several cases and immediately concluded the general result.

This lemma is almost identical with the inequality-result Cauchy used in his proof of (1) [1821, Note II, Theorem XII, p. 368]. See the Appendix for our quotation of Cauchy's result. Cauchy did not acknowledge Ampère when giving his own inequality result. He may not have been conscious of the relationship. See [Freudenthal 1971] for a discussion of Cauchy's way of working.

Once he had proved this Lemma, Ampère specified that $b-a = c-b = \ldots = k-h = i$; thus, the Lemma showed that there was some x on the interval $[a,k]$ such that $f(x+i) - f(x)/i$ was less than $K-A/k-a$, and another x such that $f(x+i) - f(x)/i$ was greater than $K-A/k-a$. Now Ampère appealed to the Lagrange property of the derivative to go from this result about finite differences to the corresponding result about $f'(x)$.

> Since $f(x+i) - f(x)/i$ becomes equal to $f'(x)$ when $i=0$, it can be represented in general by $f'(x) + I$, where I is a function of x and i which vanishes with i, and which, therefore, can become as small as desired by taking i sufficiently small. [1806, 154-155]

Compare also Lagrange's characterization of the function iP going to zero with i [1813, 28-9] in the *Fonctions analytiques*. Thus, by taking i sufficiently small, Ampère concluded that he could find some x on $[a,k]$ such that

$$f'(x) \leq K-A/k-a,$$

and, for some other value of x on the interval,

$$f'(x) \geq K-A/k-a.$$

(Ampère expressed these inequalities verbally, saying "plus grande," "plus petite," but said specifically that he meant "greater than or equal to," or "less than or equal to" [1806, 152]. I have, additionally, substituted the notation $[a,k]$ for his verbal description). This completes Ampère's proof of (7). I have tried to isolate the theorem we are interested in. However, part of the difficulty in reading Ampère's paper is that the proof we have just gone through is embedded in the proof of the theorem that $f'(x)$ cannot be zero or infinite on the whole interval. Of course (7) does imply this.

The key steps in the proof are the perfectly valid lemma on fractions, and the passage from the inequalities valid in the case of the ratio of finite differences $f(x+i) - f(x)/i$ to the inequalities stated for the case of the derivative $f'(x)$. The second of these steps is valid *if* the convergence of the ratio $f(x+i) - f(x)/i$ to its limit is uniform. But Ampère had no more reason to assume that he could find a value of i "sufficiently small" to work for all x in the interval than Lagrange had had before him, or Cauchy would have later.

Ampère's paper is important for us because it transmitted Lagrange's proof methods, together with a new and useful lemma on fractions, to Cauchy. It is also striking that in this paper, Ampère advocated defining the derivative f'(x) uniquely and unexceptionably: *not* by the verbal limit concept used by Newton and D'Alembert but rejected by Lagrange, *not* in terms of the Taylor series, but by an inequality that f'(x), and f'(x) alone, could satisfy. However, the choice of the inequalities which were appropriate to support the whole logical structure of the differential calculus was not made by Ampère, but by Cauchy.

Ampère had taken the Lagrange property of the derivative, and noticed how Lagrange had used it in the *Calcul des fonctions* to derive Taylor's theorem with remainder. His own proof methods, Ampère argued, could serve as a way of providing a theory of derivatives, "freed not only from the consideration of infinitesimals, but also from that of the formula of Taylor" [1806, 162]. For, since Ampère had shown that the derivative f'(x) satisfied his defining inequality (7) *if* it had the Lagrange property, he then--incorrectly--felt himself justified in using the Lagrange property to deduce further theorems--including Taylor's. Ampère in effect assumed the equivalence of his definition with the Lagrange property. If Cauchy was going to use the Ampère-Lagrange proof methods, which rest on the Lagrange property of the derivative, Cauchy's own definition of f'(x) would somehow have to justify the use of that property. Cauchy's definition was designed to do just that.

