Asking and Answering Quetions in the History of Mathematics
Aarhus University, Denmark

Tinne Hoff Kjeldsen   Henrik Kragh Sørensen

Summer 2012

Contents

Historiography of Mathematics **1**

1 T. H. Kjeldsen, S. A. Pedersen, and L. M. Sonne-Hansen (2004). "Introduction". In: *New Trends in the History and Philosophy of Mathematics*. *New Trends in the History and Philosophy of Mathematics*. Ed. by T. H. Kjeldsen, S. A. Pedersen, and L. M. Sonne-Hansen. University of Southern Denmark Studies in Philosophy 19. Odense: University Press of Southern Denmark, pp. 11–25. 1

2 T. H. Kjeldsen (July 2012). "Uses of History for the Learning of and about Mathematics. Towards a theoretical framework for integrating history of mathematics in mathematics education". Plenary address at the HPM International Congress in Korea. 16

3 I. Grattan-Guinness (May 2004). "The mathematics of the past: distinguishing its history from our heritage". *Historia Mathematica*, vol. 31, no. 2, pp. 163–185. 38

4 J. Lützen and W. Purkert (1989). "Conflicting Tendencies in the Historiography of Mathematics: M. Cantor and H. G. Zeuthen". In: *The History of Modern Mathematics*. Ed. by D. E. Rowe, J. McCleary, and E. Knobloch. Vol. 3. Proceedings of the Symposium on the History of Modern Mathematics, *Vassar College, Poughkeepsie, New York, June 20–24, 1989*. 3 volumes. Academic Press, pp. 1–42. 61

5 M. Epple (2011). "Between Timelessness and Historiality: On the Dynamics of the Epistemic Objects of Mathematics". *Isis*, vol. 102, no. 3, pp. 481–493. 83

6 Anachronical and diachronical history. From H. Kragh (1994). *An Introduction to the Historiography of Science*. First published 1987. Cambridge etc.: Cambridge University Press, pp. 89–107. 96

7 Discussion of geometric algebra. From J. Lützen and K. Ramskov, eds. (1999). *Kilder til matematikkens historie*. 2nd ed. København: Matematisk Afdeling, Københavns Universitet, pp. 10–17. 115

Mini course on History of Analysis **123**

8 The Method of Archimedes. From T. L. Heath, ed. (1953). *The Works of Archimedes with the Method of Archimedes*. New York: Dover Publications, pp. 12–21. Adopted from J. Lützen and K. Ramskov, eds. (1999). *Kilder til matematikkens historie*. 2nd ed. København: Matematisk Afdeling, Københavns Universitet, pp. 22–26. 123

9 K. M. Pedersen (1980). "Techniques of the Calculus, 1630–1660". In: *From the Calculus to Set Theory, 1630–1910. An Introductory History*. Ed. by I. Grattan-Guinness. Princeton and Oxford: Princeton University Press. Chap. 1, pp. 10–48. 127

10 T. H. Kjeldsen (2011). "Does history have a significant role to play for the learning of mathematics? Multiple perspective approach to history, and the learning of meta level rules of mathematical discourse". In: *History and Epistemology in Mathematics Education. Proceedings of the Sixth European Summer University ESU 6*. Ed. by E. Barbin, M. Kronfellner, and C. Tzanakis. Vienna: Verlag Holzhausen GmbH, pp. 51–62. 166

11 Fermat on maxima and minima. From D. J. Struik (1969). *A Source Book in Mathematics. 1200–1800*. Cambridge (Mass.): Harvard University Press, pp. 222–225. 177

12 Fermat on maxima and minima. From J. Fauvel and J. Gray, eds. (1987). *The History of Mathematics: A Reader*. London: Macmillan Press Ltd., pp. 359–360. 181

13 Wallis on interpolation. From D. J. Struik (1969). *A Source Book in Mathematics. 1200–1800*. Cambridge (Mass.): Harvard University Press, pp. 244–246. 183

14 Roberval on the quadrature of the parabola. From E. Walker (1932). *A Study of the Traité des Indivisibles of Gilles Persone de Roberval*. New York, pp. 181–182. 186

15 H. J. M. Bos (1980). "Newton, Leibniz and the Leibnizian Tradition". In: *From the Calculus to Set Theory, 1630–1910. An Introductory History*. Ed. by I. Grattan-Guinness. Princeton and Oxford: Princeton University Press. Chap. 2, pp. 49–93. 187

16 N. Guicciardini (2003). "Newton's Method and Leibniz's Calculus". In: *A History of Analysis*. Ed. by H. N. Jahnke. History of Mathematics 24. Providence (Rhode Island): American Mathematical Society. Chap. 3, pp. 73–103. 232
17 Newton on fluxions and fluents. From M. E. Baron and H. J. M. Bos, eds. (1974). *Newton and Leibniz*. History of Mathematics: Origins and Development of the Calculus 3. The Open University Press, pp. 22–25. . . 263
18 Newton on the method of drawing tangents. From D. T. Whiteside, ed. (1964). *The Mathematical Works of Isaac Newton*. Vol. 1. Johnson Reprint Corp. 267
19 Leibniz' process of discovery. From M. E. Baron and H. J. M. Bos, eds. (1974). *Newton and Leibniz*. History of Mathematics: Origins and Development of the Calculus 3. The Open University Press, pp. 42–43. 268
20 Bishop Berkeley's The Analyst. From D. E. Smith (1959). *A source book in mathematics*. 2nd ed. 2 vols. New York: Dover Publications, Inc., pp. 627–634. Adopted from J. Lützen and K. Ramskov, eds. (1999). *Kilder til matematikkens historie*. 2nd ed. København: Matematisk Afdeling, Københavns Universitet, pp. 95–99. 270
21 J. Lützen (2003). "The Foundation of Analysis in the 19th Century". In: *A History of Analysis*. Ed. by H. N. Jahnke. History of Mathematics 24. Providence (Rhode Island): American Mathematical Society. Chap. 6, pp. 155–195. 274
22 I. Grattan-Guinness (1970). "Bolzano, Cauchy and the "New Analysis" of the Early Nineteenth Century". *Archive for History of Exact Sciences*, vol. 6, no. 5, pp. 372–400. 315
23 H. Freudenthal (1970–1971). "Did Cauchy Plagiarize Bolzano?" *Archive for History of Exact Sciences*, vol. 7, no. 5, pp. 375–392. 344

Mini course on History of Algebra **362**
24 Old Babylonian problems from BM 13901. From E. Robson (2007). "Mesopotamian Mathematics". In: *The Mathematics of Egypt, Mesopotamia, China, India, and Islam: A Sourcebook*. Ed. by V. J. Katz. Princeton and Oxford: Princeton University Press. Chap. 2, pp. 57–186, pp. 104–107. . . 362
25 Islamic algebra. From J. L. Berggren (2007). "Mathematics in Medieval Islam". In: *The Mathematics of Egypt, Mesopotamia, China, India, and Islam: A Sourcebook*. Ed. by V. J. Katz. Princeton and Oxford: Princeton University Press. Chap. 5, pp. 515–675, pp. 542-545. 366
26 E. Robson (2005). "Influence, ignorance, or indifference? Rethinking the relationship between Babylonian and Greek mathematics". *Bulletin of the British Society for the History of Mathematics*, vol. 4, pp. 1–17. 370

27 R. Franci (2010). "The history of algebra in Italy in the 14th and 15th centuries. Some remarks on recent historiography". *Actes d'història de la ciència i de la tècnica*, vol. 3, no. 2, pp. 175–194. DOI: `10.2436/20.2006.01.158`. 387
28 B. M. Kiernan (1971). "The Development of Galois Theory from Lagrange to Artin". *Archive for History of Exact Sciences*, vol. 8, no. 1–2, pp. 40–154, pp. 40–66. 407
29 Abel's 1824 proof. From P. Pesic (2003). *Abel's Proof. An Essay on the Sources and Meaning of Mathematical Unsolvability*. Cambridge (MA)/London: The MIT Press, pp. 155–169. 434
30 J. Lützen (2009). "Why was Wantzel Overlooked for a Century? The Changing Importance of an Impossibility Result". *Historia Mathematica*, vol. 36, no. 3, pp. 374–394. 442
31 P. M. Neumann (1999). "What groups were: A study of the development of the axiomatics of group theory". *Bull. Austral. Math. Soc.* Vol. 60, pp. 285–301. 463

Text 8: The Method of Archimedes. From T. L. Heath, ed. (1953). *The Works of Archimedes with the Method of Archimedes*. New York: Dover Publications, pp. 12–21. Adopted from J. Lützen and K. Ramskov, eds. (1999). *Kilder til matematikkens historie*. 2nd ed. København: Matematisk Afdeling, Københavns Universitet, pp. 22–26.

The Method of Archimedes treating of Mechanical Problems — to Eratosthenes

Archimedes to Eratosthenes greeting.

I sent you on a former occasion some of the theorems discovered by me, merely writing out the enunciations and inviting you to discover the proofs, which at the moment I did not give. [. . .]

Archimedes then describes some theorems that he has found and mentions that he has included the proofs. He continues

[. . .] Seeing moreover in you, as I say, an earnest student, a man of considerable eminence in philosophy, and an admirer [of mathematical inquiry], I thought fit to write out for you and explain in detail in the same book the peculiarity of a certain method, by which it will be possible for you to get a start to enable you to investigate some of the problems in mathematics by means of mechanics. This procedure is, I am persuaded, no less useful even for the proof of the theorems themselves; for certain things first became clear to me by a mechanical method, although they had to be demonstrated by geometry afterwards because their investigation by the said method did not furnish an actual demonstration. But it is of course easier, when we have previously acquired, by the method, some knowledge of the questions, to supply the proof than it is to find it without any previous knowledge. This is a reason why, in the case of the theorems the proof of which Eudoxus was the first to discover, namely that the cone is a third part

Text 8: The Method of Archimedes. From T. L. Heath, ed. (1953). *The Works of Archimedes with the Method of Archimedes*. New York: Dover Publications, pp. 12–21. Adopted from J. Lützen and K. Ramskov, eds. (1999). *Kilder til matematikkens historie*. 2nd ed. København: Matematisk Afdeling, Københavns Universitet, pp. 22–26.

of the cylinder, and the pyramid of the prism, having the same base and equal height, we should give no small share of the credit to Democritus who was the first to make the assertion with regard to the said figure* though he did not prove it. I am myself in the position of having first made the discovery of the theorem now to be published [by the method indicated], and I deem it necessary to expound the method partly because I have already spoken of it† and I do not want to be thought to have uttered vain words, but equally because I am persuaded that it will be of no little service to mathematics; for I apprehend that some, either of my contemporaries or of my successors, will, by means of the method when once established, be able to discover other theorems in addition, which have not yet occurred to me.

This is followed by some theorems about centers of gravity and the argument for the above mentioned theorem. This argument concludes with the following remark:

Now the fact here stated is not actually demonstrated by the argument used; but that argument has given a sort of indication that the conclusion is true. Seeing then that the theorem is not demonstrated, but at the same time suspecting that the conclusion is true, we shall have recourse to the geometrical demonstration which I myself discovered and have already published.‡

Proposition 2

We can investigate by the same method the proposition that

(1) *Any sphere is (in respect of solid content) four times the cone with base equal to a great circle of the sphere and height equal to its radius; and*

(2) *the cylinder with base equal to a great circle of the sphere and height equal to the diameter is* $1\frac{1}{2}$ *times the sphere.*

(1) Let $ABCD$ be a great circle of a sphere, and AC, BD diameters at right angles to one another.

Let a circle be drawn about BD as diameter and in a plane perpendicular to AC, and on this circle as base let a cone be described with A as vertex. Let the

*περὶ τοῦ εἰρημένου σχήματος, in the singular. Possibly Archimedes may have thought of the case of the pyramid as being the more fundamental and as really involving that of the cone. Or perhaps "figure" may be intended for "type of figure."

†Cf. Preface to *Quadrature of Parabola.*

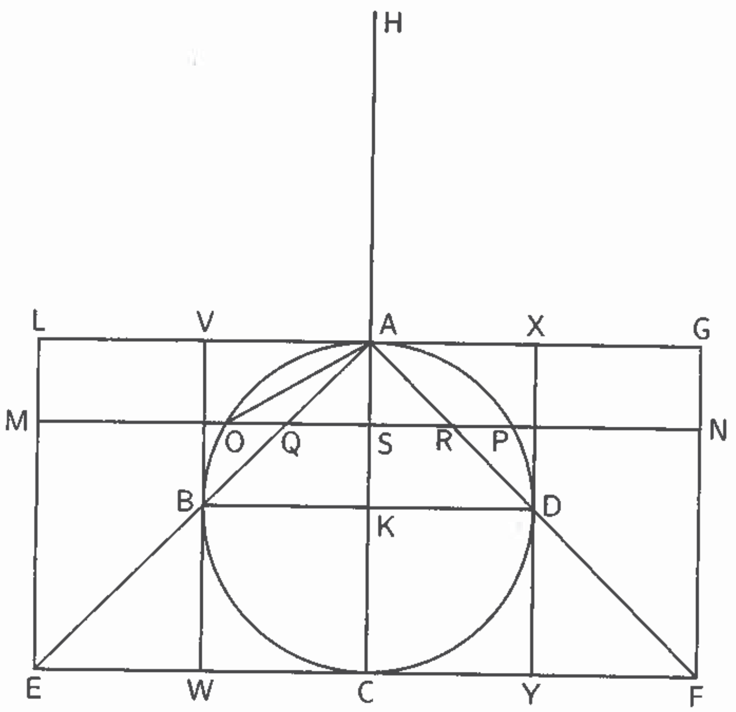
‡The word governing τὴν γεωμετρουμένην ἀπόδειξιν in the Greek text is τάξομεν, a reading which seems to be doutbful and is certainly difficult to translate. Heiberg translates as if τάξομεν meant "we shall give lower down" or "later on", but I agree with Th. Reinach (*Revue générale des sciences pures et appliquées*, 30 November 1907, p. 918) that it is questionable whether Archimedes would really have written out in full once more, as an appendix a proof which, as he says, had already been published (i.e. presumably in the *Quadrature of a Parabola*). τάξομεν, if correct, should apparently mean "we shall appoint", "prescribe" or "assign."

surface of this cone be produced and then cut by a plane through C parallel to its base; the section will be a circle on EF as diameter. On this circle as base let a cylinder be erected with height and axis AC, and produce CA to H, making AH equal to CA.

Let CH be regarded as the bar of a balance, A being its middle point.

Draw any straight line MN in the plane of the circle $ABCD$ and parallel to BD. Let MN meet the circle in O, P, the diameter AC in S, and the straight lines AE, AF in Q, R respectively. Join AO.

Through MN draw a plane at right angles to AC; this plane will cut the cylinder in a circle with diameter MN, the sphere in a circle with diameter OP, and the cone in a circle with diameter QR.



