CHAPTER 19

JOSEPH LOUIS LAGRANGE, *THÉORIE DES FONCTIONS ANALYTIQUES*, FIRST EDITION (1797)

Craig G. Fraser

In this volume, based upon his first teaching at the *Ecole Polytechnique*, Lagrange both popularised and extended his view that the differential and integral calculus could be based solely on assuming the Taylor expansion of a function in an infinite power series and on algebraic manipulations thereafter. He also made some applications to problems in geometry and mechanics.

First publication. Paris: L'Imprimerie de la République, 1797. 277 pages.

Reprint. As *Journal de l'Ecole Polytechnique, cahier 9, 3* (1801), 1–277.

Later editions. 2nd 1813, 3rd 1847. Both Paris: Bachelier. Also as Lagrange *Oeuvres*, vol. 9, Paris: Gauthiers–Villars, 1881.

Portuguese translation. *Theorica das funções analyticas* (trans. M.J. Noguiera da Gama), 2 vols., Lisbon: J.C.P. da Silva, 1798.

German translations. 1) Of 1st ed.: *Theorie der analytischen Functionen* (trans. J.P. Gruson), 2 vols., Berlin: Lagarde, 1798–1799. 2) Of 2nd ed.: *Theorie der analytischen Functionen* (trans. A.L. Crelle), Berlin: Reimer, 1823.

Related articles: Leibniz (§4), MacLaurin (§10), Euler (§12–§14), Lagrange on mechanics (§16), Lacroix (§20), Cauchy on real-variable analysis (§25).

1 INTRODUCTION

At the end of the 18th century Joseph Louis Lagrange (1736–1813) published a book in which he developed a systematic foundation of the calculus, his *Théorie des fonctions analytiques* (1797). Parts of it were further developed in his *Leçons sur le calcul des functions* (1801; revised edition 1806).

Landmark Writings in Western Mathematics, 1640–1940
I. Grattan-Guinness (Editor)
© 2005 Elsevier B.V. All rights reserved.

By 1790 a critical attitude had developed both within mathematics and within general scientific culture. As early as 1734 Bishop George Berkeley in his work *The Analyst* had called attention to what he perceived as logical weaknesses in the reasonings of the calculus arising from the employment of infinitely small quantities (§8). Although his critique was somewhat lacking in mathematical cogency, it at least stimulated writers in Britain and the Continent to explain more carefully the basic rules of the calculus. In the 1780s a growing interest in the foundations of analysis was reflected in the decisions of the academies of Berlin and Saint Petersburg to devote prize competitions to the metaphysics of the calculus and the nature of the infinite. In philosophy Immanuel Kant's *Kritik der reinen Vernunft* (1787) set forth a penetrating study of mathematical knowledge and initiated a new critical movement in the philosophy of science.

The contents of the *Théorie des fonctions analytiques* is summarised in Table 1. The book was divided in three parts, the first part devoted to analysis and the second and third

Table 1. Contents of Lagrange's book. The original pagination is given.

Page	**Topics**
Part One	*Exposition of the theory, with its principal uses in analysis.*
1	Preliminaries, series developments, and derived functions.
15	Series expansions for algebraic and transcendental functions.
28	Composite functions, exceptional cases.
41	Expression for the remainder.
50	Equations among derived functions.
80	Primitive functions.
91	Functions of several variables.
Part Two	*Application of the theory to geometry and to mechanics.*
	Application to geometry.
117	Theory of contacts.
147	Developable curves.
150	Maxima and minima of a function of a single variable.
155	Areas and path-lengths.
161	Differential geometry of surfaces.
187	Maxima and minima.
200	Method of variations.
	Application to mechanics.
223	Speed and acceleration.
232	Particle dynamics.
241	Motion in a resisting medium.
251	Constrained motion.
256	Conservation theorems.
271	Impact of bodies, machine performance. [End 277.]

parts devoted to geometry and mechanics. The *Leçons sur le calcul des fonctions* concentrated almost exclusively on analysis, and included a detailed account of the calculus of variations. The material in both books originated in lectures that Lagrange delivered at the *Ecole Polytechnique*. He wrote them when he was in his sixties, still an active mathematician but certainly past his prime creative period of scientific research.

Although Lagrange's books appeared at the dawn of the new century, they encapsulated the prevailing understanding of analysis, refining conceptions that had been set forth by Leonhard Euler (1707–1783) in his textbooks of the 1740s and 1750s (§13, §14). Lagrange's fundamental axiom involving the Taylor-series expansion of a function originated in a memoir he published in 1774 in the proceedings of the Berlin Academy. Near the beginning of the *Théorie*, he stated that Louis Arbogast had submitted a detailed memoir to the Paris Academy developing these ideas. The memoir was never published, although Arbogast discussed it in his book [1800]; because it did not appear, and because Lagrange himself happened to be involved in a study of the general principles of analysis as a result of 'particular circumstances' (presumably his teaching duties), he decided to write a treatise generalizing and extending his earlier ideas.

Part One of the *Théorie* begins with some historical matters and examines the basic expansion of a function as a Taylor power series. There is considerable discussion of values where the expansion may fail, and a derivation of such well-known results as l'Hôpital's rule. Lagrange then turned to methods of approximation and an estimation of the remainder in the Taylor series, followed by a study of differential equations, singular solutions and series methods, as well as multi-variable calculus and partial differential equations. He outlined and supplemented topics explored in some detail in memoirs of the 1760s and 1770s.

Part Two on geometry opens with an investigation of the geometry of curves. Here Lagrange examined in detail the properties that must hold at a point where two curves come into contact—the relationships between their tangents and osculating circles. Corresponding questions concerning surfaces are also investigated, and Lagrange referred to Gaspard Monge's memoirs on this subject in the *Académie des Sciences*. He derived some standard results on the quadrature and rectification of curves. The theory of maxima and minima in the ordinary calculus, a topic Lagrange suggested could be understood independently of geometry as part of analysis, is taken up. Also covered are basic results in the calculus of variations, including an important theorem of Adrien-Marie Legendre in the theory of sufficiency. The topic of the calculus of variations was treated on an analytical level much more extensively in the *Leçons*.