CAUCHY'S THEORY OF DERIVATIVES: PROVING BASIC THEOREMS

Cauchy's rigorous proofs about f'(x) are vastly more than just the culmination of a development of the work of Lagrange and Ampère. Still, viewing Cauchy's proofs in the light of the Lagrange-Ampère work helps us to understand why they took the form they did. Cauchy used his definition of limit to define the derivative in such a way that it would have the Lagrange property. He used this property of the derivative--now, for the first time, justified by a definition--in order to prove the mean-value inequality, (1). And he used the Lagrange-Ampère proof technique in this and several other proofs. Cauchy's proofs begin by translating his verbal definitions into the language of algebraic inequalities. The proofs then proceed, as did Ampère's and Lagrange's, through the use of algebraic techniques. But the techniques are no longer ad hoc. Cauchy's basic concepts are designed, whether consciously or not, to support the proof methods worked out by Lagrange and Ampère.

We have reproduced Cauchy's proof of Theorem (1) in the Appendix; we shall now discuss its major features. [The reader may wish to read Cauchy's proof before proceeding.] Cauchy stated the theorem as follows:

> If, f(x) being continuous between the limits $x=x_0$, $x=X$, we designate by A the smallest, and by B the largest, value that the derived function f'(x) receives in the interval, the ratio of the finite differences $f(X) - f(x_0)/X-x_0$ will necessarily be included between A and B. [1823, 44] [7]

Cauchy, like Lagrange and Ampère, used (2b), which is a translation into inequalities of the Lagrange property of the derivative, to prove this theorem. But for Cauchy, the procedure was justified by his own definition of f'(x) as a limit. Cauchy's proof is technically like Ampère's but much easier to follow. Here is a shortened version; the δ-ε notation is Cauchy's, introduced in this proof for the first time in history.

Given $\varepsilon > 0$, Cauchy said, we can choose δ such that

(9) $$f'(x) - \varepsilon < f(x+i) - f(x)/i < f'(x) + \varepsilon, \text{ if } |i| < \delta.$$

Statement (9) is valid since it just translates Cauchy's definition of the derivative into an algebraic inequality--a justification immeasurably superior to those given by Lagrange and Ampère.

We must note again that Cauchy took his definition of f'(x) for a particular x and applied it to the whole interval; he assumed that given an ε, he could find a δ that worked for *every* x in the interval. This assumes that f'(x) is the uniform limit of the quotients $f(x+i) - f(x)/i$ in the interval, a confusion which we have already found in the work of Ampère and Lagrange. The confusion arises from not precisely specifying on what the variable δ depends.

Once having made the algebraic statement (9), Cauchy went on to interpose n-1 new values of the variable x, namely, $x_1, x_2, \ldots, x_{n-1}$, between x_0 and X, such that (x_1-x_0), (x_2-x_1), $\ldots(X-x_{n-1})$, are all less than δ. (This differs from what Lagrange and Ampère did; the subintervals they had used were all equal.) Now applying the property (9) to each subinterval, we have

(10)
$$f'(x_0) - \varepsilon < f(x_1) - f(x_0)/x_1-x_0 < f'(x_0) + \varepsilon$$
$$f'(x_1) - \varepsilon < f(x_2) - f(x_1)/x_2-x_1 < f'(x_1) + \varepsilon$$
$$\ldots$$
$$f'(x_{n-1}) - \varepsilon < f(X) - f(x_{n-1})/X-x_{n-1} < f'(x_{n-1}) + \varepsilon.$$

If A and B are the minimum and maximum values of f'(x) on the given interval, then each of the fractions in (10) will be greater than $A-\varepsilon$, and less than $B+\varepsilon$.

Now Cauchy applied his version of Ampère's lemma (8) on fractions to the fractions $f(x_1) - f(x_0)/x_1-x_0, \ldots,$ $f(X) - f(x_{n-1})/X-x_{n-1}$, all of which have positive denominators. (Note that Cauchy did not refer to Ampère, but to his own result in the *Cours d'analyse*, Note II [1821, 27]). He combined this result with the telescoping property of the sum $f(X) - f(x_{n-1}) + \ldots + f(x_1) - f(x_0)$, used by Lagrange. These two results led him from (10) to the inequality

$$A - \varepsilon < f(X) - f(x_0)/X-x_0 < B + \varepsilon.$$

But since this is true no matter how small ε is, Cauchy concluded that

$$A \leq f(X) - f(x_0)/X-x_0 \leq B.$$

This completes Cauchy's proof of (1).