Now, since $MS = AC$, and $QS = AS$,

$$\begin{aligned} MS.SQ &= CA.AS \\ &= AO^2 \\ &= OS^2 + SQ^2. \end{aligned}$$

And $HA = AC$,

$$\begin{aligned} HA : AS &= CA : AS \\ &= MS : SQ \\ &= MS^2 : MS.SQ \\ &= MS^2 : (OS^2 + SQ^2) \text{ from above} \\ &= MN^2 : (OP^2 + QR^2) \\ &= (\text{circle, diam. } MN) : (\text{circle, diam. } OP + \text{circle, diam. } QR). \end{aligned}$$

That is, $HA : AS =$ (circle in cylinder):(circle in sphere + circle in cone).

Therefore the circle in the cylinder, placed where it is, is in equilibrium, about A, with the circle in the sphere together with the circle in the cone, if both the latter circles are placed with their centres of gravity at H.

Similarly for the three corresponding sections made by a plane perpendicular to AC and passing through any other straight line in the parallelogram LF parallel to EF.

If we deal in the same way with all the sets of three circles in which planes perpendicular to AC cut the cylinder, the sphere and the cone, and which make up those solids respectively, it follows that the cylinder, in the place where it is, will be in equilibrium about A with the sphere and the cone together, when both are placed with their centres of gravity at H.

Therefore, since K is the centre of gravity of the cylinder,

$$HA : AK = (\text{cylinder}) : (\text{sphere} + \text{cone } AEF).$$

But $HA = 2AK$; therefore

$$\text{cylinder} = 2\ (\text{sphere} + \text{cone } AEF).$$

Now

$$\text{cylinder} = 3\,(\text{cone } AEF); \qquad \text{[Eucl. XII.10]}$$

therefore

$$\text{cone } AEF = 2\,(\text{sphere}).$$

But, since $EF = 2BD$,

$$\text{cone } AEF = 8\,(\text{cone } ABD);$$

therefore

$$\text{sphere} = 4\,(\text{cone } ABD).$$

(2) Through B, D draw VBW, XDY parallel to AC; and imagine a cylinder which has AC for axis and the circles on VX, WY as diameters for bases.

Then

$$\begin{aligned} \text{cylinder } VY &= 2\,(\text{cylinder } VD) \\ &= 6\,(\text{cone } ABD) \quad \text{[Eucl. XII.10]} \\ &= \tfrac{3}{2}\,(\text{sphere}), \quad \text{from above.} \end{aligned}$$

Q.E.D.

From this theorem, to the effect that a sphere is four times as great as the cone with a great circle of the sphere as base and with height equal to the radius of the sphere, I conceived the notion that the surface of any sphere is four times as great as a great circle in it; for, judging from the fact that any circle is equal to a triangle, with base equal to the circumference and height equal to the radius of the circle, I apprehended that, in like manner, any sphere is equal to a cone with base equal to the surface of the sphere and height equal to the radius.[§]

Text 9: K. M. Pedersen (1980). "Techniques of the Calculus, 1630–1660". In: *From the Calculus to Set Theory, 1630–1910. An Introductory History*. Ed. by I. Grattan-Guinness. Princeton and Oxford: Princeton University Press. Chap. 1, pp. 10–48.

Chapter 1

Techniques of the Calculus, 1630–1660

Kirsti Møller Pedersen

1.1. *Introduction*

During the first six decades of the 17th century mathematics was in a state of rapid development. In this period ideas were born and developed which were to be taken up later by Isaac Newton and G. W. Leibniz. Many methods were developed to solve calculus problems ; common to most of them was their *ad hoc* character. It is possible to find examples from the time before Newton and Leibniz which, when translated into modern mathematical language, show that differentiation and integration are inverse procedures ; however, these examples are all related to specific problems and not to general theories. The special merit of Newton and Leibniz was that they both worked out a general theory of the infinitesimal calculus. However, it cannot be said that either Newton or Leibniz gave to his calculus a higher degree of mathematical rigour than their predecessors had done.

As the ideas which were the basis of the methods preceding the work of Newton and Leibniz came to bear fruit, the methods themselves fell into oblivion. In this chapter, therefore, great importance will be attached to the earlier ideas, and the methods will be illustrated by simple examples. The picture of what the mathematicians of the time achieved may thus appear somewhat distorted, but a rendering of the more complicated examples would be all too easily submerged in calculations. That it is possible to find simple problems is due to the fact that it was the practice of the mathematicians of the time to verify their methods by applying them to problems of which the solutions were known beforehand. Then the next step was to find new results by means of these methods.

It is impossible to deal comprehensively with this topic in a single chapter. My approach will be to exemplify the calculus of the period by relatively few methods, which are described in some detail. This implies that the methods of many important mathematicians will have to be left unmentioned. A more general survey giving a more profound impression of the development of the calculus from 1630 to 1660 may be found in the rich literature on this subject.[1] I have made my choice on the assumption that to give even a tolerably satisfactory general survey in a single chapter would mean listing names and outlining techniques in a way which could not possibly give a proper impression of the methods and style of the time to a reader who is not acquainted with the period.

One criterion for the selection of methods has been that they should render a picture of the way in which the mathematicians of the time did actually solve the problems with which they were most heavily engaged; another has been that they should inform the reader of the ideas which were to become sources of inspiration for later methods. Where different methods are based on similar ideas, I have tried to select the writer who first formulated the idea.

Of the period 1630–1660, no less than of all other periods, it holds true that if you really want to set its mathematics into relief then you must know the mathematics which preceded it. The mathematics of the period in question were greatly influenced by classical Greek mathematics[2] and also by that of the previous period. The reason for the importance of Greek mathematics was that during the 16th century it had become usual for the mathematicians to acquire a knowledge of this discipline, and it formed a basic element in the mathematical equipment of most of them. Greek mathematics was especially admired for its great stringency. But its methods were not heuristic; they were not well-fitted to suggest ideas as to how to attack a new problem, a fact which will be illustrated later in connection with quadratures and cubatures.

It was natural, therefore, to search for other methods which, if they could not live up to the Greek requirement of exactness, were at least able to suggest ideas as to the solution of problems. The seeds of such methods are to be found in the previous period, the end of the 16th and the beginning of the 17th centuries, which was a fertile time for the exact sciences as a whole. Astronomy made great progress through the work of Johannes Kepler; Simon Stevin contributed much to statics with his treatise *De Beghinselen der Weeghconst* ('The elements of the art of weighing': *1586a*). In mechanics Galileo Galilei's deduction of the laws of freely falling bodies and of the parabolic paths of projectiles meant a break with Aristotelian physics and the beginning of a new epoch, where mathematics was to be extensively used in physics.

[1] See, for example, Baron *1969a*, Boyer *1939a* and Whiteside *1961a*, and their bibliographies.

[2] There are excellent bibliographies of Greek mathematics in Boyer *1968a* and Kline *1972a*.

Kepler made use of infinitesimal methods in his works. The interest he took in estimating the volumes of wine casks resulted in the book *Nova stereometria doliorum vinariorum* ('New measurement of large wine casks': *1615a*). There he considered solids of revolution as composed in various ways of infinitely many constituent solids. For example, he regarded a sphere as made up of an infinite number of cones with vertices at the centre and bases on the surface of the sphere. This led to the result that the sphere is equal in volume to the cone which has the radius of the sphere as altitude and as base a circle equal to the surface of the sphere, that is, a circle with the diameter of the sphere as radius (Kepler *1615a*, Prima Pars, Theorem 11; *Works*$_1$, vol. 4, 563, or *Works*$_2$, vol. 9, 23 f.).

Galileo planned to write a book on indivisibles, but this book never appeared; however, his ideas had a great influence on his pupil Cavalieri, with whose work we shall deal later.

1.2. *Mathematicians and their society*

A great many mathematicians of the 17th century were not mathematicians by profession. This tendency was especially noticeable in France; there only Gilles Personne de Roberval occupied a chair of mathematics, while great mathematicians like Pierre de Fermat, René Descartes and Blaise Pascal worked without any official connection with their discipline. Like the mathematician who inspired him, François Viète, Fermat was a lawyer, and worked as such in Toulouse for most of his career. Descartes and Pascal were men of private means and, apart from mathematics, were also occupied with physics and philosophy. Descartes spent a large part of his time outside France, living for long periods in Holland and elsewhere.

This stay of Descartes in Holland served to inspire several Dutch mathematicians, among whom was Frans van Schooten. He was a member of the School of Engineering at Leyden, while his more important pupils, whose treatises he published along with his own, mostly worked professionally outside mathematics. However, the most illustrious of his pupils, Christiaan Huygens, devoted his whole life to mathematics and physics. In 1666 the *Académie des Sciences* was founded in Paris, and Huygens was offered a membership which he accepted. As a member of the *Académie* he received an ample stipend. In Italy, the most outstanding mathematicians and physicists, such as Galileo Galilei, Bonaventura Cavalieri and Evangelista Torricelli, held offices within their own fields, partly at universities and partly as court mathematicians.

The development of that part of mathematics with which this chapter is concerned started later in England than on the Continent. Hence the only English mathematician with whom we shall deal in this chapter is John Wallis, who was Savilian Professor of Geometry at Oxford from 1649. It should be mentioned that in Thomas Harriot England had a brilliant scientist whose work both in algebra and the calculus preceded some of the methods discussed in this chapter. But only his *Artis analyticae praxis* (' Practice of the analytical art ' : *1631a*), which contains his less important work, was published (posthumously) at this time ; thus his unpublished results will not be considered.

The period provides several good examples of the independent and almost simultaneous discovery of methods with striking resemblance, which often gave rise to disputes about priority and charges of plagiarism. Today, we are able to establish that as a rule these charges were unfounded ; but at the time this was not possible, since it was not common to publish one's treatises. For this there were two principal reasons. First, after 1640 publishers were reluctant to print mathematical literature, which was not very profitable ; and second, mathematicians were reticent about publishing their new methods, wanting to release the results only. Many treatises had to wait a very long time for their publication : several were left unprinted until the end of the 19th and the beginning of the 20th centuries, and some remain unpublished to this day.

Not until the last third of the 17th century did scientific periodicals come into existence ; before that time mathematicians communicated by letter. Here the Frenchman Marin Mersenne played an important part, for he kept in touch with many European scientists by correspondence and meetings which he held at his convent in Paris. To the mathematicians he sent the problems which he could not solve himself, and took care that the results and manuscripts he received were circulated among those interested in them.

1.3. *Geometrical curves and associated problems*

In the 17th century the calculus was closely bound up with the investigation of curves, since there was as yet no explicit concept of the variable or of functional relationships between variables. The first curves to be dealt with were those inherited from the Greeks : the conic sections, Hippias's quadratrix, the Archimedean spiral, the conchoid of Nicomedes, and the cissoid of Diocles. (For the definition and the history of these and the following curves see, for example, Loria *1902a*.)

As the century went on, these curves were augmented by, among others, the cycloid, the higher parabolas and hyperbolas ($y^m = kx^n$ and $ky^m x^n = 1$ respectively, m and n being natural numbers and k a constant), the spiral of Galileo, and the conchoid to a circle, also termed ' the limaçon of [Etienne] Pascal ', which is in turn a variant of the curves called ' the ovals of Descartes '.

Next to the conic sections the cycloid, the curve traced by a point on the circumference of a circle which rolls along a horizontal line, was the curve most often investigated. Its early history is connected with a problem called ' Aristotle's wheel ' (see Drabkin *1950a*). When solving this problem Roberval generalised the motion which generates the curve, and considered the curtate and the prolate cycloid (which are traced by points on a radius and respectively outside and inside the circle) as well as the ordinary cycloid. In 1658 Blaise Pascal arranged a competition designed to find the area of a section of the cycloid, its centre of gravity, the volumes of solids obtained by revolving the section about certain axes, and the centres of gravity of these volumes (Pascal *1658a* and *1658b*).

In *La géométrie* (*1637a*) Descartes introduced his oval as a curve involved in the solution of various optical problems. One of these problems was to determine the form of a lens which makes all the rays that come from a single point or that are parallel converge at another unique point, after having passed through the lens (Descartes *1637a*, 362 ; *1925a*, 135).

Similarly, Galileo's spiral was the attempted solution of a physical problem concerning the path of a body which moves uniformly around a centre and at the same time descends towards the centre with constant acceleration. The recognition of the shape of another of Galileo's curves, namely, the catenary, caused the mathematicians many difficulties. This curve has the form of a chain suspended from two points (see section 2.8).

The three last-mentioned curves are examples of an interplay between physics and mathematics. Before discussing this topic further we shall answer the question : what kind of problems concerning curves did the mathematicians solve in the period before 1660 ?

Pascal's competition of 1658 relates to certain typical problems which were solved. Other problems consisted in finding tangents, surface areas and extreme values ; furthermore, some inverse tangent problems (that is, to find a curve which has tangents with a specific property) were considered. Finally, about the middle of the century, the rectification of arcs became a question of interest. Although there are earlier examples of rectifications, Christopher Wren's rectification of the cycloidal arc in the late 1650s was the first widely known one. He sent the result to Pascal outside the competition (see Wren *1659a*, or Wallis *Works*, vol. 1, 532–541).

Even though the solutions to these problems could be applied both to physics and to astronomy, their inspiration owed more to Greek mathematics than to physics and astronomy. The Greeks had worked on all the types of problem mentioned above ; one may therefore consider work on them as a continuation of the tradition of the Greek mathematicians. This does not mean that there was no correlation between mathematics and physics. This continued to happen, if for no other reason than that in this period important physicists were often also important mathematicians. It is nevertheless difficult to point unambiguously to a concrete physical problem which inspired the mathematicians to take up the above-mentioned problems. In the late 1650s, however, a new mathematical problem cropped up which sprang from physics, namely the study of evolutes, which was started by Huygens in connection with his work on the pendulum clock.

1.4. *Algebra and geometry*

When the Greeks came to realise the existence of incommensurable magnitudes, which meant that the rational numbers are not sufficient for purposes of measurement, they made geometry the foundation of that part of mathematics which was not number theory, the straight line being a substitute for a continuous field of numbers. This attitude resulted in the geometric algebra on which Euclid, Archimedes and Apollonius based their calculations.

In the course of time the theory of equations became separated from geometry, and a good deal of symbolism was gradually developed for this discipline. Viète contributed much to the introduction of symbols with his work *In artem analyticen isagoge* (' Introduction to the analytic art ' : *1591a*), in which he emphasised the advantage of using symbols to indicate not only unknown but also known quantities (Viète *1591a*, ch. V, 5 ; *Works*, 8, or *1973a*, 52). In this way he could deal with equations in general.