The third part on dynamics is somewhat anticlimactic, given the publication nine years earlier of his major work *Méchanique analitique* (§16). In this part Lagrange presented a rather kinematically-oriented investigation of particle dynamics, including a detailed discussion of the Newtonian problem of motion in a resisting medium. He also derived the standard conservation laws of momentum, angular momentum and live forces. The book closes with an examination of the equation of live forces as it applies to problems of elastic impact and machine performance.

In our account of the *Théorie* we will concentrate on some of the major original contributions of this work: the formulation of a coherent foundation for analysis; Lagrange's conception of theorem-proving in analysis; his derivation of what is today called the Lagrange remainder in the Taylor expansion of a function; his formulation of the multiplier rule in the calculus and calculus of variations; and his account of sufficiency questions in the calculus of variations. Pages of this book, and also from the *Leçons,* are cited from the *Oeuvres* edition.

2 ALGEBRAIC ANALYSIS AND THE FUNCTION CONCEPT

The full title of the *Théorie* explains its purpose: 'Theory of analytical functions containing the principles of the differential calculus disengaged from all consideration of infinitesimals, vanishing limits or fluxions and reduced to the algebraic analysis of finite quantities'. Lagrange's goal was to develop an algebraic basis for the calculus that made no reference to infinitely small magnitudes or intuitive geometrical and mechanical notions. In a treatise on numerical equations published in 1798 he set forth clearly his conception of algebra [1798, 14–15]:

> [Algebra's] object is not to find particular values of the quantities that are sought, but the system of operations to be performed on the given quantities in order to derive from them the values of the quantities that are sought. The tableau of these operations represented by algebraic characters is what in algebra is called a *formula*, and when one quantity depends on other quantities, in such a way that it can be expressed by a formula which contains these quantities, we say then that it is a *function* of these same quantities.

Lagrange used the term 'algebraic analysis' to designate the part of mathematics that results when algebra is enlarged to include calculus-related methods and functions. The central object here was the concept of an analytical function. Such a function $y = f(x)$ is given by a single analytical expression, constructed from variables and constants using the operations of analysis. The relation between y and x is indicated by the series of operations schematized in $f(x)$. The latter possesses a well-defined, unchanging algebraic form that distinguishes it from other functions and determines its properties.

The idea behind Lagrange's theory was to take any function $f(x)$ and expand it in a power series about x:

$$f(x+i) = f(x) + pi + qi^2 + ri^3 + si^4 + \cdots. \tag{1}$$

The 'derived function' or derivative $f'(x)$ of $f(x)$ is defined to be the coefficient $p(x)$ of the linear term in this expansion. $f'(x)$ is a new function of x with a well-defined algebraic form, different from but related to the form of the original function $f(x)$. Note that this conception is very different from that of the modern calculus, in which the derivative of $f(x)$ is defined at each value of x by a limit process. In the modern calculus the relationship of the derivative to its parent function is specified in terms of correspondences defined in a definite way at each value of the numerical continuum.

Lagrange's understanding of derived functions was revealed in his discussion in the *Leçons* of the method of finite increments. This method was of historical interest in the background to his programme. Brook Taylor's original derivation in 1715 of Taylor's theorem was based on a passage to the limit of an interpolation formula involving finite increments. Lagrange wished to distinguish clearly between an approach to the foundation of the calculus that uses finite increments and his own quite different theory of derived functions. In taking finite increments, he noted, one considers the difference $f(x_{n+1}) - f(x_n)$ of the same function $f(x)$ at two successive values of the independent argument. In the differential calculus the object Lagrange referred to as the derived function was traditionally obtained by letting $dx = x_{n+1} - x_n$ be infinitesimal, setting $dy = f(x_{n+1}) - f(x_n)$, dividing dy by dx, and neglecting infinitesimal quantities in the resulting reduced expression for dy/dx. Although this process leads to the same result as Lagrange's theory, the connection it presumes between the method of finite increments and the calculus obscures a more fundamental difference between these subjects: in taking $\Delta y = f(x_{n+1}) - f(x_n)$ we are dealing with one and the same function $f(x)$; in taking the derived function we are passing to a new function $f'(x)$ with a new algebraic form. Lagrange explained this point as follows [1806, 270, 279]:

> [...] the passage from the finite to the infinite requires always a sort of leap, more or less forced, which breaks the law of continuity and changes the form of functions.
>
> [...] in the supposed passage from the finite to the infinitely small, functions actually change in nature, and [...] dy/dx, which is used in the differential Calculus, is essentially a different function from the function y, whereas as long as the difference dx has any given value, as small as we may wish, this quantity is only the difference of two functions of the same form; from this we see that, if the passage from the finite to the infinitely small may be admitted as a mechanical means of calculation, it is unable to make known the nature of differential equations, which consists in the relations they give between primitive functions and their derivatives.

In Lagrange's conception of analysis, one is given a universe of functions, each expressed by a formula $y = f(x)$ and consisting of a single analytical expression involving variables, constants and algebraic and transcendental operations. During the 18th century such functions were called continuous, and the *Théorie* is devoted exclusively to functions that are continuous in this sense. (Mathematicians were aware of the possibility of other sorts of functions, but alternate definitions never caught on.) Such functions were naturally suited to the usual application of calculus to geometrical curves. In studying the curve the calculus is concerned with the connection between local behaviour, or slope, and global behaviour, or area and path-length. If the curve is represented by a function $y = f(x)$ given by a single analytical expression then the relation between x and y is permanently established in the form of f. Local and global behaviour become identified in this functional relation.

It is also necessary to call attention to the place of infinite series in Lagrange's system of analysis. Each function has the property that it may be expanded as the power series (1). Nevertheless, an infinite series as such is never defined to be a function. The logical concept of an infinite series as a functional object defined *a priori* with respect to some criterion such as convergence or summability was foreign to 18th-century analysis. Series expansions were understood as a tool for obtaining the derivative, or a way of representing functions that were already given.

For the 18th-century analyst, functions are things that are given 'out there', in the same way that the natural scientist studies plants, insects or minerals, given in nature. As a general rule, such functions are very well-behaved, except possibly at a few isolated exceptional values. It is unhelpful to view Lagrange's theory in terms of modern concepts (arithmetical continuity, differentiability, continuity of derivatives and so on), because he did not understand the subject in this way.