There are many differences between Cauchy's proof and Ampère's, besides the crucial fact that Cauchy defined $f'(x)$ to justify the proof procedure. First, in notation: using the delta instead of saying "a value of *i*" makes it much easier to follow the proof; similarly, the index notation for the values of the variable helps the reader. Much more important are the conceptual differences. Cauchy made his hypotheses explicit. The proof is crystal clear. And he understood the difference between "≤," "<," and "bounded away from," as is shown in the last lines of the proof, where he skillfully used epsilons to indicate that certain functions were bounded away from their limiting values. As we contrast Cauchy's proof of this theorem with the proofs given by Lagrange and Ampère, we are especially struck by Cauchy's ability to cull precise concepts from work written with ill-defined and hazy ideas, and to present instead proofs which are models of clarity both in the notation and the ideas.

One consequence of Cauchy's Theorem (1) is a corollary, (5), the mean-value theorem for derivatives. Cauchy derived (5) in the form

$$f(x+h) - f(x)/h = f'(x+\theta h), \quad 0 \leq \theta \leq 1,$$

using the intermediate-value theorem for continuous functions which he had proved in his *Cours d'analyse* [1821]. This result (5) is of importance to us not only for Cauchy's rigorous proof, but also for its applications. Cauchy, like Lagrange, obtained most of his applications of the derivative from properties of $f'(x)$ such as the mean-value theorem (5), and its higher-order analogue, Taylor's series with Lagrange remainder:

(11) $f(x+h) = f(x) + hf'(x) + \ldots + h^n/n!\, f^{(n)}(x+\theta h), \; 0 \leq \theta \leq 1.$

(For instance, see [Cauchy 1823, 88-9, 217]. The notation is

Cauchy's, save for "≤" which he expressed verbally). When expressions like (5) and (11) appear in Cauchy's work, the arguments used are often descendents of arguments given by Lagrange. (Compare [Lagrange 1813, 193-6] and [Cauchy 1826, 115-116]; [Lagrange 1813, 183] and [Cauchy 1826, 77-80]; [Lagrange 1813, 258-267] and [Cauchy 1823, 47-49, 221-222]. (Bolzano also adapted some of Lagrange's arguments of this type. See, for instance, [Bolzano 1930, 155ff]. For a specific acknowledgement by Bolzano of Lagrange's prior use of such techniques in [1797] and [1806], see [Bolzano 1930, 170]). Once one has the mean-value theorem and the Lagrange remainder, it is possible to justify most of the common applications of the calculus to problems of geometry and extrema. (For instance, consider the problem of extrema. For Lagrange, see [1813, 233-7]. For Cauchy, see [1823, 88-92]. And for a modern treatment using Lagrange's procedures, see [Widder 1947, 26-7, 99]). Cauchy thus was able to use his theory of the derivative to make rigorous many of the major results and applications of eighteenth-century calculus.

CONCLUSION

For the task of rigorizing analysis, Cauchy received many suggestions from the work of his predecessors. In the case of the theory and applications of the derivative, he found much of the work done for him. And it had been done chiefly by Lagrange. Starting with what we have been calling the Lagrange property of the derivative, Lagrange had been the first to apply what we since Cauchy recognize as delta-epsilon techniques to the calculus. Cauchy, by turning this property into a definition and by using his own theories of limit and continuity, was able to adapt Lagrange's techniques into his own rigorous proofs about the concepts of the calculus. Basing his work logically on a consistent set of definitions, Cauchy proved the major results and justified the major applications of the differential calculus. There was, of course, more to be done, and it was not all done by Cauchy [8]. But we now have come to understand how Cauchy created his theory of derivatives. He did not create it out of nothing, but by extending and transforming the work of his predecessors.

Yet all this should not diminish our high opinion of Cauchy's work. Cauchy's proofs rest on a clear and rigorous definition of both limit and derivative. Lagrange and Ampère, on the other hand, really could not justify the proof techniques they used. Moreover, the clarity of Cauchy's proofs, and the logically connected system of theorems which constitutes his rigorous calculus, should make his achievement sufficiently great by any mortal standards. It is by understanding the roots of his achievements in the work of others that we can obtain a clearer understanding of the nature of these achievements. In the theory of derivatives, his accomplishment was not so much in inventing wholly new techniques as in clearly understanding the meaning of a number of existing, previously ad hoc techniques. Cauchy took what was known and what had been done, but he used it and improved upon it to build something new. That new creation, the first rigorous theory of derivatives and their applications, was the central achievement of Cauchy's rigorization of the calculus.