Viète also connected algebra and geometry by determining the equations which correspond to various geometrical constructions. He only employed this technique when the geometrical problems were determinate and led to determinate equations in one unknown quantity. The next step was to use an indeterminate equation in two unknown quantities when solving problems concerning geometric loci. Fermat and Descartes took this step almost simultaneously.

Fermat's treatise *Ad locos planos et solidos isagoge* (' Introduction to plane and solid loci ' : *1637a*) contains a pedagogic introduction to analytic geometry and some of its applications. However, the treatise did not have any great influence, for the simple reason that Descartes's *La géométrie* was published before it was generally known. *La géométrie* treats many subjects with supreme skill, but it starts with an introduction to analytic geometry that was not easy for the uninitiated to follow. Notwithstanding this fact, the work had a tremendous influence, especially after van Schooten had published it in Latin translation and with commentaries in 1659. Its success was mainly due to Descartes's notation, which bore the hallmark of genius. It will not surprise the modern reader, as it is the beginning of the notation still in use ; but for the time it was revolutionary. There is no doubt that the notation and the thoughts embodied in *La géométrie* had a positive—if only indirect—influence on the development of the calculus.

1.5. *Descartes's method of determining the normal, and Hudde's rule*

In *La géométrie* Descartes described his technique of determining the normal to an algebraic curve at any point. He attached great importance to the method, as can be seen from the following introductory remarks (*1637a*, 341 ; *1925a*, 95) :

> This is my reason for believing that I shall have given here a sufficient introduction to the study of curves when I have given a general method of drawing a straight line making right angles with a curve at an arbitrarily chosen point upon it. And I dare say that this is not only the most useful and most general problem in geometry that I know, but even that I ever desired to know.

Let the algebraic curve ACE be given and let it be required to draw the normal to the curve at C (see figure 1.5.1). Descartes supposed the line CP to be the solution of the problem. Let $CM=x$, $AM=y$,

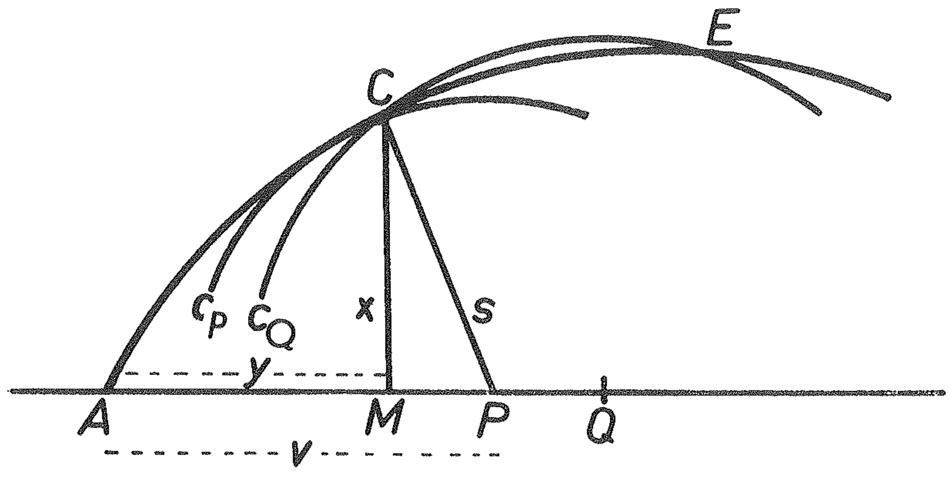


Figure 1.5.1.

$AP=v$ and $CP=s$. Although he always used a particular example, for the sake of convenience we shall suppose the curve to have the following equation :

$$x=f(y). \tag{1.5.1}$$

We shall also modernise his notation to some extent.

Besides the curve, Descartes considered the circle c_P with centre at P and passing through C ; that is, the circle with the equation

$$x^2+(v-y)^2=s^2. \tag{1.5.2}$$

This circle will touch the curve CE at C without cutting it, whereas the circle c_Q

$$x^2+(v_Q-y)^2=s_Q{}^2 \tag{1.5.3}$$

with centre at a point Q different from P and passing through C will cut the curve not only at C but also in another point. Let this point be E. This means that the equation obtained by eliminating x from (1.5.1) and (1.5.3),

$$(f(y))^2+(v_Q-y)^2-s_Q{}^2=0, \tag{1.5.4}$$

has two distinct roots ;[1] but ' the more C and E approach each other, the smaller the difference of the two roots, and at last, when the points coincide, the roots are exactly equal, that is to say when the circle through C touches the curve at the point C without cutting it ' (Descartes *1637a*, 346–347 ; *1925a*, 103–104).

Thus the analysis has brought Descartes to the conclusion that CP will be a normal to the curve at C when P (that is, v) is so determined that the equation

$$(f(y))^2+(v-y)^2-s^2=0 \tag{1.5.5}$$

has two roots equal to y_0 (or the corresponding equation with y eliminated has one pair of equal roots). With modern conceptions it is not difficult to realise that this requirement gives the correct expression,

$$v-y_0=f'(y_0)\cdot f(y_0), \tag{1.5.6}$$

for the sub-normal MP.

Descartes illustrated his method by finding, among other things, the normal to the ellipse (*1637a*, 347 ; *1925a*, 104). Putting its equation in the form

$$x^2=ry-\frac{r}{q}y^2, \tag{1.5.7}$$

he found the equation corresponding to (1.5.5) to be

[1] Descartes only considered curves for which $(f(y))^2$ is a polynomial in y or y^2 a polynomial in x.

$$y^2+\left(\frac{rq-2vq}{q-r}\right)y+\frac{qv^2-qs^2}{q-r}=0. \tag{1.5.8}$$

This equation has two roots equal to y_0 when

$$\frac{rq-2vq}{q-r}=-2y_0 \quad\text{and}\quad \frac{qv^2-qs^2}{q-r}=y_0{}^2; \tag{1.5.9}$$

because the point C is given, the value y_0 is known, and from (1.5.9) the sub-normal $v-y_0$ can be determined:

$$v-y_0=\frac{r}{2}-\frac{r}{q}y_0. \tag{1.5.10}$$

Although an indication, not to say a full account, of what happens when the two points C and E coincide would involve limit-considerations,[1] Descartes, by taking the double contact of the circle with the curve as a characteristic of the normal, has avoided the use of infinitesimals and obtained an algebraic method. His correspondence indicates that in solving some of his problems he did employ methods which involved the use of infinitesimals. However, he did not consider them precise enough to be published.

In principle, Descartes's method is applicable to any algebraic curve. But when the equation of the curve is not a simple algebraic equation, the method becomes tedious because of the laborious calculations which it is necessary to carry out in order to determine v by comparing the coefficients.

The Dutch mathematician (later Burgomaster of Amsterdam) Johann Hudde invented a rule for determining double roots. He described his method in a letter to Frans van Schooten, who published it in his 1659 Latin edition of Descartes's *La géométrie* (Hudde *1659a*, 507):

> If in an equation two roots are equal, and if the equation is multiplied by any arithmetical progression in such a way that the first term of the equation is multiplied by the first term of the progression and so on, I say that the product will be an equation in which the given root is found again.

[1] If we let the coordinates of E be $(y_0+\Delta y, f(y_0+\Delta y))$, then the requirement that C and E be on the same circle with centre at Q on the axis gives us the condition:

$$AQ=y_0+\frac{\Delta y}{2}+\left(\frac{f(y_0+\Delta y)-f(y_0)}{\Delta y}\right)\left(\frac{f(y_0+\Delta y)+f(y_0)}{2}\right).$$

(To obtain this result, let F be the mid-point of CE and note that $QF\perp CE$.) P and v are then determined by the coincidence of the points C and E, that is:

$$v=AP=\lim_{\Delta y\to 0} AQ=f'(y_0)f(y_0)+y_0.$$

For this rule Hudde gave a proof which in modern notation may be rendered as follows. Let $x=x_0$ be a double root in the polynomial $p(x)$, that is,

$$p(x)=(x-x_0)^2 \sum_{i=0}^{n} \alpha_i x^i$$
$$= \sum_{i=0}^{n} \alpha_i (x^{i+2}-2x_0x^{i+1}+x_0{}^2x^i), \qquad (1.5.11)$$

and let a, $a+d, \ldots, a+(n+2)d$ be an arbitrary arithmetical progression. We then multiply the constant term $\alpha_0 x_0{}^2$ in $p(x)$ by a, the term of the first degree by $a+d$, and so on. Let the result of this procedure be denoted by $(p(x), a, d)$; that is,

$$(p(x), a, d)= \sum_{i=0}^{n} \alpha_i\{(a+(i+2)d)x^{i+2}-2(a+(i+1)d)x_0x^{i+1}$$
$$+(a+id)x_0{}^2x^i\}. \qquad (1.5.12)$$

(Note that

$$(p(x), a, d)=ap(x)+dxp'(x), \qquad (1.5.13)$$

where $p'(x)$ is the derivative of $p(x)$ and ' dx ' means $d\times x$.) If we put $x_0=x$, the expression in curled brackets in (1.5.12) vanishes. We therefore have $(p(x_0), a, d)=0$.

This necessary condition for a polynomial to have one pair of equal roots made Descartes's method easier to apply, because one might so arrange the arithmetical progression that a difficult term might be multiplied by 0. We see that in his studies in autumn 1664 Newton found the sub-normal to a curve by using a combination of Descartes's method and Hudde's rule (Newton *Papers*, vol. 1, 217 ff.).

Hudde applied his rule to the determination of extreme values, acting on the assumption that if α is a value which makes $p(x)$ extreme, then the equation $p(x)=p(\alpha)$ has two equal roots (see Haas *1956a*, 250–255). He also extended his procedure to a rule for determining sub-tangents (*1659b*). He did not prove this rule, but it is interesting because it is one of the first general rules. Let the equation of the curve be $p(x, y)=0$, where p is a polynomial in x and y ; Hudde's rule then states that the sub-tangent t to a point (x, y) is given by

$$t=\frac{-x(p(x, y), a, d)_y}{(p(x, y), a, d)_x}. \qquad (1.5.14)$$

The subscripts mean that in the numerator $p(x, y)$ is to be considered as a polynomial in y and in the denominator as a polynomial in x. From (1.5.13) we have

$$t = \frac{-x(ap(x, y) + dyp_y'(x, y))}{ap(x, y) + dxp_x'(x, y)} \tag{1.5.15}$$

(where the prime indicates differentiation with respect to the subscript variable), or, since $p(x, y) = 0$,

$$t = \frac{-yp_y'(x, y)}{p_x'(x, y)}. \tag{1.5.16}$$

Hudde's method was not forgotten after the introduction of the differential calculus; for example, l'Hôpital commented on it in his *1696a*, ch. 10, para. 192 (see also section 2.5 below).

1.6. *Roberval's method of tangents*

In the late 1630s Gilles Personne de Roberval and Evangelista Torricelli independently found a method of tangents which used arguments from kinematics. In 1644, in his *Opera geometrica*, Torricelli published an application of his method to the parabola (Torricelli *1644a*, 119–121; *Works*, vol. 2, 122–124). In the same year Mersenne, in his *Cogitata physico mathematica* ('Physico-mathematical thoughts'), mentioned Roberval's method and applied it also to the parabola (Mersenne *1644a*, 115–116; see Jacoli *1875a*). One of Roberval's pupils, François du Verdus, wrote a treatise on Roberval's method. It was eventually published in 1693 (Roberval *Observations*) and became quite well-known, so the kinematic method came to bear Roberval's name.

The method rests on two basic ideas. The first is to consider a curve as the path of a moving point which is simultaneously impressed by two motions. The second is to consider the tangent at a given point as the direction of motion at that very point. If the two generating motions are independent, then the direction of the resultant motion is found by the parallelogram law for compounding motions. However, Roberval also applied his method to curves like the quadratrix and the cissoid, where the generating motions which he considered were dependent. He ingeniously compensated for the dependence when compounding the motions, as we shall see.

Roberval succeeded in determining the correct tangents to all the curves which were generally considered at his time. For the conic sections, however, the tangents were not determined correctly, because he took the generating motions to be the motions away from the foci or from the focus and the directrix, and wrongly used the parallelogram rule in compounding these motions (see Pedersen *1968a*, 165 ff.).

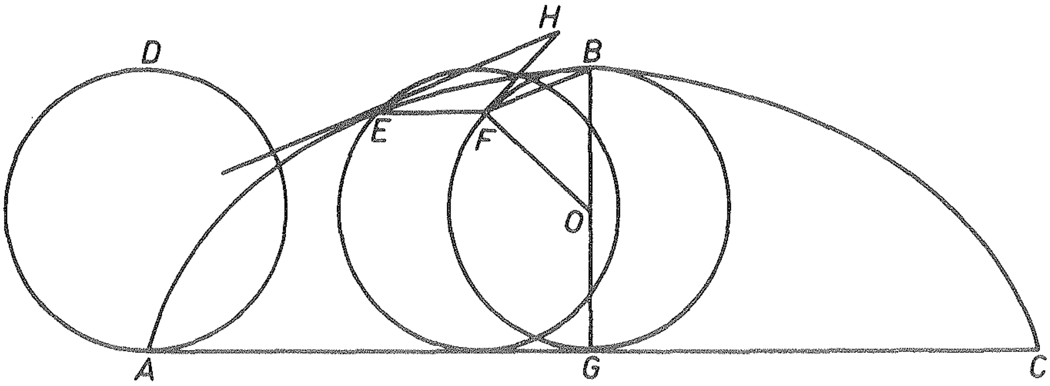


Figure 1.6.1.

To illustrate the method, we shall first see how Roberval determined the tangents to the cycloids (Roberval *Works*$_2a$, 58–63). Let ABC be a cycloid generated by the circle AD; that is, ABC is the path of the point A when the circle makes one turn on the line AC (compare figure 1.6.1, where the ordinary cycloid is drawn). The motion of A is then compounded of a uniform motion with direction AC or EF, and a uniform rotation about the centre of the generating circle, the direction of this at a point E being the tangent to the generating circle at E or the line FH. The ratio between the speeds of these motions is equal to the ratio between AC and the perimeter ADA, so if the point H is determined by

$$EF : FH = AC : \text{perimeter } ADA, \tag{1.6.1}$$

then EH will be the tangent to the cycloid at E. For the ordinary cycloid, the ratio on the right hand side is equal to unity, and Roberval proved geometrically that EH is parallel to FB.

Thus the method is easily applied to the cycloid; but to see how general it is, let us also consider Roberval's determination of the tangent to the quadratrix. In figure 1.6.2 we let the two sides AD and CD of a square $ABCD$ move simultaneously, AD being rotated uniformly about A and CD being paralleledly displaced in such a way that AD and CD coincide with AB at the same time. The point of intersection between the two lines will then describe a quadratrix DFH. Let F—the point of intersection between IN and AD_1—be one of the points of the quadratrix and let us see how he determines the tangent at F. (Actually he considers a point on DFH's prolongation, but the principle is the same.)