3 THEOREMS OF ANALYSIS

3.1 Expansions

Lagrange was aware that the expansion of a function as the series (1) may fail at particular values of x, and he discussed this point at some length in the *Théorie*. He reasoned that the expansion of $f(x+i)$ can contain no fractional powers of i. He illustrated this conclusion by means of the example $f(x) = \sqrt{x}$. Suppose indeed that we had a relation of the following form for the expansion of $\sqrt{(x+i)}$:

$$\sqrt{x+i} = \sqrt{x} + ip + i^2 q + i^3 r + \cdots + i^{m/n}. \tag{2}$$

This equation establishes a relation of equality between the 2-valued function on the left side, and the $2n$-valued function on the right side, a result that is evidently absurd. Hence it must be the case that the powers of i in the expansion of $\sqrt{(x+i)}$ are all integral.

Lagrange noted that the 'generality' and 'rigour' of this argument require that x be indeterminate (p. 8: it is interesting that he associates generality and rigour, a point of view characteristic of 18th-century algebraic analysis). In particular cases such as $x = 0$ we will have fractional powers of i, but this arises because certain formal features of the function—in the case at hand the radical $\sqrt{x}$—disappear at $x = 0$.

3.2 Taylor's theorem

Lagrange's understanding of what it meant to prove a theorem of analysis differed from the understanding which developed in later analysis and which is customary today. To prove a theorem was to establish its validity on the basis of the general formal properties of the relations, functions, and formulae in question. The essence of the result was contained in its general correctness, rather than in any considerations about what might happen at particular numerical values of the variables.

The derived function $f'(x)$ is the coefficient of the linear term in the expansion of $f(x+i)$ as a power series in i. By definition, the second derived function $f''(x)$ is the coefficient of i in the expansion of $f'(x+i)$, the third derived function $f'''(x)$ is the coefficient of i in the expansion of $f''(x+i)$, and so on.

In art. 16 Lagrange related the coefficients $q, r, s, \ldots$ in (1) to the higher-order derived functions $f''(x)$, $f'''(x)$, $f^{(\text{iv})}(x), \ldots$. If we replace i by $i+o$ in (1) we obtain

$$\begin{aligned} f(x+i+o) &= f(x) + (i+o)p = (i+o)^2 q + (i+o)^3 r + (i+o)^4 s + \cdots \\ &= f(x) + ip + i^2 q + i^3 r + i^4 s + \cdots \\ &\quad + op + 2ioq + 3i^2 or + 4i^3 os + \cdots . \end{aligned} \tag{3}$$

Suppose now the we replace x by $x+o$. $f(x)$, p, q, r then become

$$f(x)+op+\cdots, \quad p+op'+\cdots, \quad q+oq'+\cdots, \quad r+or'+\cdots. \tag{4}$$

If we next increase $x+o$ by i we obtain (using $x+i+o=(x+o)+i$)

$$\begin{aligned} f(x+i+o) = {} & f(x)+op+\cdots+i(p+op'+\cdots)+i^2(q+oq'+\cdots) \\ & +i^3(r+or'+\cdots)+\cdots. \end{aligned} \tag{5}$$

Equating (3) and (5) we obtain

$$q=\frac{1}{2}p', \quad r=\frac{1}{3}q', \quad s=\frac{1}{4}r', \quad \ldots . \tag{6}$$

The derived functions $f'(x)$, $f''(x)$, $f'''(x), \ldots$ are the coefficients of i in the expansions of $f(x+i)$, $f'(x+i)$, $f''(x+i), \ldots$. Hence

$$q=\frac{1}{2}f''(x), \quad r=\frac{1}{2\cdot 3}f'''(x), \quad s=\frac{1}{2\cdot 3\cdot 4}f^{(\mathrm{iv})}(x), \quad \ldots . \tag{7}$$

Thus the series (1) becomes

$$f(x+i)=f(x)+if'(x)+\frac{i^2}{2}f''(x)+\frac{i^3}{2\cdot 3}f'''(x)+\frac{i^4}{2\cdot 3\cdot 4}f^{(\mathrm{iv})}+\cdots, \tag{8}$$

which is the Taylor series for $f(x+i)$.

3.3 The theorem on the equality of mixed partial derived functions

In art. 86 Lagrange considered a function $f(x, y)$ of the two variables x and y. He observed that $f(x+i, y+o)$ can be expanded in two ways. First, we expand $f(x+i, y+o)$ with respect to i, and then expand the expression which results with respect to o. The expansion for $f(x+i, y+o)$ obtained in this way is presented at the top of p. 93:

$$\begin{aligned} f(x+i, y+o) = {} & f(x, y)+if'(x, y)+of_{\prime}(x, y)+\frac{i^2}{2}f''(x, y)+iof'_{\prime}(x, y) \\ & +\frac{o^2}{2}f_{\prime\prime}(x, y)+\frac{i^3}{2\cdot 3}f'''(x, y)+\frac{i^2 o}{2}f''_{\prime}(x, y) \\ & +\frac{io^2}{2}f'_{\prime\prime}(x, y)+\frac{o^3}{2\cdot 3}f_{\prime\prime\prime}(x, y)+\&c. \end{aligned} \tag{9}$$

One of the terms in this expansion is $f'_{\prime}(x, y)$, where the superscript prime denotes partial differentiation with respect to x and the subscript prime denotes partial differentiation with respect to y, and where the differentiation occurs first with respect to x and second with respect to y.

However, we could also expand $f(x+i, y+o)$ with respect to o, and then expand the expression which results with respect to i. In the expansion obtained in this way, we again have the term $f'{}_{\prime}(x, y)$, except that here the partial differentiation occurs first with respect to y and second with respect to x. By equating the two series expansions for $f(x+i, y+o)$ we are able to deduce that the two quantities $f'{}_{\prime}(x, y)$ are equal. In modern notation we have $\partial^2 f/\partial x \partial y = \partial^2 f/\partial y \partial x$. Lagrange's notation system does not allow one to indicate the order of differentiation, but fortunately the order does not matter.