NOTES

1. Similarly, since the α^2 term can be made to exceed the sum of all which follow it, Euler argued that if x is a relative maximum, d^2y/dx^2 must be negative; if x is a minimum, the second derivative must be positive. He added that similar considerations applied to higher-order examples [1755, section 255]. The use of these considerations in Euler's treatment of maxima and minima was highlighted by A. P. Yushkevich [1954]. Yushkevich also pointed out that there was a kinship between this work of Euler's and Lagrange's treatment of maxima and minima using the Taylor series remainder. We shall give details of Lagrange's theory of extrema below, pp. 393-395.

2. L. F. A. Arbogast had earlier (1789) tried to prove this too, though in a very different way than Lagrange. Arbogast's method was to choose i so that each term of the series was more than twice the following term. See [Zimmermann 1934, 47-8]. The conclusion, once i is so chosen, follows from the term-by-term comparison with the geometric series $\Sigma\ 1/2^k$. On the use of the comparison with this series to insure the good behavior of infinite series, see [Newton 1964, 24]; also, Euclid, *Elements*, X, 1. For evidence that Lagrange knew Arbogast's unpublished memoir, see his own statement in the *Fonctions analytiques* [1797, 5; 1813, 19].

3. Lagrange defined $f'(x)$ as the coefficient of i in the Taylor-series expansion for $f(x+i)$, $f''(x)$ as the coefficient of i in the Taylor-series expansion for $f'(x+i)$, etc. He then gave a formal proof that $f^{(k)}(x)/k!$ is the coefficient of i^k in the Taylor expansion of $f(x+i)$.

4. In his [1797], Lagrange derived the Lagrange property from Euler's criterion. In the *Leçons sur le calcul des fonctions* [1801] [Second edition 1806], he derived the Lagrange property of the derivative directly from the existence of the Taylor series of a function, and then proved Euler's criterion as a corollary of Taylor's theorem with Lagrange remainder; see [1806, 100-101]. The earlier derivation supports our

conclusion that Euler's criterion led Lagrange to the crucial property of the derivative, while the later one shows that he recognized the equivalence of the two properties, provided that the function f(x) is represented by its Taylor series.

5. Compare [1813, 78-80]: "If a prime function of x, like f'(x), is always positive for all values of x from $x=a$ to $x=b$, b greater than a, then the difference of primitive functions corresponding to these two values of x, that is, f(b) - f(a), will necessarily be positive." We choose to discuss the *Calcul des fonctions* version because the proof is better, specifically in the deriving of the relevant inequalities.

6. Lagrange had elsewhere given arguments based on calculating quantities like D on the basis of maximum or minimum properties of f'(x), so this explanation is likely. See, for instance, the argument about P immediately following. Alternatively, Lagrange might have considered D to depend on i and n, in which case D could be calculated. If this were his reason, the rest of the proof would be invalid, since i, and therefore n, must be chosen *after* D. Bolzano, incidentally, believed that this latter had been Lagrange's rationale, and had criticized Lagrange's proof on these grounds [1817, 12].

7. Cauchy said "comprise" for what I have translated as "included"; he meant "$\leq$." For strict inequality, he said "renfermé," which I have translated as "lying between." Lagrange did not make this distinction clear in his proofs; Ampère at least tried to. I use $|i|$ for Cauchy's "numerical value."

8. Cauchy also tried to rigorize the integral calculus, by defining the integral as the limit of sums and by proving what we now call the Fundamental Theorem of Calculus. His work on integrals was considerably improved upon by Riemann [Hawkins 1970, 9-12, 17-20]. Cauchy also did not distinguish at first between convergence and uniform convergence, a distinction first made in the 1840's. Weierstrass and his school completed the job of making analysis rigorous. See, for instance, [Boyer 1949, Chapter VII].

APPENDIX

THE FIRST RIGOROUS PROOF ABOUT DERIVATIVES

Cauchy [1823], Leçon 7, *Oeuvres* (2) IV, pp. 44-45.