Roberval starts by letting the line FK represent the velocity of the line IN. From the definition of the quadratrix follows that F describes the line FK in the same time as D_1 describes the arc D_1B, whence arc D_1B represents the speed of D_1's circular motion. As

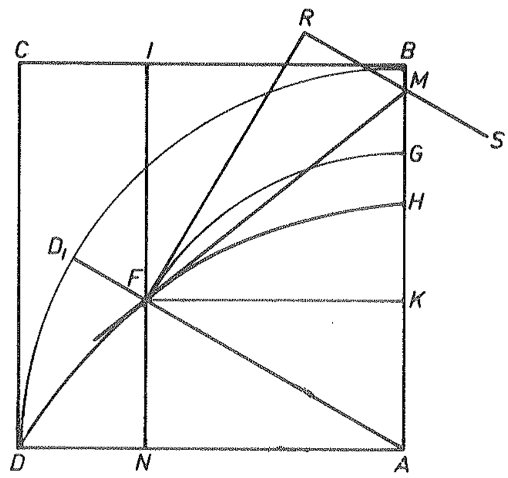


Figure 1.6.2.

$$(\text{the speed of the circular motion of } F) : (\text{the speed of the circular motion of } D_1) = AF : AD_1 = \text{arc } FG : \text{arc } D_1B, \quad (1.6.2)$$

the arc FG represents the speed of F's circular motion ; and further, as the direction of this latter motion is perpendicular to AF, the circular motion of F will be represented by the line-segment FR on the perpendicular with length equal to arc FG. To obtain F's direction of movement he then draws the line RS through R parallel to AF and seeks the point of intersection, M, between RS and AB (which is the line through K parallel to IF) and connects F and M. FM will then be the tangent.

Roberval used this general approach in other cases too. His argument for it is not quite clear, but it has a great deal in common with the following. F's motion can be considered in two ways :

(1) F's motion on the quadratrix is compounded of the motion F has by taking part in AF's motion (with the instantaneous velocity FR) and the motion F has on AF because it has to be the point of intersection; the direction of the last motion is AF or RS. By compounding these two motions we see that the line of direction of the movement of F starts at F and ends on the line RS.

(2) Similarly, it is realised, by compounding the motion F has when it takes part in the motion of IF with its motion on IF, that its direction of motion is a line starting at F and ending on AB.

As both the conclusion of (1) and (2) must be fulfilled, the above construction follows.

By taking the instantaneous direction of motion as known, Roberval and Torricelli had avoided the use of infinitesimals in their method. Their method had the further advantage of being applicable to curves which are not referred to a Cartesian coordinate system. The method, however, was not general as long as the velocities could not be generally determined.

It is interesting to note that Newton's method of tangents from 1666 is inspired by the same ideas as Roberval's. For algebraic curves Newton only had to use the method once to obtain the sub-tangent expressed by a formula ; but for transcendental curves like the quadratrix he found the tangent in almost the same manner as had Roberval (Newton *Papers*, vol. 1, 416–418).

1.7. *Fermat's method of maxima and minima*

About 1636 there was circulated among the French mathematicians a memoir of Fermat entitled *Methodus ad disquirendam maximam et minimam* (' Method of investigating maxima and minima ' : *Methodus*). It was remarkable, for it gave the first known general method of determining extreme values. It contained another striking feature, namely, the idea of giving an increment to a magnitude, which we might interpret as the independent variable.

The memoir opens with the sentence : ' The entire theory of determining maxima and minima is based on two positions expressed in symbols and this single rule '. The rule is the following :

I. Let A be a term related to the problem ;
II. The maximum or minimum quantity is expressed in terms containing powers of A ;
III. A is replaced by $A+E$, and the maximum or minimum is then expressed in terms involving powers of A and E ;
IV. The two expressions of the maximum or minimum are made ' adequal ', which means something like ' as nearly equal as possible ' ;[1]
V. Common terms are removed ;
VI. All terms are divided by a power of E, so that at least one term does not contain E ;
VII. The terms which still contain E are ignored ;
VIII. The rest are made equal.

The solution of the last equation will give the value of A which makes the expression take an extreme value. Fermat illustrated his method by finding the point E on the line-segment AC which makes the rectangle $AE \,.\, EC$ a maximum. Let $AC=b$ and let us replace Fermat's A by x (so that $AE=x$), and his E by e ; we then have to

[1] Fermat used the word ' adaequo '. Mahoney has translated this as ' set adequal ' (*1973a*, 162). The idea of adequality derives from Diophantus (*ibid.*, 163–165).

maximize the expression $x(b-x)$. In accordance with the method, we have

$$(x+e)(b-(x+e)) \approx x(b-x), \tag{1.7.1}$$

where $\approx$ signifies the adequality. Removing common terms, we have

$$be \approx 2xe + e^2, \tag{1.7.2}$$

and dividing by e,

$$b \approx 2x + e. \tag{1.7.3}$$

Finally we ignore the term e and obtain $b = 2x$.

It is tempting to reproduce Fermat's method by letting $A = x$, $E = \Delta x$, and the quantity $= f(x)$; the rule then tells us

$$\text{IV, V} \qquad f(x+\Delta x) - f(x) \approx 0, \tag{1.7.4}$$

$$\text{VI} \qquad \frac{f(x+\Delta x) - f(x)}{\Delta x} \approx 0, \tag{1.7.5}$$

$$\text{VII, VIII} \qquad \left(\frac{f(x+\Delta x) - f(x)}{\Delta x}\right)_{\Delta x = 0} = 0. \tag{1.7.6}$$

For differentiable functions this might be interpreted in modern terms as if the x which makes $f(x)$ a local extreme value is determined by the equation

$$f'(x) = \lim_{\Delta x \to 0} \left\{\frac{f(x+\Delta x) - f(x)}{\Delta x}\right\} = 0. \tag{1.7.7}$$

However, this would be to read too much into the method. Primarily, Fermat did not think of a quantity as a function. Secondly, he did not say anything about E being an infinitesimal, or even a small magnitude, and the method does not involve any concept of limits ; it is purely algebraic. Thirdly, the statement in VI makes no sense in this interpretation, as we always have to divide by E to the first degree. Nevertheless, his examples show us that on occasion he divided by higher powers of E than one. The reason for this is that, if the quantity contained a square root, he squared the adequality before applying the last steps of the rule. Note that he did not emphasise that his method gave only a necessary condition.

Few results in the history of science have been so closely examined as Fermat's method of maxima and minima. He wrote about a dozen short memoirs where he explained and applied his method. Historians have been puzzled by his very short descriptions, and disagree about the dating of the memoirs and about the order of his ideas. To me it seems probable that he developed his ideas in the way that he intimated in his manuscript ' Syncriseos et anastrophes ' (*Syncriseos* ; see Mahoney *1973a*, 145–165).

Fermat says here that he got the idea of a process for determining extreme values by studying Viète's theory of equations and combining it with the expression ' *μοναχός* ' used by Pappus to characterise a minimal ratio (see Pappus *Collections*, book VII, theorem 61). Fermat takes ' *μοναχὸς* ' to mean ' singular ' in the sense of ' unique ' (see his *Works*, vol. 1, 142, 147), and gives an illustrative example of what he meant. The line-segment of the length B has to be divided by a point so that the product of the segments is maximum. The required point is the midpoint which makes the maximum equal to $B^2/4$. If $Z < B^2/4$, then the equation

$$X(B-X)=Z \tag{1.7.8}$$

will have two roots. Let them be A and E. Following Viète, Fermat obtains

$$A(B-A)=E(B-E) \tag{1.7.9}$$

or

$$BA-BE=A^2-E^2. \tag{1.7.10}$$

By dividing by $A-E$, it is seen that $B=A+E$. The closer that Z approaches $B^2/4$, the smaller will be the difference between A and E; at last, when $Z=B^2/4$, A will be equal to E, and $B=2A$, which is the unique solution leading to the maximum product. In other words, to find the maximum you have to equate the two roots.

As it can be complicated to divide by the binomial $A-E$, Fermat chose to let the two roots be A and $A+E$; then he divided by E, and finally equated the two roots by putting $E=0$. After these considerations he repeated his procedure from *Methodus* sketched in I–VIII at the beginning of this section. In this procedure he did not put $E=0$, but ignored the terms still containing E. However, the process is the same, and it became common practice to put E, or a corresponding magnitude, equal to 0 when his method was applied.

Until it was realised that the important process is

$$\lim_{\Delta x \to 0} \left\{ \frac{f(x+\Delta x)-f(x)}{\Delta x} \right\}, \tag{1.7.11}$$

the procedure that involved dividing by E and putting $E=0$ was a thorn in the mathematicians' side. They were severely criticised for it, and they admitted that it was unsatisfactory.

Huygens who knew, applied and simplified Fermat's method, tried in vain to justify it logically (manuscript from 1652 printed in Huygens *Works*, vol. 12, 61). Instead he found another method, and one of which he could give a proof (*ibid.*, 62 ff.). This method combined Fermat's idea of an extreme value as unique with Descartes's idea of a double-root which he used in his method of normals. Briefly and in modern terms : Let $p(x)$ be a polynomial and let $p(x_0)$ be a maximum ; when $a < p(x_0)$, the equation $p(x) = a$ has two roots which will be equal when $a = p(x_0)$. By a comparison of coefficients, x_0 may then be determined from the relation

$$p(x) - p(x_0) = (x^2 - 2xx_0 + x_0^2)p_1(x), \qquad (1.7.12)$$

where $p_1(x)$ is again a polynomial. As the applicability of this method is very limited, and as it is intricate to use, Huygens admitted that Fermat's method was easier to operate, and he himself accepted it.

Among others, Pierre Brûlart requested Fermat to give a proof of his method. In his answer *1643a* Fermat took another line, considering the coefficients of the powers of E in the development of $f(A \pm E)$. Although he could not prove it rigorously, he made it seem plausible that a maximum or minimum can be determined from the equation obtained by putting the coefficient of E equal to 0. Further, he showed that he understood that the coefficient of E^2 must be smaller than 0 for a maximum and greater for a minimum.

To Fermat it was more important to see that a method worked in practice than to give an exact proof of it. The method of maxima and minima had proved its value, for it gave the correct results when applied to a series of problems. Among these was the determination of the points of inflection of a curve in the manuscript ' Doctrinam tangentium ' (Fermat *Works*, vol. 1, 166–167).

Fermat, however, did not stop at that ; he extended the use of the procedure III–VIII from *Methodus* to other fields. This enabled him to determine tangents to curves (as will be seen in the next section), centres of gravity (*1638a*), and the sine law of refraction (*1662a*).

1.8. *Fermat's method of tangents*

In *Methodus*, Fermat made a determination of the tangent to the parabola, and presented this as an application of his method of maxima and minima. Before discussing the method we shall consider the example (Fermat *Works*, vol. 1, 134–136). Let the parabola DB with axis DC be given as in figure 1.8.1. Fermat wants to find the tangent at B ; suppose it to be BE, and let the sub-tangent be EC. He takes an arbitrary point O on BE and draws IO parallel to the ordinate BC. Let P be a point of intersection of IO with the parabola.

From the inequality $IO > IP$, and from the property of the parabola

$$DC : DI = CB^2 : IP^2, \qquad (1.8.1)$$

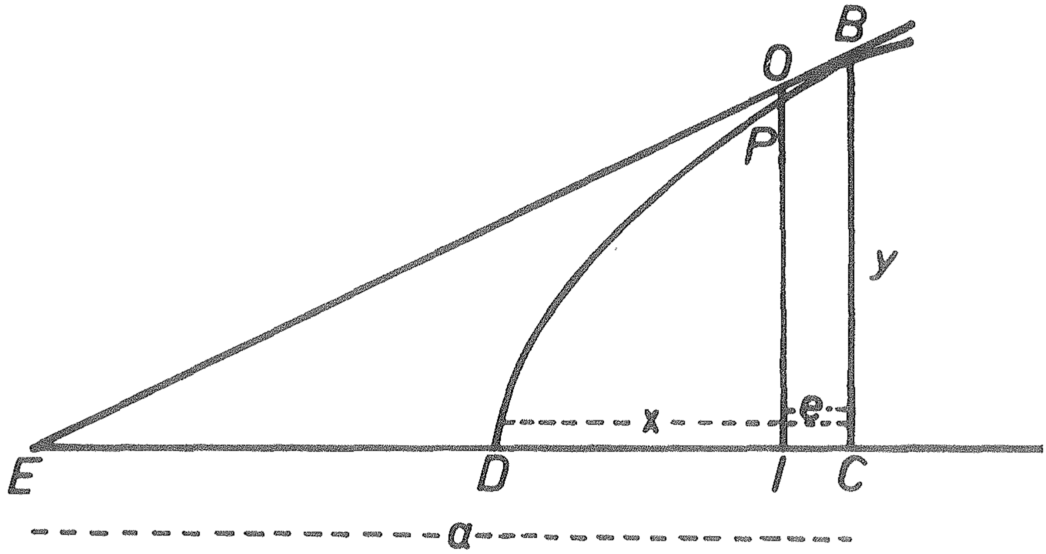


Figure 1.8.1.

it follows that

$$DC : DI > CB^2 : IO^2. \tag{1.8.2}$$

Since the triangles EIO and ECB are similar, we have

$$CB^2 : IO^2 = EC^2 : EI^2. \tag{1.8.3}$$

Thus

$$DC : DI > EC^2 : EI^2. \tag{1.8.4}$$

Let $DC = x$ (x is known since the point B is given), $EC = a$ (the unknown quantity) and $IC = e$. Then (1.8.4) becomes

$$x : (x - e) > a^2 : (a - e)^2, \tag{1.8.5}$$

or

$$xa^2 + xe^2 - 2xae > xa^2 - a^2e. \tag{1.8.6}$$

Fermat replaces this inequality by the adequality

$$xa^2 + xe^2 - 2xae \approx xa^2 - a^2e. \tag{1.8.7}$$

By using the procedure of the method of maxima and minima he obtains $a = 2x$, and thereby determines the tangent.

In a letter to Mersenne of January 1638 Descartes objected to this determination, maintaining that it did not solve the problem of an extreme value (see Fermat *Works*, vol. 2, 126–132, or Descartes *Works*, vol. 1, 486–493). He also accused Fermat of not having used the specific property of the curve, so that the determination would give the same result for all curves. The last objection is clearly wrong, and may be ascribed to the hostile attitude which Descartes took to Fermat after

Fermat had criticised his *Dioptrique* (*1637a*). The first objection, however, is worth examining.