This result evidently applies to all functions $f(x, y)$ for all ranges of the variables x and y, except possibly at isolated exceptional values. Lagrange considered a couple of examples. Suppose $f(x, y) = x\sqrt{(2xy + y^2)}$. If we differentiate f with respect to x and then with respect to y we obtain

$$\frac{x+y}{\sqrt{2xy+y^2}} + \frac{x^2 y}{(2xy+y^2)^{3/2}}. \tag{10}$$

However, if we differentiate f with respect to y and then with respect to x we have

$$\frac{2x+y}{\sqrt{2xy+y^2}} - \frac{(x^2+xy)y}{(2xy+y^2)^{3/2}}. \tag{11}$$

Although these two expressions appear to be different, it is not difficult to see that both reduce to the one and the same expression

$$\frac{3x^2y + 3xy^2 + y^3}{(2xy+y^2)^{3/2}}. \tag{12}$$

Lagrange supplied a second example to provide further confirmation of his theorem.

4 METHODS OF APPROXIMATION

4.1 Lagrange's form of the remainder

In arts. 45–53 Lagrange developed results that belong to the core of any modern course in real analysis. Indeed, it is likely that this part of the treatise influenced Augustin-Louis Cauchy when he wrote his famous textbooks initiating modern analysis 25 years later (§25). However, for Lagrange the results in question were not fundamental: they did not belong to the foundation of the subject. His purpose rather was essentially practical, to obtain a result that would be useful in the approximation of functions. Thus he derived an expression for the remainder in the Taylor series when the series is terminated after the nth term. The result allowed for a general method for approximating functions by obtaining the bounds on the error committed if one approximates a function by the first n terms of its Taylor expansion.

Lagrange first proved the following lemma. If $f'(x)$ is positive and finite throughout the interval $a \leqslant x \leqslant b$, then the primitive function $f(x)$ satisfies the inequality $f(b) - f(a) \geqslant 0$. Consider the expansion

$$f(z+i) = f(z) + if'(z) + \frac{i^2}{2} f''(z) + \&c. \tag{13}$$

For sufficiently small i, the linear term in the expansion on the right side will dominate the sum of the remaining terms. Thus if $f'(z)$ is positive, and i is taken to be a sufficiently small positive quantity, it follows that $f(z+i) - f(z)$ will be positive. Consider the succession of values $a, a+i, a+2i, \ldots, a+ni$. By assumption, $f'(a+i)$, $f'(a+2i), \ldots, f'(a+ni)$ are positive. Thus if i is taken positive and small enough, each of the quantities $f(a+i) - f(a)$, $f(a+2i) - f(a+i), \ldots, f(a+(n+1)i) - f(a+ni)$ will be positive. (Lagrange is evidently assuming a uniformity property with respect to $f(z+i) - f(z)$.) If we let $a + (n+1)i = b$ and add together all of the quantities, it follows that $f(b) - f(a) \geqslant 0$. Hence the lemma is proved.

Lagrange explicitly stated the condition that the derived function $f'(x)$ be finite on the given interval because it was clear from examples that the lemma fails otherwise. In the *Leçons* he cited the example $y = 1/(a-z) - 1/a$ $(a > 0)$ [Ovaert, 1976, 222]. The derived function is $1/(a-z)^2$, which is positive everywhere. Nevertheless, for the interval $[0, b]$ $(b > a)$, it is clear that $f(b) - f(0)$ is negative. However, in this example the derived function is infinite at $x = a$, and so the conditions of the lemma do not hold.

We turn now to Lagrange's derivation of the remainder in the Taylor power series. He first introduced a second variable z and wrote $x = (x - xz) + xz$. Series (8) becomes

$$f(x) = f(x-xz) + xzf'(x-xz) + \frac{x^2z^2}{2\cdot 1}f''(x-xz) + \frac{x^3z^3}{3\cdot 2\cdot 1}f'''(x-xz) + \cdots. \quad (14)$$

We rewrite (14) in the form

$$f(x) = f(x-xz) + xP(x,z). \quad (15)$$

If we differentiate (15) with respect to z we obtain

$$0 = -xf'(x-xz) + xP'(x,z), \quad (16)$$

so that

$$P'(x,z) = f'(x-xz). \quad (17)$$

Suppose that z belongs to the interval $[a, b]$, $a \geqslant 0$. Let N and M be the maximum and minimum values of $P'(x,z)$ on this interval. We have the inequalities

$$N \leqslant P'(x,z) \leqslant M, \quad a \leqslant z \leqslant b. \quad (18)$$

It follows that $P'(x,z) - N \geqslant 0$ and $M - P'(x,z) \geqslant 0$ for $a \leqslant z \leqslant b$. Applying the above lemma to the functions $P - N$ and $M - P$ we obtain

$$P(x,a) + N(b-a) \leqslant P(x,b) \leqslant P(x,a) + M(b-a). \quad (19)$$

Now $P(x,z) = 0$ if $z = 0$. Setting $a = 0$ and $b = 1$ in (19) we obtain

$$N \leqslant P(x,1) \leqslant M. \quad (20)$$

As z goes from 0 to 1, $(x - xz)$ goes from x to 0. From (19) and (20) it follows that $f'(x - xz)$ takes on all values between N and M. (Lagrange is assuming here that $f'(x - xz)$

satisfies an intermediate-value property.) Hence for some u with $0 \leqslant u \leqslant x$ we have, by (20),

$$P(x, 1) = f'(u). \tag{21}$$

Hence the original series (8) may be written for $z = 1$ as

$$f(x) = f(0) + xf'(u), \quad 0 \leqslant u \leqslant x. \tag{22}$$

(22) expresses what is today called 'the mean-value theorem'.

Let us now write (14) in the form

$$f(x) = f(x - xz) + xzf'(x - xz) + x^2 Q(x, z). \tag{23}$$

By differentiating each side of (23) with respect to z we easily deduce that

$$Q'(x, z) = zf''(x - xz). \tag{24}$$

Let N_1 and M_1 be the minimum and maximum values of $f''(x - xz)$ for $a \leqslant z \leqslant b$:

$$N_1 \leqslant f''(x - xz) \leqslant M_1. \tag{25}$$

Since $z \geqslant a \geqslant 0$, we have

$$zN_1 \leqslant zf''(x - xz) \leqslant zM_1, \quad \text{or} \quad zN_1 \leqslant Q'(x, z) \leqslant zM_1. \tag{26}$$

From the lemma we conclude that

$$Q(x, a) + \frac{N_1(b^2 - a^2)}{2} \leqslant Q(x, b) \leqslant Q(x, a) + \frac{M_1(b^2 - a^2)}{2}. \tag{27}$$