Theorem. If the function $f(x)$ is continuous between the limits[A1] $x=x_0$, $x=X$, and if we let A be the smallest, B the largest, value of the derivative $f'(x)$ in that interval, the ratio of the finite differences

(4) $$f(X) - f(x_0)/X-x_0$$

must be included [A2] between A and B.

Proof. Let δ, ε be two very small numbers; the first is chosen so that, for all numerical [i.e., absolute] values of i less than δ, and for any value of x included between the limits x_0, X, the ratio

$$f(x+i) - f(x)/i$$

will always be greater than $f'(x) - \varepsilon$[A3], and less than $f'(x) + \varepsilon$. If we interpose n-1 new values of the variable x between the limits x_0, X, that is

$$x_1,\ x_2,\ \ldots,\ x_{n-1},$$

so that the difference $X-x_0$ is divided into elements

$$x_1-x_0,\ x_2-x_1,\ \ldots,\ X-x_{n-1},$$

which all have the same sign and which have numerical values less than δ; then, since of the fractions

(5) $f(x_1)-f(x_0)/x_1-x_0$, $f(x_2)-f(x_1)/x_2-x_1,\ldots,f(X)-f(x_{n-1})/X-x_{n-1}$.

the first will be included between the limits $f'(x_0) - \varepsilon$, $f'(x_0) + \varepsilon$, the second between the limits $f'(x_1) - \varepsilon$, $f'(x_1) + \varepsilon$, ..., etc., each of the fractions will be greater than $A-\varepsilon$, and less than $B+\varepsilon$. Moreover, since the fractions (5) have denominators of the same sign, if we divide the sum of their numerators by the sum of their denominators, we obtain a *mean* fraction, that is, one included between the smallest and the largest of those under consideration [see *Analyse algébrique*, Note II, Theorem XII].

[Note: Here Cauchy was referring to this theorem: "If b, b', b''... are n quantities with the same sign, and if a, a', a''... are any n quantities, we have

$$\frac{a+a'+a''+\ \ldots}{b+b'+b''+\ \ldots} = M(a/b,\ a'/b',\ a''/b'',\ \ldots\ )."$$

He had defined a *mean* of *c*, *c'*, *c''*, ..., *M*(*c*, *c'*, *c''*, ...), as "a new quantity included between the smallest and the largest of those under consideration." *Cours d'analyse*, in *Oeuvres* (2) III, p. 27.]

The expression (4), with which that mean coincides, will thus itself lie between the limits $A-\varepsilon$, $B+\varepsilon$, and since this conclusion holds no matter how small ε may be, we can conclude that the expression (4) will be included between *A* and *B*.

NOTES TO THE APPENDIX

A1. Cauchy uses "limites" for "bounds" or "endpoints."

A2. I have translated Cauchy's "comprise" as "included," and his "renfermé" as "lying between." The context of the proof makes clear that *c* "included" between a and b means $a \leq c \leq b$, and that *c* "lies between" *a* and *b* means $a < c < b$.

A3. Cauchy's *Oeuvres* has $f(x) - \varepsilon$, a misprint.

REFERENCES

Abel, Niels Hendrik 1826 Recherches sur la série $1 + m/1\ x + m(m-1)/1.2\ x^2 + m(m-1)(m-2)/1.2.3\ x^3 + \ldots$, *Crelles Journal* Bd. 1; reprinted in *Oeuvres completes*, ed. L. Sylow & S. Lie, Cristiana, 1881, pp. 221-250

Ampère, André-Marie 1806 Recherches sur quelques points de la théorie des fonctions derivées qui conduisant à une nouvelle démonstration de la série de Taylor, et à l'expression finie des terms qu'on neglige lorsqu'on arrête cette série à une terme quelconque *Journal de l'école polytechnique* 13e Cahier, Tome VI, pp. 148-181.