The inequality $IO > IP$ holds for curves concave with respect to the axis, and the inequality $IO < IP$ for convex curves. For curves without points of inflection it is possible from these inequalities to find a magnitude depending on $a-e$ and $x-e$ which has an extreme value for $x-e=x$ (see Itard *1947a*, 597, and Mahoney *1973a*, 167). As $x(=DC)$ is known, a may be determined from the requirement for an extreme value. Neither in *Methodus* nor in Fermat's later writings, however, is there any indication that this was the way he related his method of tangents to his method of maxima and minima. In the memoir *1638b* of June 1638, Fermat, after having explained his method, wanted to show that there was a relation between the method of maxima and minima and that of tangents. However, by solving a problem of extrema he did not find the tangent to the curve, but rather the normal. This gave an algorithm quite different from the one used in *Methodus* and explained in the memoir. He is therefore not likely to have used this relation when he established his method of tangents. (By the way, the problem of extreme values which Fermat solved was suggested by Descartes in his first attack on Fermat's method.) So Descartes was right after all in raising the objection that the method of tangents was not a direct application of the method of maxima and minima.

When, in the memoir just mentioned, Fermat explained his method of tangents to Descartes, he clearly showed that he used only the procedure drawn from the method of maxima and minima. Descartes thereafter accepted the method. In modern notation Fermat's explanation can be reproduced in the following way. Let B be the point (x, y) on the curve $f(x, y)=0$ and let $DI=x-e$ (see figure 1.8.1). From the similar triangles EOI and EBC we obtain

$$IO=\frac{y(a-e)}{a}. \qquad (1.8.8)$$

Since IO is almost equal to PI, Fermat writes

$$f\left(x-e, \frac{y(a-e)}{a}\right)\approx 0. \qquad (1.8.9)$$

This is the *adequality* to which he applied his procedure from the method of maxima and minima. It is not difficult to see that it will lead to an expression for a corresponding to

$$a=-\frac{yf_y'}{f_x'}. \qquad (1.8.10)$$

If we have the parabola $\alpha x = y^2$, we obtain from (1.8.9)

$$\alpha(x-e) - \frac{y^2(a-e)^2}{a^2} \approx 0, \qquad (1.8.11)$$

or

$$y^2(a-e)^2 \approx a^2\alpha(x-e)\,; \qquad (1.8.12)$$

and since $y^2 = \alpha x$, then

$$x(a-e)^2 \approx a^2(x-e), \qquad (1.8.13)$$

which is (1.8.7).

As the method requires a development of

$$f\left(x-e, \frac{y(a-e)}{a}\right),$$

it was in its original presentation only applicable to algebraic curves (because in Fermat's time only algebraic functions were developed). However, in ' Doctrinam tangentium ' Fermat extended its field of application to include some transcendental curves. He introduced two principles (Fermat *Works*, vol. 1, 162), stating that it was allowed

> (1) . . . to replace the ordinates to the curves by the ordinates to the tangents [already] found . . .
>
> (2) . . . to replace the arc lengths of the curves by the corresponding portions of tangents already found . . . .

These two principles enabled him to determine the tangent to the cycloid (*ibid.*, 163). Let HCG be a cycloid with vertex C and generating circle CMF (figure 1.8.2), and RB be the tangent at an arbitrary point R. For the sake of convenience we reproduce his analysis with use of some modern symbols. Let $CD = x$, $RD = f(x)$, $MD = g(x)$, and the magnitude to be investigated $DB = a$. The specific property of the cycloid is the following :

$$f(x) = RM + MD = \text{arc } CM + g(x). \qquad (1.8.14)$$

Let $DE = e$, and draw NE parallel to RD intersecting RB at N and the circle at O ; as usual in the method of tangents, we have that

$$NE = \frac{f(x)(a-e)}{a} \approx f(x-e), \qquad (1.8.15)$$

where

$$f(x-e) = \text{arc } CO + g(x-e) = \text{arc } CM - \text{arc } OM + g(x-e). \qquad (1.8.16)$$

Let MA be the tangent to the circle at M intersecting NE at V, and let $MA = d$ and $AD = b$.

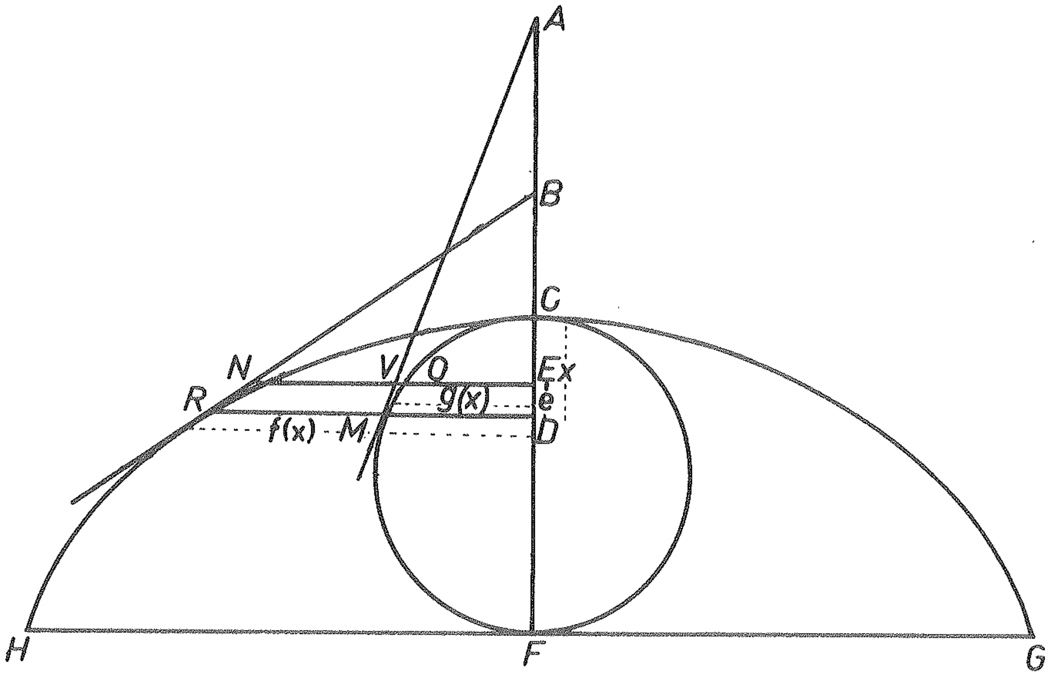


Figure 1.8.2.

From the first principle Fermat obtains

$$g(x-e) \approx EV = \frac{g(x)(b-e)}{b}, \tag{1.8.17}$$

and from the second

$$\text{arc } OM \approx MV = \frac{de}{b}. \tag{1.8.18}$$

Thus

$$f(x-e) \approx \text{arc } CM - \frac{de}{b} + \frac{g(x)(b-e)}{b}, \tag{1.8.19}$$

which together with (1.8.14) and (1.8.15) gives

$$\frac{(\text{arc } CM + g(x))(a-e)}{a} \approx \text{arc } CM - \frac{de}{b} + \frac{g(x)(b-e)}{b}. \tag{1.8.20}$$

Hence, by the standard procedure,

$$\frac{\text{arc } CM + g(x)}{a} = \frac{d+g(x)}{b}, \tag{1.8.21}$$

or

$$\frac{f(x)}{a} = \frac{d+g(x)}{b}. \tag{1.8.22}$$

Geometrically it is seen that

$$\frac{d+g(x)}{b}=\frac{g(x)}{x}, \tag{1.8.23}$$

so that the tangent at R is parallel to MC.

1.9. *The method of exhaustion*

The method of geometrical integration which was considered in the first part of the 17th century to be ideal was the exhaustion method, which had been invented by Eudoxus and improved by Archimedes. The name is unfortunate because the idea of the method is to avoid the infinite, and the method therefore does not lead to an exhaustion of the figure to be determined, as will be seen from the following outline of the idea behind it (see Dijksterhuis *1956a*, 130–132).

The method aims at showing that an area, a surface or a volume to be investigated, X, is equal to a known magnitude of the same kind K (for example, X may be the surface of a sphere and K four great circles on the sphere). A monotone ascending sequence I_n and a monotone descending sequence C_n of, respectively, inscribed and circumscribed figures to X are constructed. Thus we have the result :

$$\text{for all } n,\ I_n < X < C_n. \tag{1.9.1}$$

It is then shown either that for any magnitude $\epsilon > 0$ there exists a number N such that

$$C_N - I_N < \epsilon\ ; \tag{1.9.2}$$

or that for any two magnitudes of the same kind μ and ν where $\mu > \nu > 0$, there exists a number N such that

$$C_N : I_N < \mu : \nu, \tag{1.9.3}$$

and further that

$$\text{for all } n, \quad I_n < K < C_n. \tag{1.9.4}$$

From (1.9.1), (1.9.2) or (1.9.3), and (1.9.4), it follows by a *reductio ad absurdum* that $K = X$.

This last demonstration always proceeds in the same manner, independent as it is of the magnitudes in question. Nevertheless, whenever applying the method, the Greek mathematicians wrote out the argument down to the last detail. The reason may be that they did not have a notation which made it easy for them to deal with the general case. Furthermore, it is rather complicated to establish the basic inequalities of the proof, especially (1.9.4), and the method can be used only if K

is known in advance. This means that it needs to be supplemented by another method, if results are to be produced.

Among the mathematicians of the early 17th century there was a desire to find such a method of obtaining results which, in contrast to the method of exhaustion, would be direct. It would be as well if the new method, apart from giving results, could be used to prove the relations achieved. Such a direct method might have been obtained had it been realised that

$$\lim_{n\to\infty} C_n = \lim_{n\to\infty} I_n, \tag{1.9.5}$$

and had X been put equal to that limit ; however, this was not within the style of expression and power of abstraction of 17th-century mathematicians.

The path which they followed was that of an intuitive understanding of the geometric magnitudes. They imagined an area to be filled up, for example, by an infinite number of parallel lines. When, in 1906, Heiberg found Archimedes's *The method*, it was discovered that Archimedes too had adopted this point of view in his search for results. However, he did not regard it as sufficiently rigorous to be applied in proofs. Kepler, too, had used techniques involving such intuitive considerations, and it was the purpose of the first systematic exposition of the method of indivisibles to legitimise such techniques. This exposition, *Geometria indivisibilibus continuorum nova quadam ratione promota* (' Geometry by indivisibles of the continua advanced by a new method ' : *1635a*, hereafter referred to as *Geometria*), by Cavalieri, appeared in 1635, when he was a professor of mathematics at the University of Bologna. The ideas that it contained were developed in 1627, as can be seen in a letter from Cavalieri to Galileo (Galileo *Works*, vol. 13, 381).

The mathematicians differed on the importance to attach to a proof by the method of indivisibles. Most of those who thought about the matter regarded the method of indivisibles as heuristic, and thought that an exhaustion proof was still necessary. The exhaustion method was therefore modified and extended during the 17th century (see Whiteside *1961a*, 333–348). In many cases, however, mathematicians confined themselves to the remark that the results achieved by the method of indivisibles could be easily demonstrated by an exhaustion proof.

1.10. *Cavalieri's method of indivisibles*

Geometria, and Cavalieri's later work *Exercitationes geometricae sex* (' Six geometrical exercises ' : *1647a*), became well-known among mathematicians. The works inspired many of them to find their own methods, whereas others like Fermat and Roberval found their integration methods independently of Cavalieri.

Cavalieri presented two methods of indivisibles in his *Geometria*, and called them the 'collective' and the 'distributive' methods respectively. The first six of the seven books of *Geometria* embody the collective method, and a summary of it is given in *Exercitationes*, Book I. The framework of this section cannot possibly allow for a full account of the wide spectrum of concepts and ideas which Cavalieri introduced and developed in these six books, but the following outline gives a rough idea of his approach.

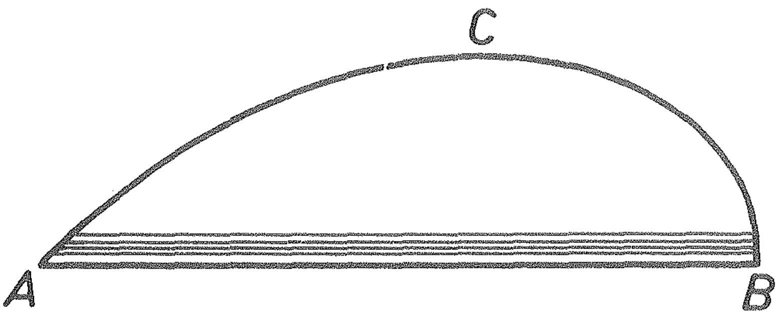


Figure 1.10.1.

Let there be given a plane figure $F = ABC$ limited by the curve ABC, and the straight line AB, called the 'regula' (figure 1.10.1). Cavalieri imagined that a straight line starting along AB is uniformly displaced parallel to AB, and considered the bunch of parallel line-segments which made up the section between F and the line during the motion. He named these line-segments 'all the lines of the given figure' ('omnes lineae propositae figurae'), and sometimes referred to them as 'the indivisibles of the given figure'; let us denote them by $\mathcal{O}_F(l)$.

Expressed in modern terms, Cavalieri constructed a mapping

$$F \to \mathcal{O}_F(l) \qquad (1.10.1)$$

from the set of plane figures into a set consisting of bunches of parallel line-segments. He then extended Eudoxus's theory of magnitudes (see book V of Euclid's *Elements*) to include his new magnitudes $\{\mathcal{O}_F(l)\}$. Thereafter he established—although not in a mathematically satisfactory manner—the fundamental relation

$$F_1 : F_2 = \mathcal{O}_{F_1}(l) : \mathcal{O}_{F_2}(l) \qquad (1.10.2)$$

between two plane figures (Cavalieri *1635a*, Book II, Theorem 3). By letting the regula be a plane he obtained in a similar way the relation

$$S_1 : S_2 = \mathcal{O}_{S_1}(p) : \mathcal{O}_{S_2}(p), \tag{1.10.3}$$

where S_i is a solid and $\mathcal{O}_{S_i}(p)$ all the planes belonging to it, $i = 1, 2$.

Cavalieri's aim was to find the ratio on the left hand side of (1.10.2) by calculating the ratio on the right hand side. In doing so he was greatly helped by a postulate which leads to 'Cavalieri's theorem' (described below), a skilful use of previous results, theorems about similar figures, and the concept of powers of line-segments.