Setting $a = 0$ and $b = 1$ in (27) there follows

$$\frac{N_1}{2} \leqslant Q(x, 1) \leqslant \frac{M_1}{2}. \tag{28}$$

It is clear from (25) and (28) that $Q(x, 1) = f''(u)/2$ for some $u \in [0, x]$. Hence for $z = 1$ series (8) becomes

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2 \cdot 1} f''(u), \quad 0 \leqslant u \leqslant x. \tag{29}$$

Lagrange proceeded to extend the reasoning used to obtain (22) and (29) to derive the equation

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2 \cdot 1} f''(0) + \frac{x^3}{3 \cdot 2 \cdot 1} f'''(u), \quad 0 \leqslant u \leqslant x. \tag{30}$$

As he indicated on p. 49, equations (20), (29) and (30) may be iterated to obtain expressions for $f(x)$ involving derivatives of f of any order evaluated at u for $0 \leqslant u \leqslant x$. He observed that there results 'a theorem which is new and remarkable for its simplicity and generality'. The theorem gives what is today called 'the Lagrange form' of the remainder in the Taylor series.

By taking the function $g(x) = f(x + z)$ and applying the preceding result to $g(x)$ we obtain immediately

$$\begin{aligned} f(z+x) &= f(z) + xf'(u) = f(z) + xf'(z) + \frac{x^2}{2} f''(u) \\ &= f(z) + xf'(z) + \frac{x^2}{2} f''(z) + \frac{x^3}{2 \cdot 3} f'''(u), \end{aligned} \tag{31}$$

&c., where $0 \leqslant u \leqslant x$.

Lagrange called attention to the importance of (31) for methods of approximation and emphasized its utility in geometrical and mechanical problems. Although he gave no examples, the usefulness of (31) is evident in one of the functions that he introduced earlier, the exponential function $f(x) = e^x$. We use it to approximate $f(1) = e$, following the account in [Courant, 1937, 326–327]. For $z = 0$ and $x = 1$ (31) becomes

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \frac{e^u}{(n+1)!}, \tag{32}$$

where $0 \leqslant u \leqslant 1$. We have

$$\begin{aligned} e &= 1 + 1 + 1/2! + 1/3! + 1/4! + \cdots < 1 + 1 + 1/2 + 1/2^2 + 1/2^3 + \cdots \\ &= 1 + 2 = 3, \quad \text{or} \quad e < 3. \end{aligned} \tag{33}$$

Hence the error committed in neglecting the remainder term in (32) will be less than $3/(n+1)!$ To obtain an approximation of e with an error smaller $1/10{,}000$, we observe that $8! > 30{,}000$, and arrive at the estimate

$$\begin{aligned} e &\approx 1 + 1 + \frac{1}{2} + \frac{1}{3 \cdot 2} + \frac{1}{4 \cdot 3 \cdot 2} + \frac{1}{5 \cdot 4 \cdot 3 \cdot 2} + \frac{1}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2} + \frac{1}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2} \\ &= 2.71822. \end{aligned} \tag{34}$$

4.2 Mean-value theorem

Lagrange obtained equation (21), the mean-value theorem (a term he never used), from his fundamental axiom concerning the expansion of a function in a Taylor series. In modern analysis this result is derived in a different way from the basic properties of continuous and differentiable functions. Today the mean-value theorem is a cornerstone of the foundation of real analysis. To prove a theorem is to establish its correctness for each value of the functional variable. The law of the mean is used in theorem-proving in order to take whatever property or relation that is under consideration and localize it a given number. A typical application of this law is found in the modern proof of the theorem on the equality of mixed partial derivatives (discussed above in section 3.3). One takes a point of the plane, establishes the equality in question for finite increments, and then extends this result to derivatives using the law of the mean. Lagrange's reasoning involved very different conceptions and assumptions, and indicates the large difference between his perspective and the modern one.

There is in the *Théorie* one place where Lagrange's derivation resembles the proof-structure of modern analysis. This occurs in art. 134, in the second part of the book devoted to the application of analysis to geometry. Consider a function $y = f(x)$, and suppose that y and x are the coordinates of a curve in a rectangular coordinate system. Let $F(x)$ be the area under the curve from $x = 0$ to $x = x$. Using the mean-value theorem and some reasoning involving inequalities, Lagrange was able to show that $F(x)$ is a primitive function for $f(x)$. Of course, this is the result known today as the fundamental theorem of calculus and Lagrange's proof is not dissimilar to the reasoning involved in a modern derivation. It is nevertheless important to appreciate the distinctiveness of his approach. First, the result is not fundamental nor is it even a theorem of analysis, but an application of analysis to geometry. Second, the fundamental notion for Lagrange is always the primitive as an antiderivative, rather than the integral as a sum or an area.

5 MULTIPLIER RULE

5.1 *Mechanics*

The idea for the multiplier rule seems to have originated in an interesting way in Lagrange's study of statics in Part One of the *Méchanique analitique* [1788, 45–49]; compare §16.3. Suppose that we are given a system of bodies or points, with coordinates x, y, z; x', y', z'; and so on. The system is subject to the constraints or conditions $L = 0$, $M = 0$, $N = 0$, and so on. External forces act on each of the points of the system. According to the principle of virtual velocities (what is know today as virtual displacements or virtual work), equilibrium will subsist if the following equation holds:

$$P\,dp + Q\,dq + R\,dr + \&c. = 0. \tag{35}$$

Here $P\,dp$ is the virtual work (which he called 'moment') that results when the force P acts on the point x, y, z with a corresponding virtual displacement dp. Similar interpretations hold for $Q\,dq$, $R\,dr$ and so on. On way to proceed to a solution is the following. We set $dL = 0$, $dM = 0$, $dN = 0$ and use these relations to eliminate some of the differentials dx, dy, dz; dx', dy', dz'; ... (as many differentials are eliminated as there are constraints $dL = 0$). If we substitute for the eliminated differentials in (35), we will obtain a relation in which each of the resulting differentials may be regarded as independent. By equating to zero the coefficients of these differentials, we obtain the desired equations of equilibrium.