Bolzano, Bernhard 1817 *Rein analytischer Beweis des Lehrsatzes dass zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewaehren, wenigstens eine reele Wurzel der Gleichung liege* Prag. Reprinted in Ostwalds Klassiker, Leipzig, 1905

——— 1930 *Functionenlehre, Band I, Bernard Bolzanos Schriften* ed. Karel Rychlik Prague From manuscripts dating from the 1830's

Boyer, Carl B 1949 *The Concepts of the Calculus* New York *Reprinted (1959) as The History of the Calculus and Its Conceptual Development (Dover)*

Cauchy, A-L 1821 *Cours d'analyse de l'Ecole royale polytechnique 1re partie:* Analyse algébrique [all published] Paris Reprinted in *Oeuvres d'Augustin Cauchy* (2), tome 3

——— 1823 *Résumé des leçons données à l'école royale polytechnique sur le Calcul infinitésimal, Tome Premier* [all published] Paris Reprinted in *Oeuvres* (2), tome 4, 5-261

——— 1826 *Leçons sur les applications du calcul infinitésimal à la géometrie* Paris 1826-1828 Reprinted in *Oeuvres* (2), Tome 5

——— 1829 *Leçons sur le calcul différentiel* Paris Reprinted in *Oeuvres* (2), Tome 4, 263-609

D'Alembert, Jean-le-Rond 1789 *Dictionnaire encyclopédique des mathématiques* Paris A selection of the mathematical articles from the Diderot-D'Alembert *Encyclopédie*

Euler, Leonhard 1748 *Introductio in analysin infinitorum* Lausanne. Reprinted in OPERA (1), volumes VIII-IX

——— 1755 *Institutiones calculi differentialis* St. Petersburg Reprinted in *Opera* (1), volume X

Freudenthal, Hans 1971 Cauchy *Dictionary of Scientific Biography* 3, 131-148

Grabiner, Judith V 1974 Is Mathematical Truth Time-Dependent? *American Mathematical Monthly* 81, 354-365

Hawkins, Thomas 1970 *Lebesgue's Theory of Integration: Its Origins and Development* Madison, Wisconsin

Klein, Felix 1926 *Vorlesungen über die Entwicklung der Mathematik im 19ten Jahrhundert* Berlin 1926-1927 Reprinted Chelsea 1967

Lagrange, Joseph-Louis 1797 *Théorie des fonctions analytiques* Paris (first edition)

——— 1806 *Leçons sur le calcul des fonctions, nouvelle édition* Paris Reprinted in *Oeuvres de Lagrange* X (page citations from *Oeuvres*)

——— 1813 *Théorie des fonctions analytiques, nouvelle edition* Paris Reprinted in *Oeuvres* IX (page citations from *Oeuvres*)

Maclaurin, Colin 1742 *Treatise of Fluxions* 2 volumes Edinburgh

Newton, Isaac 1934 *Mathematical Principles of Natural Philosophy* tr. Andrew Motte, revised by Florian Cajori, Berkeley [translation of the third Latin edition, 1726]

——— 1964 *Of Analysis by equations of an infinite number of terms* tr. J. Stewart, 1745, from manuscript of 1669; reprinted in D. T. Whiteside, ed., *Mathematical Works of Isaac Newton* vol. 1 New York & London

Pringsheim, A & Molk, J 1909 Principes fondamentaux de la théorie des fonctions *Encyclopédie des sciences mathématiques pures et appliquées* Tom. 2, vol. 1 Fasc. 1

Stolz, O 1881 B. Bolzanos Bedeutung in der Geschichte der Infinitesimalrechnung *Math. Annalen* 18, 255-279

Valson, C-A 1868 *La vie et les travaux du Baron Cauchy* Paris

Widder, David 1947 *Advanced Calculus* New York

Yushkevich, A P 1954 Euler und Lagrange über die Grundlagen der Analysis *Sammelband der zu ehren des 230 Geburtstages Leonhard Eulers* Deutschen Akademie der Wissenschaft zu

Berlin, 224-244
——— 1971 Lazare Carnot and the Competition of the Berlin Academy in 1786 on the Mathematical Theory of the Infinite, in C. C. Gillispie *Lazare Carnot Savant* Princeton
Zimmerman, K 1934 *Arbogast als Mathematiker und Historiker der Mathematik* Heidelberg

ACKNOWLEDGEMENT

An earlier version of this paper was circulated at the Smithsonian Conference on the History of Mathematics, October, 1976. The author acknowledges the assistance of a grant from the American Council of Learned Societies, 1971-1972.