The postulate (*1635a*, Corollarium to Theorem 4 of Book II) states that if in two figures F_1 and F_2 with the same altitude every pair of corresponding line-segments (that is, line-segments at equal distances from the common regula) has the same ratio, then $\mathcal{O}_{F_1}(l)$ and $\mathcal{O}_{F_2}(l)$ have this ratio too. In modern notation and using figure 1.10.2,

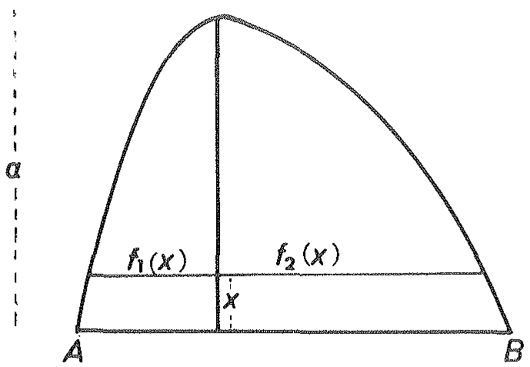


Figure 1.10.2.

if $f_1(x) : f_2(x) = b : c$ for all x $0 < x < a$,

$$\text{then } \mathcal{O}_{F_1}(l) : \mathcal{O}_{F_2}(l) = b : c. \tag{1.10.4}$$

This, together with (1.10.2), immediately gives 'Cavalieri's theorem':

If $f_1(x) : f_2(x) = b : c$ for all x $0 < x < a$,

$$\text{then } F_1 : F_2 = b : c \tag{1.10.5}$$

(*1635a*, Book II, Theorem 4).

Cavalieri's skilful employment of his previous results may be illustrated by a simple example. It is easily realised from figure 1.10.3 that

$$\mathcal{O}_{ACF}(l) = \mathcal{O}_{CDF}(l) \quad \text{and} \quad \mathcal{O}_{ACDF}(l) = \mathcal{O}_{ACF}(l) + \mathcal{O}_{CDF}(l). \tag{1.10.6}$$

From these relations follows the theorem that the parallelogram $ACDF$ is the double of each of the triangles ACF and CDF. However, Cavalieri was capable of interpreting them in a more general way.

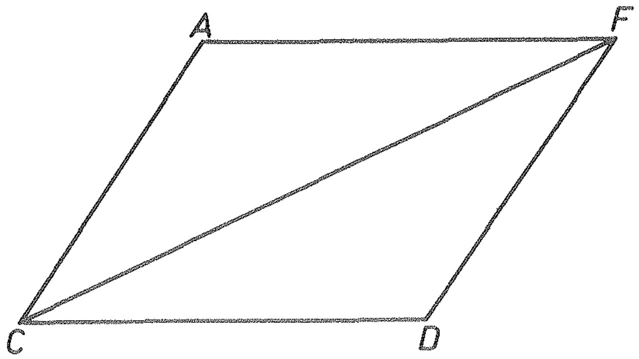


Figure 1.10.3.

By setting $AC = CD$ and using concepts which we cannot go into here he obtained a result which he could use every time he needed a proportion corresponding to

$$\int_0^a x\,dx : \int_0^a a\,dx = 1 : 2 \qquad (1.10.7)$$

(*1635a*, Corollarium II to Theorem 19 of Book II: compare figure 1.10.4).

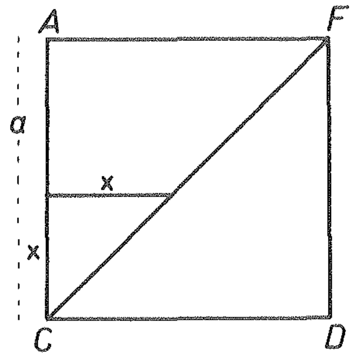


Figure 1.10.4.

Cavalieri found an alternative to integrating x^2 by introducing the squares of line-segments. If, instead of considering the line-segments of $\mathcal{O}_F(l)$, we take their squares situated in parallel planes, we obtain what he called ' all the quadrates of the given figure ' (' omnia quadrata propositae figurae ') ; this aggregate will be denoted by $\mathcal{O}_F(\Box l)$.

Let us illustrate the use of this concept by an example. For each l in the parallelogram $ACGE$ in figure 1.10.5 we have

$$\Box R_l T_l + \Box T_l V_l = 2\Box R_l S_l + 2\Box T_l S_l, \qquad (1.10.8)$$

where $\Box R_l T_l$ means the square on the side $R_l T_l$. From this relation Cavalieri concluded that

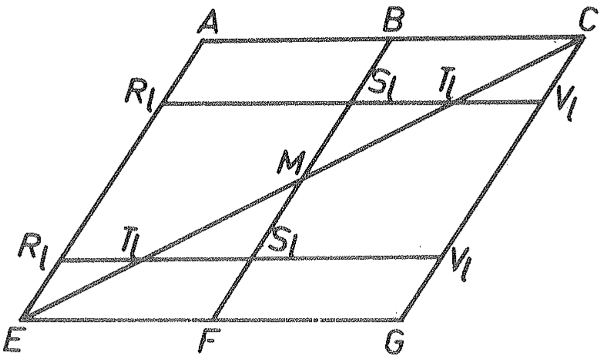


Figure 1.10.5.

$$\mathcal{O}_{AEC}(\square l)+\mathcal{O}_{CEG}(\square l)= 2\mathcal{O}_{ABFE}(\square l)+2(\mathcal{O}_{MEF}(\square l)+\mathcal{O}_{CBM}(\square l)). \qquad (1.10.9)$$

Since the triangles AEC and CEG are congruent, we have

$$\mathcal{O}_{AEC}(\square l)=\mathcal{O}_{CEG}(\square l), \qquad (1.10.10)$$

and similarly

$$\mathcal{O}_{MEF}(\square l)=\mathcal{O}_{CBM}(\square l). \qquad (1.10.11)$$

He further proved that, since the triangles CEG and MEF are similar, the following relation holds :

$$\mathcal{O}_{CEG}(\square l) : \mathcal{O}_{MEF}(\square l)=EG^3 : EF^3=8 : 1. \qquad (1.10.12)$$

In the same way he found that

$$\mathcal{O}_{ACGE}(\square l) : \mathcal{O}_{ABFE}(\square l)=EG^2 : EF^2=4 : 1. \qquad (1.10.13)$$

From (1.10.9)–(1.10.13) it follows that

$$\mathcal{O}_{ACGE}(\square l)=3\mathcal{O}_{CEG}(\square l) \qquad (1.10.14)$$

(*1635a*, Book II, Theorem 24). This result has as an immediate consequence that a cylinder is three times the inscribed cone. Cavalieri applied the relation (1.10.14) to a series of problems concerning conics, interpreting it by analogy with (1.10.6) as a relation which was an alternative to

$$\int_0^a x^2\,dx=\tfrac{1}{3}a^3. \qquad (1.10.15)$$

The first six books of *Geometria* are in their general style a copy of the Greek classical mathematical works, built up of definitions and postulates from which the theorems are carefully deduced, all verbally. Although Cavalieri ingeniously used his concepts to obtain many results, this made the reading of the book rather tedious. Perhaps he felt this himself ; at least, he wrote to Galileo in 1634 that he composed the seventh book of *Geometria* to help those who found the concept of 'all the lines' too difficult (Galileo *Works*, vol. 16, 113). In this last book, concerning the distributive method, he turned to a more intuitive treatment of the indivisibles.

As we saw in the relation (1.10.2), by the collective method Cavalieri found the ratio between two figures by comparing the aggregates of indivisibles. In the distributive method, two figures with the same altitude were compared by comparing *corresponding* indivisibles. The basic relation in this theory was Cavalieri's theorem (1.10.5), for which he gave a new proof without using the concepts from the collective theory.

A part of the criticism to which Cavalieri's methods were exposed was levelled against the nature of his indivisibles and the problem of the structure of the continuum. Some mathematicians took him to mean that a plane figure was made up of indivisibles and that these were line-segments. This was against the Aristotelian view of a continuum as divisible into parts of the same kind as the original magnitude, the parts again being infinitely divisible. To avoid his seeming error of dimensionality, they tried tentatively to conceive a plane figure as composed of rectangles with infinitesimal breadth. But the distinction was of theoretical interest only, for it remained usual to consider the ratio between two areas, so that an eventually missing Δx was cancelled by the relation

$$\frac{A}{B}=\frac{\sum a_n\Delta x}{\sum b_n\Delta x}=\frac{\sum a_n}{\sum b_n}, \tag{1.10.16}$$

where a_n and b_n are the altitudes in the rectangles of which the areas A and B are composed.

The conception of an area as a kind of a sum $\sum a_n\Delta x$ did not solve the problem, because it was still uncertain what was meant by an infinitesimal magnitude and by an infinite sum. Despite the lack of rigour in their foundations, the methods were useful insofar as they provided the mathematicians with new results.

1.11. *Wallis's method of arithmetic integration*

To determine the area under the spiral of Galileo, Fermat used an arithmetic quadrature which he described in a letter to Mersenne in 1638 (Mersenne *Correspondence*, vol. 7, 377–380). In his *Traité des indivisibles*[1] Roberval squared many figures on the basis of arithmetical

[1] *Traité* uses a method of infinitesimals which Roberval worked out about 1630. The date of the composition of the *Traité* is, however, unknown. It was first printed in 1693 (Roberval *Works*$_1$) and reprinted in 1730 (*Works*$_2$).

considerations. Pascal observed in his treatise *Potestatum numericarum summa* (' sum of numerical powers ') that his results concerning the sums $\sum_{i=0}^{m}(A+id)^n$ (where A, d and n are natural numbers) could be applied to the quadratures of curves (Pascal *Works*$_1$, vol. 3, 364 ; *Works*$_2$, vol. 2, 1272). Using proofs by complete induction he also established the rules for determining the binomial coefficients (*1654a*).

But most of the results based on a method of arithmetic integration were achieved by John Wallis. His treatise on the subject, *Arithmetica infinitorum* (' The arithmetic of infinites ' : *1655a*), is not burdened with proofs, for he relied boldly and confidently on his really astounding intuition as to the correlation between the sums of different series. He called his favourite method in the treatise ' modus inductionis ' : later it was termed ' incomplete induction '. One might also call it ' conclusion by analogy '.

Wallis started the treatise by establishing by this method that

$$\frac{\sum_{i=0}^{l} i}{(l+1)l}=\tfrac{1}{2},\quad \frac{\sum_{i=0}^{l} i^2}{(l+1)l^2}=\tfrac{1}{3}+\frac{1}{6l},\quad \frac{\sum_{i=0}^{l} i^3}{(l+1)l^3}=\tfrac{1}{4}+\frac{1}{4l},\qquad (1.11.1)$$

and similarly that

$$\frac{\sum_{i=0}^{l} i^n}{(l+1)l^n}=\frac{1}{n+1}+\frac{a_1}{l}+\dots\frac{a_{n-1}}{l^{n-1}},\qquad (1.11.2)$$

where the a_i's are rational numbers and $n=4$, 5, 6 (Wallis *1655a*. Propositions I, XIX and XXXIX ; *Works*, vol. 1, 365, 373 and 382), From this he concluded that

$$\lim_{l\to\infty}\left\{\frac{\sum_{i=0}^{l} i^n}{(l+1)l^n}\right\}=\frac{1}{n+1}\qquad (1.11.3)$$

(*ibid.*, 384). This relation enabled him to square the curves $y=x^n$ in figure 1.11.1 to obtain

$$\frac{\sum_{x=0}^{a} y}{\sum_{x=0}^{a} b}=\lim_{l\to\infty}\left\{\frac{\left(\frac{a\,.\,0}{l}\right)^n+\left(\frac{a\,.\,1}{l}\right)^n+\dots\left(\frac{a\,.\,l}{l}\right)^n}{a^n+a^n+\dots a^n}\right\}$$

$$=\lim_{l\to\infty}\left\{\frac{\sum_{i=0}^{l} i^n}{(l+1)l^n}\right\}=\frac{1}{n+1},\qquad (1.11.4)$$

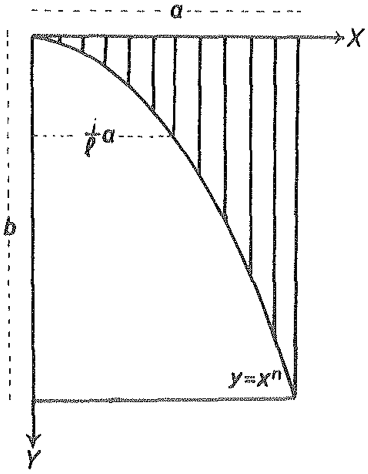


Figure 1.11.1.

a result which corresponds to

$$\frac{\int_0^a x^n\,dx}{a^{n+1}} = \frac{1}{n+1}. \qquad (1.11.5)$$

This result was not new, and indeed it had been found by many of Wallis's predecessors; but he did not stop there. He extended the range of n in (1.11.3) to include at least all rational numbers except -1. The foundation of his extension is an observation which he made in connection with the formula (1.11.3), namely: If the numbers l^{n_1}, $l^{n_2}, \ldots, l^{n_r}$ are in geometric progression (where $n_1, n_2, \ldots, n_r$ are non-negative whole numbers), then

$$\lim_{l\to\infty}\left\{\frac{\sum_{i=0}^{l+1} l^{n_j}}{\sum_{i=0}^{l} i^{n_j}}\right\}, \quad j=1, 2, \ldots r, \qquad (1.11.6)$$

will be in arithmetic progression (*ibid.*, 387). Further, from the fact that for $0 \leqslant p \leqslant q$ $(q=1, 2, 3, \ldots)$

$l^0, l^{1/q}, l^{2/q}, \ldots l^{p/q}, \ldots l^1$ are in geometric progression,

$1, 1+\frac{1}{q}, 1+\frac{2}{q}, \ldots 1+\frac{p}{q}, \ldots 2$ are in arithmetic progression,

and the first and last members of the latter sequence are the reciprocals of the values of the right hand side of (1.11.3) for $n=0$ and $n=1$ respectively, he concluded that

$$\lim_{l\to\infty}\left\{\frac{\sum_{i=0}^{l+1} i^{p/q}}{\sum_{i=0}^{l} l^{p/q}}\right\}=\frac{1}{1+\dfrac{p}{q}} \quad (1.11.7)$$

(*ibid.*, 390). He did not doubt that the relation (1.11.7) held good for all $p/q \geqq 0$; he even said that it was valid for an irrational exponent, such as $\sqrt{3}$ (*ibid.*, 395), and as he extended the concept of power to include negative powers he considered (1.11.7) to be valid for them too—except -1 (*ibid.*, 408). By means of (1.11.7), he was now able to determine, when p/q was a rational number different from -1, the ratios between the areas under the curves $y = x^{p/q}$ and the circumscribed rectangles. He could also determine the ratios between the volumes obtained by a revolution of these areas about an axis and the circumscribed cylinders.

After that, Wallis proceeded to study polynomials in x; he applied the formula (1.11.7) to binomial expansions of $(x^p(D^n - x^n))^m$ when p, n and m are small natural numbers and D is a constant, and by analogy deduced that

$$\frac{\int_0^D [x^p(D^n-x^n)]^m\,dx}{D^{m(n+p)+1}} = \frac{n\,.\,2n\,.\,\ldots\,mn}{(mp+1)(mp+n+1)(mp+2n+1)\ldots(mp+mn+1)} \quad (1.11.8)$$

(*ibid.*, 419–420 and 425–430), a result which he put into various tables. (For clarity I render the last of his sums as integrals.) He further extended (1.11.8) to include the case where p and n are positive rational numbers (*ibid.*, 433).