It may well be that the required elimination of differentials is not that easy to carry out. The method of multipliers provides an alternative way of deriving the conditions of equilibrium. We multiply $dL = 0$ by the indeterminate quantity λ, dM by the quantity μ, dN by the quantity υ, and so on. Lagrange asserted that 'it is not difficult to prove by the theory of the elimination of linear equations' that the general equation of equilibrium will become

$$P\,dp + Q\,dq + R\,dr + \mu\,dM + \nu\,dN + \&c. = 0. \tag{36}$$

We now equate the coefficient of each of the differentials in (36) to zero; the resulting equations, in combination with the constraints $dL = dM = dN = \cdots = 0$, allow us to eliminate the multipliers and arrive at the desired equilibrium conditions.

Lagrange provided a rather natural physical interpretation of the multiplier constants λ appearing in (36) [1788, 48–49]. The effect of the constraint is to produce a force that acts on the point, producing the increment of virtual work or moment $\lambda\, dL$. The moments due to the constraints $L = 0$ balance the moments resulting from the external forces P. Thus the system can be regarded as subject to the constraints $L = 0$, or alternatively it can be regarded as entirely free with the constraints replaced by the forces to which they gives rise. In the latter case we obtain the free equation (36). According to Lagrange, this interpretation provides the 'metaphysical reason' why the addition of the terms $\lambda\, dL$ to the left side of (35) allows one to treat the system as entirely free in (36)—indeed, 'it is in this that the spirit of the method [of multipliers] consists'. Thus the method is justified in a natural way using physical considerations, in contrast to the analytical approach usually associated with Lagrange's mathematics [Fraser, 1981, 263–267].

5.2 Calculus

In arts. 131–184 of the *Théorie*, Lagrange moved from the domain of mechanics to analysis in his exposition of methods of maxima and minima in the calculus. Although this investigation occurred in the part of the book devoted to the applications of analysis to geometry, he noted the independence of the basic problem of optimization from considerations of curves (pp. 150–151).

In art. 167 Lagrange took up the problem of maximizing or minimizing a function of the form $f(x, y, z, \ldots)$ of any number of variables that are subject to the constraint $\phi(x, y, z, \ldots) = 0$. If we increase x, y, z and so on by the small increments p, q and r we obtain from the constraint $\phi = 0$ the relation

$$p\phi'(x) + q\phi'(y) + r\phi'(z) + \&c. + \textit{higher-order terms} = 0, \tag{37}$$

where $\phi'(x)$, $\phi'(y)$ and $\phi'(z)$ denote as usual the partial derivatives $\partial\phi/\partial x$, $\partial\phi/\partial y$ and $\partial\phi/\partial z$ of ϕ with respect to x, y and z. To arrive at the equations of maxima and minima we can neglect the higher-order terms in (37). Because f is a maximum or minimum we have as well the condition

$$pf'(x) + qf'(y) + rf'(z) + \&c. = 0. \tag{38}$$

One solution would be to solve (37) for p in terms of $q, r, \ldots,$ substitute the resulting expression for p in (38), and equate to zero the coefficients of $q, r, \ldots$. Another solution is obtained by multiplying (37) by the multiplier a and adding the resulting expression to (38). In the equation which results all of the variables $x, y, z, \ldots$ may be regarded as free, and the coefficients of each of the $p, q, r, \ldots$ may be set equal to zero:

$$f'(x) + a\phi'(x) = 0, \quad f'(y) + a\phi'(y) = 0, \quad f'(z) + a\phi'(z) = 0, \quad \&c. \tag{39}$$

(39) together with the constraint $\phi = 0$ provide the equations of solution, allowing us to determine the desired maximizing or minimizing values of a, x and y.

Lagrange did not state explicitly the precise reasoning by which he arrived at (39), but it seems to have developed along the following lines. Multiply (37) by a and add to (38) to obtain:

$$p\big(f'(x)+a\phi'(x)\big)+q\big(f'(y)+a\phi'(y)\big)+r\big(f'(z)+a\phi'(z)\big)+\&c.=0 \qquad (40)$$

In (40), define a so that $f'(x)+a\phi(x)=0$. The first term in (40) then disappears, and we can assume that the remaining variables $y, z, \ldots$ may be varied arbitrarily and independently. Hence we obtain the equations

$$f'(y)+a\phi(y)=0, \qquad f'(z)+a\phi(z)=0, \qquad \ldots . \qquad (41)$$

There are, as is well known today, geometrical ways of justifying the multiplier rule. Thus it is immediately evident (to anyone, for example, who has drawn a trail line on a topographical map) that the maximum or minimum of $f(x, y)$ along the path $\phi(x, y)=0$ will occur when this path runs parallel to a member of the family of contours or level curves $f=$ constant. At the point where this is the case the unit normals to the two curves coincide, and so we obtain equations (39). Such geometrical reasoning was not used by Lagrange, whose approach in the *Théorie* was analytical throughout.

5.3 Calculus of variations

In the ordinary calculus it would be possible to do without the multiplier rule, this rule being a useful but finally unessential technique of solution. In the calculus of variations the situation is very different: in problems of constrained optimization, where the side conditions are differential equations, the multiplier rule is the only general method for obtaining the variational equations. The situation in the ordinary calculus is similar to that of the calculus of variations when the side constraints are finite equations. In introducing the rule in the variational calculus [Courant and Hilbert, 1953, 221] write: 'Up to now [i.e. in the case of finite constraints] the multiplier rule has been used merely as an elegant artifice. But multipliers are indispensable if the subsidiary condition takes the general form $G(x, y, z, y', z')=0$, where the expression $G(x, y, z, y', z')$ cannot be obtained by differentiating an expression $H(x, y, z)$ with respect to x, i.e. where G is a *nonintegrable differential expression*'.

In art. 181 of the *Théorie*, Lagrange formulated the multiplier rule for problems of constrained optimization in the calculus of variations. In the *Leçons* he provided an extended treatment of this subject, including the presentation of detailed examples. The multiplier rule proved to be an extremely powerful and effective tool, enabling one to derive results that could only be obtained with considerable difficulty otherwise.

In a variational problem with two dependent variables it is necessary to optimize the primitive or integral of $f(x, y', y'', \ldots, z, z', z'', \ldots)$ evaluated between $x=a$ and $x=b$. The solution must satisfy the Euler–Lagrange differential equations, which Lagrange wrote as

$$f'(y)-\big[f'(y')\big]'+\big[f'(y'')\big]''-\big[f'(y''')\big]'''+\&c=0, \qquad (42)$$

$$f'(z)-\big[f'(z')\big]'+\big[f'(z'')\big]''-\big[f'(z''')\big]'''+\&c=0. \qquad (43)$$

(In modern notation the term $[f'(y)]'$ is $d(\partial f/\partial y')/dx$.) To obtain these equations Lagrange used algebraic, analogical reasoning very different from the modern methods of real analysis.

Suppose now that the variables $x, y, z, \ldots$ satisfy a constraint of the form

$$\phi(x, y, y', \&c., z, z', \&c.) = 0, \tag{44}$$

consisting of a differential equation of arbitrary order connecting the variables of the problem. To obtain the Euler–Lagrange equations in this situation, we multiply (44) by the multiplier function $\Delta(x)$ (note that Δ is a function of x and not a constant) and form the differential equations

$$f'(y) - \left[f'(y')\right]' + \left[f'(y'')\right]'' - \left[f'(y''')\right]''' + \&c + \Delta\phi'(y) - \left[\Delta\phi'(y')\right]' + \left[\Delta\phi'(y'')\right]'' - \&c. = 0; \tag{45}$$

$$f'(z) - \left[f'(z')\right]' + \left[f'(z'')\right]'' - \left[f'(z''')\right]''' + \&c + \Delta\phi'(z) - \left[\Delta\phi'(z')\right]' + \left[\Delta\phi'(z'')\right]'' - \&c. = 0. \tag{46}$$

In the *Leçons*, Lagrange applied the multiplier rule to two examples involving the motion of a particle descending through a resisting medium. These examples had originally appeared in Chapter 3 of Euler's *Methodus inveniendi* (1744) and concerned the brachistochrone and the curve of maximum terminal velocity (§12.2). Assume the y-axis is measured horizontally and the x-axis is measured vertically downward, and let z equal the square of the speed. We have the dynamical constraint equation

$$z' - 2g + 2\phi(z)\sqrt{1 + y'^2} = 0, \tag{47}$$

giving z as a function of x, y and y'. In each of the examples in question Lagrange wrote down the variational equations (45)–(46). Using them and the constraint equation he obtained differential equations for the multiplier function and the trajectory. Because no restriction was placed on the end value of z in the class of comparison arcs, we obtain another equation, one that allows us to calculate a constant appearing in the expression for the multiplier.

Lagrange was led by means of his multiplier rule to quite straightforward solutions of these problems, arrived at independently of the specialized methods that had appeared in Euler's and his own earlier writings. It is reasonable to assume that the successful treatment of such advanced examples would have confirmed in his mind the validity of the rule and instilled a confidence in the basic correctness of the analytical procedure involved in its application.

In the later calculus of variations the multiplier rule would assume an even more fundamental role and become the basic axiom of the whole subject. Alfred Clebsch (1833–1872) showed how the multiplier rule can be used to reduce problems with higher-order derivatives to problems involving only first derivatives and side constraints [Clebsch, 1858]. In the modern subject, any problem with side constraints is known as a Lagrange problem and it is solved by means of the multiplier rule. The most general problem of the calculus of variations can be formulated as such a problem and solved in principle using the rule.

6 CALCULUS OF VARIATIONS: SUFFICIENCY RESULTS

In arts. 174–178 Lagrange took up the question of sufficiency in the calculus of variations. Given that a proposed solution satisfies some of the conditions of the problem, it is necessary to investigate what additional conditions must hold in order that there be a genuine maximum or minimum. Here Lagrange reported on results of Adrien-Marie Legendre (1752–1833) published as [Legendre, 1788], and added some important new observations of his own.

For simplicity we consider the case where there is only one dependent variable and where only the first derivative appears in the variational integrand. In a problem in the calculus of variation, a proposed solution will be optimal for Legendre if the sign of the second variation is unchanged (always positive for a minimum, or always negative for a maximum) with respect to all comparison arcs. Let the increment or variation of y be the function $w(x)$. The second variation I_2 is by definition

$$I_2 = \int_{x_0}^{x_1} \left(\frac{\partial^2 f}{\partial y^2} w^2 + 2\frac{\partial^2 f}{\partial y \partial y'} w w' + \frac{\partial^2 f}{\partial y'^2} w'^2 \right) dx. \tag{48}$$

It is necessary to investigate the sign of I_2. Let $v = v(x)$ be a function of x and consider the expression

$$\frac{d}{dx}(w^2 v), \tag{49}$$

Because $w(x_0) = w(x_1) = 0$ the integral of (49) is zero:

$$\int_{x_0}^{x_1} \frac{d}{dx}(w^2 v)\, dx = 0. \tag{50}$$

We introduce some standard abbreviations for the second partial derivatives:

$$P = \frac{\partial^2 f}{\partial y^2}, \qquad Q = \frac{\partial^2 f}{\partial y \partial y'}, \qquad R = \frac{\partial^2 f}{\partial y'^2}. \tag{51}$$

If we add the integral of (49) to the expression for the second variation I_2 given in (48) there results no change in its value:

$$I_2 = \int_{x_0}^{x_1} \left((P + v')w^2 + 2(Q + v) w w' + R w'^2 \right) dx. \tag{52}$$

The integrand is a quadratic expression in w and w'. Legendre observed that it will become a perfect square if

$$R(P + v') = (Q + v)^2. \tag{53}$$

For $v(x)$ satisfying this differential equation the second variation becomes

$$I_2 = \int_{x_0}^{x_1} R\left(w' + \frac{Q+v}{R}w\right)^2 dx. \tag{54}$$

It is evident that the given transformation is only possible if $R = \partial^2 f/\partial y'^2$ is non-zero on the interval $[x_0, x_1]$. The proposed solution will indeed be a minimum if on the interval we have

$$\frac{\partial^2 f}{\partial y'^2} > 0, \tag{55}$$

which would become known in the later subject as 'Legendre's condition'.