One of Wallis's purposes was to square the circle; he stressed that from (1.11.8) and its extension we know for $m = 0, 1, 2, 3 \ldots$ the 'sums'

$$\frac{\int_0^R (R^2-x^2)^m\,dx}{R^{2m+1}} \quad\text{and}\quad \frac{\int_0^D (xD-x^2)^m\,dx}{D^{2m+1}}, \quad (1.11.9)$$

and for $m = 1, 2, 3, \ldots$ the 'sum'

$$\frac{\int_0^R (R^{1/m}-x^{1/m})^m\,dx}{R^2}, \quad (1.11.10)$$

where R is the radius and D the diameter of the circle. He wished to find the values of these ' sums ' for $m=\frac{1}{2}$, and he introduced the symbol ' □ ' to signify the reciprocal of (1.11.10) :

$$\square = \frac{R^2}{\int_0^R (R^2-x^2)^{1/2}\,dx}\left(=\frac{4}{\pi}\right). \tag{1.11.11}$$

By a principle of interpolation which we cannot go into here, he succeeded in establishing the formula

$$\frac{R^{n-1}}{\int_0^R (R^2-x^2)^{(n/2)-1}\,dx}=a_n \quad \text{for} \quad n=1, 2, 3, \ldots, \tag{1.11.12}$$

where

$$\left.\begin{aligned} &a_1=\frac{\square}{2},\quad a_2=1,\quad a_3=\square,\\ &a_{n+2}=\frac{3\,.\,5\,.\,7\ldots(n+1)}{2\,.\,4\,.\,6\ldots\quad n},\quad n=2, 4, 6, \ldots\\ &a_{n+2}=\frac{4\,.\,6\,.\,8\ldots(n+1)}{3\,.\,5\,.\,7\ldots\quad n}\,\square,\quad n=3, 5, 7, \ldots \end{aligned}\right\} \tag{1.11.13}$$

(see Prag *1929a*, 389–392, and Whiteside *1961a*, 237–241). From the fact that

$$\frac{a_{n+2}}{a_n}=\frac{n+1}{n} \quad \text{for} \quad n=1, 2, 3 \ldots, \tag{1.11.14}$$

he concluded that the ratio a_{n+1}/a_n is continuously decreasing,[1] so that

$$\frac{n+2}{n+1}=\frac{a_{n+3}}{a_{n+1}}=\frac{a_{n+3}}{a_{n+2}}\cdot\frac{a_{n+2}}{a_{n+1}}<\left(\frac{a_{n+2}}{a_{n+1}}\right)^2<\frac{a_{n+2}}{a_{n+1}}\cdot\frac{a_{n+1}}{a_n}=\frac{a_{n+2}}{a_n}=\frac{n+1}{n}, \tag{1.11.15}$$

and hence

$$\sqrt{\left(\frac{n+2}{n+1}\right)}<\frac{a_{n+2}}{a_{n+1}}<\sqrt{\left(\frac{n+1}{n}\right)}. \tag{1.11.16}$$

From the formulae (1.11.13) he obtained for odd n the inequalities

[1] Wallis was lucky that his sequence behaved in this way, for a sequence defined by

$$a_1=k,\ a_2=1,\ a_{2n+1}=\frac{2n}{2n-1}\,a_{2n-1},\quad \text{and} \quad a_{2n+2}=\frac{2n+1}{2n}\,a_{2n}$$

will not generally have a_{n+1}/a_n continuously decreasing.

$$\frac{3\,.\,3\,.\,5\,.\,5\,.\,7\,.\,7\,.\,.\,.\,(n-2)\,.\,n\,.\,n}{2\,.\,4\,.\,4\,.\,6\,.\,6\,.\,8\,.\,.\,.\,(n-1)(n-1)(n+1)}\sqrt{\left(\frac{n+2}{n+1}\right)} < \square$$
$$< \frac{3\,.\,3\,.\,5\,.\,5\,.\,7\,.\,7\,.\,.\,.\,(n-2)\,.\,n\,.\,n}{2\,.\,4\,.\,4\,.\,6\,.\,6\,.\,8\,.\,.\,.\,(n-1)(n-1)(n+1)}\sqrt{\left(\frac{n+1}{n}\right)}. \qquad (1.11.17)$$

In the limit as $n \to \infty$, these give a result now called ' Wallis's product ' :

$$\frac{4}{\pi} = \square = \frac{3\,.\,3\,.\,5\,.\,5\,.\,7\,.\,7\,.\,9\,.\,9\,.\,.\,.}{2\,.\,4\,.\,4\,.\,6\,.\,6\,.\,8\,.\,8\,.\,10\,.\,.\,.} \qquad (1.11.18)$$

(Wallis *Works*, vol. 1, 469).

1.12. *Other methods of integration*

Most of the methods of integration in use before the time of Newton and Leibniz made use of an equidistant sub-division of intervals and compared the area or volume to be found with a known one, as we have seen with Cavalieri and Wallis. However, Fermat had a method which allowed him to make an absolute calculation of an area, employing a sub-division which meant that the areas of the infinitesimal rectangles to be summed were in a geometric progression with quotient less than unity. We may illustrate this by means of an example from his treatise on quadrature *De aequationum*, which he wrote about 1658 using ideas he had already had in the 1640s (see Mahoney *1973a*, 243 f.).

Fermat considered the hyperbolas

$$yx^n = k, \quad k \text{ is a constant}, \quad n = 2, 3, 4, \ldots\,. \qquad (1.12.1)$$

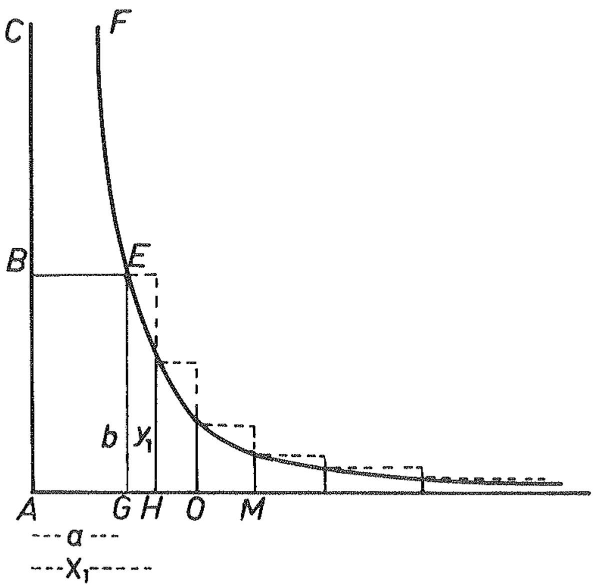


Figure 1.12.1.

For convenience I reproduce his arguments in modern terms. He divided the x-axis to the right of the point G (see figure 1.12.1) in intervals GH, HO, OM, ... of lengths $x_1 - a$, $x_2 - x_1$, $x_3 - x_2$, ... ($a = AG$), so that

$$\frac{a}{x_1} = \frac{x_1}{x_2} = \frac{x_2}{x_3} = \dots \qquad (1.12.2)$$

and hence

$$\frac{a}{x_1} = \frac{x_1 - a}{x_2 - x_1} = \frac{x_2 - x_1}{x_3 - x_2} = \dots . \qquad (1.12.3)$$

He then considered the circumscribed rectangles

$$\left.\begin{array}{l} R_1 = b(x_1 - a), \text{ where } b = GE, \\ R_r = y_{r-1}(x_r - x_{r-1}), \quad r = 2, 3, \dots . \end{array}\right\} \qquad (1.12.4)$$

From (1.12.1)–(1.12.4) it follows that

$$\frac{R_1}{R_2} = \frac{b(x_1 - a)}{y_1(x_2 - x_1)} = \frac{{x_1}^{n-1}}{a^{n-1}} = \frac{x_{n-1}}{a}, \qquad (1.12.5)$$

$$\frac{R_r}{R_{r+1}} = \frac{y_{r-1}(x_r - x_{r-1})}{y_r(x_{r+1} - x_r)} = \frac{{x_r}^n a}{{x_{r-1}}^n \, x_1} = \frac{{x_1}^n a}{a^n x_1} = \frac{x_{n-1}}{a}, \qquad (1.12.6)$$

which means that the circumscribed rectangles are in a geometric progression with quotient a/x_{n-1}.

To determine the sum S of a geometric progression with first term α and quotient u/v ($u < v$), Fermat used the following relation :

$$\frac{v - u}{u} = \frac{\alpha}{S - \alpha} \qquad (1.12.7)$$

(this is equivalent to $S = \alpha/(1 - u/v)$). Hence, if S denotes the sum of the rectangles R_r, we have :

$$\frac{x_{n-1} - a}{a} = \frac{b(x_1 - a)}{S - b(x_1 - a)} \qquad (1.12.8)$$

or

$$\frac{x_{n-1} - a}{x_1 - a} = \frac{ba}{S - b(x_1 - a)}. \qquad (1.12.9)$$

He then imagined the intervals $x_1 - a$, $x_2 - x_1$, ... to be sufficiently small and almost equal, and he concluded that the left hand side of (1.12.9) by adequality is equal to $n - 1$. Further, as the intervals are small, he concluded that $S - b(x_1 - a)$ in the relation (1.12.9) can be set equal to the area σ defined by the hyperbola and the lines GH and GE. Hence

$$n-1=\frac{ba}{\sigma}=\frac{AG\,.\,GE}{\sigma}, \tag{1.12.10}$$

and the quadrature is achieved.[1] We could have obtained (1.12.10) by taking the limit of both sides of (1.12.9) for x_1 approaching a, but he did not use limits. He observed that his method could not be applied when $n=1$ as the rectangles R_r will then be equal.

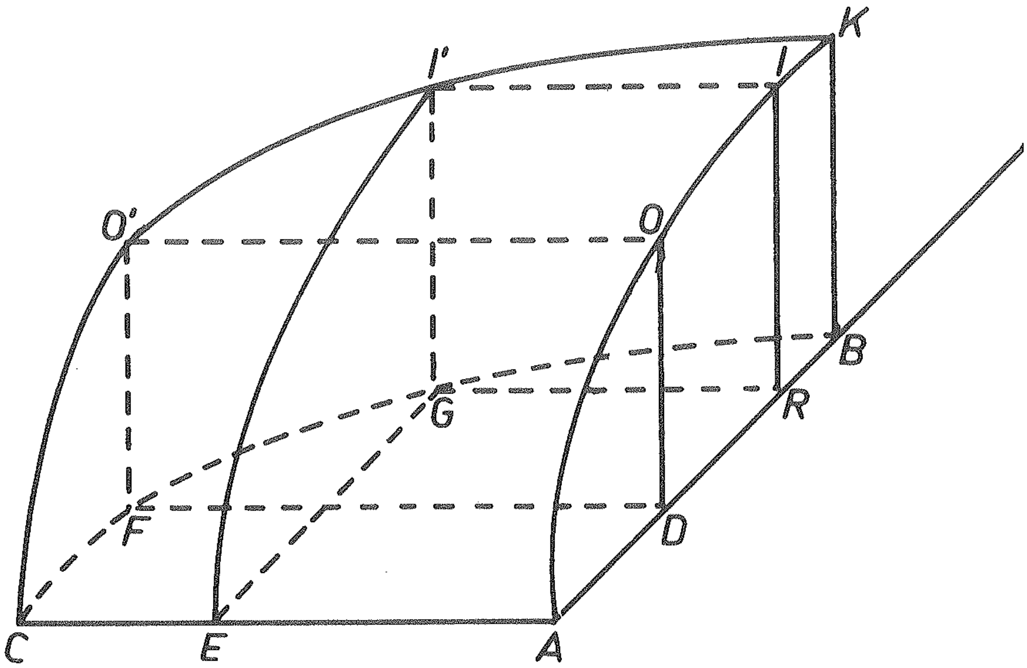


Figure 1.12.2.

[1] Fermat called his method 'logarithmic' (*Works*, Vol. 1, 265). In his time the word 'logarithmic' was used to characterise a connection between a geometric and an arithmetic progression ; hence 'logarithmic' was also used at that time where today we would say 'exponential'. Let us indicate in modern terms how his expression and proof can be interpreted. If we let $a=\exp(t_0)$ and $x_r=\exp(t_0+r\Delta t)$, $r=1, 2, 3, \ldots$, then we have a sub-division which is equivalent to (1.12.2). An easy calculation shows that

$$R_r=k\exp[-(n-1)(t_0+(r-1)\Delta t)](\exp[\Delta t]-1). \tag{1}$$

Hence

$$S=\sum_{r=1}^{\infty}R_r=k\exp[-(n-1)t_0](\exp[\Delta t]-1):(1-\exp[-(n-1)\Delta t]), \tag{2}$$

and

$$\lim_{\Delta t\to 0}S=(k\exp[-(n-1)t_0]):(n-1)=(a\,.\,b):(n-1). \tag{3}$$

An ingenious use of geometrical considerations and arguments from statics led mathematicians to many transformations corresponding to transformations of integrals, which could be applied to find connections between various problems solved by quadratures and cubatures. In his *1658c* Pascal systematically drew up schedules in which appear the sums necessary to determine areas and volumes as well as their centres of gravity. He found a fundamental theorem for these connections by conceiving the volume $KCAB$ (see figure 1.12.2) both as composed of the rectangles $FDOO' = FD \,.\, DO$ and as composed of the areas $EGI' = ARI$ (Pascal *1658c*, 'Lemme génére1' in the section 'Traité des trilignes rectangles'; *Works*$_1$, vol. 9, 3–5). That is,

$$\sum_{AB} FD \,.\, DO = \sum_{AC} EGI'. \tag{1.12.11}$$

If we put $AB = a$, $AC = b$, $AD = x$, $FD = y = f(x)$ and $DO = z = g(x)$ (both being monotone functions), the relation corresponds to

$$\int_0^a f(x)g(x)\, dx = \int_0^b \left(\int_0^{f^{-1}(y)} g(t)\, dt \right) dy, \tag{1.12.12}$$

which can be obtained by an integration by parts. Since $f(a) = 0$ we have :

$$\int_0^a f(x)g(x)\, dx = -\int_0^a \left(\int_0^x g(t)\, dt \right) f'(x)\, dx$$

$$= \int_0^b \left(\int_0^{f^{-1}(y)} g(t)\, dt \right) dy. \tag{1.12.13}$$

When $g(x) = x$, we obtain

$$\int_0^a xy\, dx = \int_0^b \frac{x^2}{2}\, dy. \tag{1.12.14}$$

Roberval found the summation form of (1.12.14) in his *Traité* in a way similar to that of Pascal (Roberval *Works*$_2$*a*, 271), and it was used by Fermat too (*Works*, vol. 1, 272). Among other things, it could be applied to the determination of the centre of gravity of the area $\int_0^a y\, dx$. Let the x-coordinate of this point be ξ; in modern notation the argument is the following (see figure 1.12.3). If we consider a lever AC and let the area $\int_0^a y\, dx$ operate on the arm ξ on the one side, and at the other let all the rectangles $y\Delta x$ of the area $\int_0^a y\, dx$ or BDC operate each on the arm x, then there will be equilibrium. Hence we have

$$\xi \int_0^a y\, dx = \int_0^a xy\, dx. \tag{1.12.15}$$

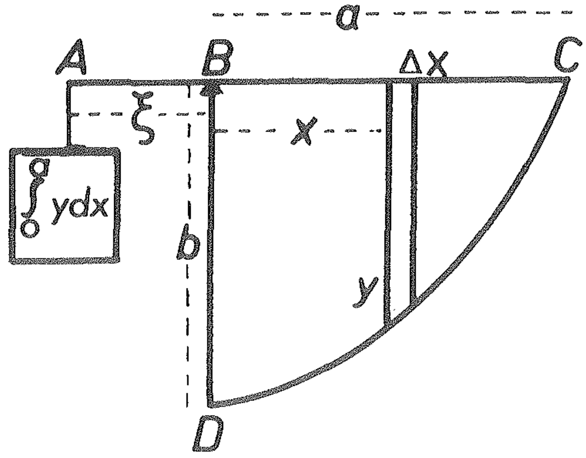


Figure 1.12.3.