In order to arrive at the expression (54) for I_2 and the associated condition (55) on $\frac{\partial^2 f}{\partial y'^2}$ it is necessary to show that solutions to the differential equation (53) exist and remain finite on the given interval. In his study of the second variation in the *Théorie*, Lagrange called attention to this point and produced examples in which no finite solutions exist (pp. 206–210). Suppose for example that $f(x, y, y') = y'^2 - y^2$. In this case $P = -2$, $Q = 0$ and $R = 2$ and (54) becomes $2(v' - 2) = v^2$. By elementary methods this equation may be integrated to produce $v = 2\tan(x + c)$, where c is a constant. It is clear that if $x_1 - x_0$ is greater than $\pi/2$, then no solution of (53) will exist.

Lagrange's exposition of Legendre's theory was important because it made known to a wide audience results that likely would otherwise have remained buried in the memoirs of the Paris *Académie*. Carl Gustav Jacobi (1804–1851), in his ground-breaking paper [1837] on sufficiency theory, began his investigation with Lagrange's formulation of the subject. Lagrange's discussion of the solutions of (53) also raised new considerations that were important stimuli for Jacobi's investigation.

7 CONCLUSION

An important and under-appreciated contribution of the history of mathematics is to provide insight into the foundations of a mathematical theory by identifying the characteristics of historically earlier formulations of the theory. The conceptual relativism of scientific theories over history is one of the major findings of the history of science since Thomas Kuhn. In this respect, history of mathematics possesses a special interest lacking in the history of other branches of science, because earlier mathematical theories are enduring objects of technical interest and even of further development.

What is primarily of note in Lagrange's *Théorie*, beyond its substantial positive achievements, is that it developed analysis from a perspective that is different from the modern one. It did so in a detailed and sophisticated way, with a self-conscious emphasis on the importance of building a sound foundation. A product of the intellectual milieu of advanced research at the end of the 18th century, it stands at the cusp between algebraic and modern analysis. It retains an historical interest for us today that transcends its contribution to technical mathematics.

Although the foundation Lagrange proposed did not achieve final acceptance, his conception of analysis exerted considerable influence in the 19th century. Cauchy was able to adapt many of Lagrange's results and methods in developing an arithmetical basis for the calculus [Grabiner, 1981]—while also refuting his belief in the generality of (1) (§25). The spread of Continental analysis to Britain relied heavily on Lagrange's writings. The formal algebraists George Peacock in England and Martin Ohm in Germany were influenced by his mathematical philosophy. Lagrange's insistence that analysis should avoid geometrical and mechanical ideas was taken up with some emphasis by the Bohemian philosopher Bernhard Bolzano. The school of researchers who used operator methods in the theory of differential equations, François Servois in France, and George Boole and Duncan Gregory in Britain, took inspiration from Lagrange's writings on analysis (§36.2). Finally, even after the consolidation of Cauchy's arithmetical foundation, Lagrange's emphasis on the algorithmic, operational character of the calculus continued to inform writers of textbooks as well as non-mathematicians such as engineers and physicists who used calculus in their research and teaching.

BIBLIOGRAPHY

Arbogast, L. 1800. *Du calcul des dérivations et de ses usages dans la théorie des suites et dans le calcul différentiel*. Strasbourg: Levrault.

Clebsch, R.F.A. 1858. 'Ueber die Reduktion der zweiten Variation auf ihre einfachste Form', *Journal für die reine und angewandte Mathematik, 60*, 254–273.

Courant, R. 1937. *Differential and integral calculus*, vol. 1, 2nd ed., New York: Interscience Publishers. [Trans. of German original.]

Courant, R. and Hilbert, D. 1953. *Methods of mathematical physics*, vol. 1, New York: Interscience Publishers. [Trans. of German original.]

Fraser, C. 1981. 'The approach of Jean D'Alembert and Lazare Carnot to the theory of a constrained dynamical system', Ph.D. dissertation, University of Toronto.

Fraser, C. 1985. 'J.L. Lagrange's changing approach to the foundations of the calculus of variations', *Archive for the history of exact sciences*, *32*, 151–191.

Fraser, C. 1987. 'Joseph Louis Lagrange's algebraic vision of the calculus', *Historia mathematica*, *14*, 38–53.

Grabiner, J. 1981. *The origins of Cauchy's rigorous calculus*, Cambridge, MA: MIT Press.

Jacobi, C.G.J. 1837. 'Zur Theorie der Variations-Rechnung und der Differential-Gleichungen', *Journal für die reine und angewandte Mathematik*, *17*, 68–82.

Lagrange, J.L. *Works*. *Oeuvres*, 14 vols., Paris: Gauthier–Villars, 1867–1892. [Repr. Hildesheim: Olms, 1968.]

Lagrange, J.L. 1774. 'Sur une nouvelle espèce de calcul relatif à la différentiation & à l'intégration des quantités variables', *Nouveaux mémoires de l'Académie des Sciences et Belles-Lettres* (*Berlin*) (1772), 185–221. [Repr. in *Works*, vol. 3, 441–476.]

Lagrange, J.L. 1788. *Méchanique analitique*, 1st ed., Paris: Desaint. [2nd ed., 'Mécanique analytique', 1811–1815, 1853–1855, and *Works*, vols. 11–12 (§16).]

Lagrange, J.L. 1798. *Traité de la résolution des équations numériques de tous les degrés*, 1st ed., Paris: Courcier. [2nd ed. 1808. 3rd ed. 1826; also *Works*, vol. 8.]

Lagrange, J.L. 1801. 'Leçons sur le calcul des fonctions', *Séances des Ecoles Normales, 10*, 1–534. [Repr. as *Journal de l'Ecole Polytechnique*, *5*, *cahier* 12 (1804).]

Lagrange, J.L. 1806. *Leçons sur le calcul des fonctions*, new ed., Paris: Courcier. [Repr. as *Works*, vol. 10. Includes 'a complete treatise on the calculus of variations'.]

Legendre, A.-M. 1788. 'Sur la manière de distinguer les maxima des minima dans le calcul des variations', *Mémoires de l'Académie des Sciences de Paris*, (1786), 7–37.

Ovaert, J.-L. 1976. 'La thèse de Lagrange et la transformation de l'analyse', in C. Houzel *et alii*, *Philosophie et le calcul de l'infini*, Paris: Maspero, 159–222.

Poinsot, L. 1806. 'Théorie générale de l'équilibre et du mouvement des systèmes', *Journal de l'Ecole Polytechnique, 6, cahier 13*, 206–241. [Critical ed. P. Bailhache: Paris: Vrin, 1975.]