Therefore, by (1.12.14),

$$\xi = \frac{\int_0^b \frac{x^2}{2}\, dy}{\int_0^a y\, dx}, \qquad (1.12.16)$$

which gives the x-coordinate of the centre of gravity. The y-coordinate can be found in a similar way.

(1.12.16) is equivalent to the relation

$$\pi \int_0^b x^2\, dy = 2\pi\xi \int_0^a y\, dx, \qquad (1.12.17)$$

which states that the volume obtained by revolving the area *BCD* about the axis *BD* (compare figure 1.12.3) is equal to the product of the area and the distance traversed by the centre of gravity. This is a special case of the theorem now known as ' Pappus-Guldin's theorem ', formulated by Paul Guldin in *Centrobaryca* (*1635–1641a*, vol. 2, 147) in the following way : ' the product of a rotating quantity and the path of rotation [that is, the circumference of the circle traversed by the centre of gravity], is equal to the quantity generated by the rotation '. The theorem is also found in Book VII of Pappus's *Collections*, but it may be a later addition (see, for example, Ver Eecke *1932a*).

1.13. *Concluding remarks*

The examples given in sections 1.5–1.8 and 1.10–1.12 illustrate the remark in the introductory section 1.1 about the special character of the infinitesimal methods in the period 1630–1660. In the case of the methods of quadrature we saw that they were all naturally founded on the conception of an area as an infinitesimal sum. However, mathematicians differed in their ways of approaching the problems raised by that concept. And not only were the methods of the various mathematicians based on different ideas; some of them also developed different methods, each one adapted to solve special problems of quadrature.

Some of the methods of solving tangent or normal problems led to fixed rules—of which the most general one was Hudde's rule for determining the sub-tangent to an algebraic curve—while others only suggested a procedure. The ideas behind the methods differed widely. Descartes used an argument about the number of points of intersection between a circle and the curve; Fermat employed similar triangles and the concept of *adequality*; while Roberval's method was founded on an intuitive conception of instantaneous velocity and the law of parallelogram of velocities. The characteristic triangle (with sides Δx, Δy and Δs) did not explicitly play a part in the deduction of the tangent methods. Nevertheless, it was applied by (for example) Pascal in connection with a transformation of a sum (see section 2.3); but not until Leibniz was the importance of this triangle fully recognised.

Thus the period did not in itself bring any perception of basic concepts which were applicable to the determination of tangents as well as to quadratures. An important reason why mathematicians failed to see the general perspectives inherent in their various methods was probably the fact that to a great extent they expressed themselves in ordinary language without any special notation and so found it difficult to formulate the connections between the problem they dealt with. As an illustration we may consider one of the results achieved by the different quadrature methods outlined in the preceding sections. This result can be expressed in modern terms as

$$\int_0^a x^n\,dx = \frac{a^{n+1}}{n+1}, \tag{1.13.1}$$

where n is a natural number different from -1. The mathematicians of that period, however, could not express their result so simply; they had to refer to areas under special parabolas. Their terminology did not prevent them from seeing connections such as that between the rectification of the parabola and the quadrature of the hyperbola, or the relation of certain inverse tangent problems to quadratures; but it may have barred their way to a deeper insight into the meaning of these connections.

These remarks are not to be taken in the negative sense at all. It is not the task of a historian of mathematics to evaluate the work of earlier mathematicians by present mathematical standards, nor to emphasise the inadequacy of their concepts as compared to modern ones. On the contrary, a historian of mathematics ought to enter into the mode of thought of the period under consideration in order to bring out the development of the mathematical ideas in its historical context. Briefly, it may be said that the mathematicians in the period preceding the invention of the calculus blazed the trail for its invention. They did so by employing heuristic methods, by making the geometry analytical, and by seeking methods for solving problems of quadratures and tangents.[1]

[1] I am grateful to Dr. John North of Oxford University for correcting some of my linguistic mistakes, and to Dr. D. T. Whiteside of Cambridge University for his valuable comments on the manuscript.

Text 10: T. H. Kjeldsen (2011). "Does history have a significant role to play for the learning of mathematics? Multiple perspective approach to history, and the learning of meta level rules of mathematical discourse". In: *History and Epistemology in Mathematics Education. Proceedings of the Sixth European Summer University ESU 6.* Ed. by E. Barbin, M. Kronfellner, and C. Tzanakis. Vienna: Verlag Holzhausen GmbH, pp. 51–62.

DOES HISTORY HAVE A SIGNIFICANT ROLE TO PLAY FOR THE LEARNING OF MATHEMATICS?

Multiple perspective approach to history, and the learning of meta level rules of mathematical discourse

Tinne Hoff KJELDSEN
IMFUFA, NSM, Roskilde University, PO-Box 260, Roskilde, Denmark
thk@ruc.dk

ABSTRACT

In the present paper it will be argued that and proposed how the history of mathematics can play a significant role in mathematics education for the learning of meta rules of mathematical discourse. The theoretical argument is based on Sfard's theory of thinking as communicating. A multiple perspective approach to history of mathematics from the practice of mathematics will be introduced along with the notions of epistemic objects and techniques. It will be argued that by having students read and analyse mathematical texts from the past within this methodology, the texts can function as "interlocutors". In such learning situations the sources can assist in revealing meta rules of (past) mathematical discourses, making them explicit objects for students' reflections. The proposed methodology and the potential of history for the learning of meta-discursive rules of mathematical discourse is exemplified by analyses of four sources from the 17th century by Fermat and Newton belonging to the calculus, and it is demonstrated how meta level rules can be made objects of students' reflections. The paper ends with a proposal for a matrix-organised design for how the introduced approach to history of mathematics for elucidating meta-discursive rules might be implemented in upper secondary mathematics education.

1 Introduction

One can think of several purposes for using history in mathematics education: (1) For pedagogical reasons; it is often argued that history motivates students to learn mathematics by bringing in a human aspect. (2) As a didactical method for the learning and teaching of the subject matter of mathematics. (3) For the development of students' historical awareness and knowledge about the development of mathematics and its driving forces. (4) For general educational goals, with respect to which the so called cultural argument makes the strongest case for history, but history can also serve general educational goals in mathematics education of developing interdisciplinary competences as a counterpart to specialisation (Beckmann 2009). These purposes are not necessarily mutually independent. In carefully designed teaching sessions all four of the above mentioned purposes can be realized in varying degrees.[1]

Regarding the question whether history promotes students' learning of mathematics I have argued in (Kjeldsen 2011), that by adopting a multiple perspective approach to history from the practice of mathematics, history has potentials in developing students' mathematical competence while providing them with genuine historical insights. In the present paper, I will go a step further and suggest that history might have a much more

[1] See (Kjeldsen 2010) where it is shown how all these four purposes can be accomplished in problem oriented and student directed project work. In (Jankvist and Kjeldsen 2011) two avenues for integrating history in mathematics education are discussed with respect to the development of students' mathematical competence and historical awareness anchored in the subject matter of mathematics, respectively, both within a scholarly approach to history. In (Kjeldsen forthcoming) a didactical transposition of history from the academic research subject to history in mathematics education is proposed for developing a framework for integrating history of mathematics in mathematics education.

profound role to play for the learning of mathematics. This suggestion is based on Sfard's (2008) theory of commognition.

In the following it will be argued that, and proposed how, the history of mathematics can play a significant role in the teaching and learning of mathematics. The theoretical argument is outlined in section 2. In section 3, the multiple perspective approach to history of mathematics from its practice is presented along with some tools of historians'. The adaptation for mathematics education is discussed in section 4. The potential of history for the learning of meta-discursive rules of mathematical discourse is exemplified in section 5 through analyses of four sources from the 17th century by Fermat and Newton belonging to the calculus. In section 6 a proposal is outlined for a so called matrix-organised design for how such an approach to history of mathematics for elucidating meta-discursive rules might be implemented in upper secondary school. The paper ends with a concluding section 7.

2 The theoretical argument for the significance of history

In Sfard's (2008, 129) theory of *Thinking as Communicating* mathematics is seen as a discourse that is regulated by discursive rules, and where the objects of mathematics are discursive constructs. There are two kinds of discursive rules both of which are important for the learning of mathematics: object-level rules and meta-discursive rules.

The object-level rules have the content of the discourse as object. In mathematics they regard the properties of mathematical objects. The meta-discursive rules have the discourse itself as object. They govern proper communicative actions shaping the discourse. The meta-discursive rules are often tacit. They are implicitly present in discursive actions when we e.g. judge if a solution or proof of a mathematical problem or statement can count as a proper solution or proof (Sfard 2000, 167). The meta-discursive rules are not necessary; they are given historically.

The meta-discursive rules are connected to the object-level of the discourse and have an impact on how participants in the discourse interpret its content. As a consequence, developing proper meta-discursive rules are indispensable for the learning of mathematics (Sfard 2008, 202). This means that designing learning situations where meta-discursive rules are elucidated is an important aspect of mathematics education. History of mathematics is an obvious method for illuminating meta-discursive rules. Because of the contingency of these rules, they can be treated at the *object level* of history discourse, and thereby be made into explicit objects of reflection. Hence, history might have a significant role to play for the learning of mathematics, precisely because meta-discursive rules can be treated as objects of historical investigations. By reading historical sources students can be acquainted with episodes of past mathematics where other meta-discursive rules governed the discourse. If students study original sources in their historical context, and try to understand the work of past mathematicians, their views on mathematics, the way they formulated and argued for mathematical statements etc. the historical texts can play the role as "interlocutors", as discussants acting according to meta rules that are different than the ones that govern the discourse of our days mathematics and (maybe) of the students. By identifying meta rules that governed past mathematics and comparing them with the rules that govern e.g. their textbook, students can be engaged in learning processes where they can become aware of their own meta rules. In case a student is acting according to non-proper meta rules he or she might experience what Sfard calls a commognitive conflict, which is "a situation in which different discursants are acting according to different metarules" (Sfard 2008, 256). Such situations can initiate a metalevel change in the learner's discourse.

This, of course, presupposes a genuine approach to history. In section 3 and 4 it will be argued that within a multiple perspective approach to the history of the practice of mathematics, and by using historian of mathematics' tools such as the idea of epistemic objects and techniques, original sources can be used in mathematics education to have students investigate and reflect upon meta-discursive rules. For further discussion of this see (Kjeldsen and Blomhøj 2011), where also some student directed problem oriented project work performed by students at degree level mathematics are analysed with respect to students' reflections about meta-discursive rules to provide empirical evidence for the theoretical claim. These projects will not be presented here. Instead I will present a proposal (see section 6) for a so called matrix-organised design for how such an approach to history of mathematics for investigating meta-discursive rules might be implemented in upper secondary school.

3 A multiple perspective approach to history

The so called whig interpretation of history has been debated at length in the historiography of mathematics.[2] In mathematics education Schubring (2008) has pointed out how translations of sources, due to an underlying whig interpretation of history, have changed the mathematics of the source. In the whig interpretation history is written from the point of view of the present, as explained by the British historian Herbert Butterfield, who coined the term in the 1930s:

> It is part and parcel of the whig interpretation of history that it studies the past with reference to the present … The whig historian stand on the summit of the twentieth century and organises his scheme of history from the point of view of his own day. (Butterfield 1931, 13)

If we want to use history to throw light on changes in meta rules from episodes of past mathematics to our days mathematics whig interpretations of history poses a problem, because, as it has been pointed out by Wilson and Ashplant (1988, 11) history then becomes "constrained by the perceptual and conceptual categories of the present, bound within the framework of the present, deploying a perceptual 'set' derived from the present". In this quote, Wilson and Ashplant emphasis exactly why one cannot design learning and teaching situations that focus on bringing out differences in meta rules of past episodes in the history of mathematics and modern ones within a whig interpretation of history. Historical sources cannot function as "interlocutors" that can be used to clarify differences in meta rules if the sources is interpreted within the framework of how mathematics is conceptualized and perceived of today.

The trap of whiggism can be avoided by investigating past mathematics as a historical product from its practice. This implies to study the sources in their proper historical context with respect to the intellectual workshop[3] of their authors, the particular mathematicians, to ask questions such as: how was mathematics viewed at the time? How did the mathematician, who wrote the source, view mathematics? What was his/hers

[2] Discussions of whig interpretations in the historiography of mathematics can be followed e.g. in the following papers (Unguru 1975), (van der Waerden 1976), (Freudenthal 1977), (Unguru and Rowe, 1981/82), (Grattan-Guiness 2004).
[3] See (Epple 2004).

intention? Why and how did mathematicians introduce certain concepts? How did they use them and for what purposes? Why and how did they work on the problems they did? Which kinds of tools were available for the mathematician (group of mathematicians)? Why and how did they employ certain strategies of proofs? Such questions can reveal underlying meta rules of the discourse at the time and place of the sources. By posing and answering such questions to the sources, possibilities for identifying meta rules that governed the mathematics of the source can emerge, and hereby also opportunities for turning meta rules into explicit objects of reflection in a teaching and learning situation.

As explained by Kjeldsen (2009b, 2011) one way of answering such questions and to provide explanations for historical processes of change is to adopt a multiple perspective approach to the history of the practice of mathematics. I have taken the term "a multiple perspective" approach from the Danish historian Jensen (2003). It signifies that episodes of the past can be studied from several perspectives, several points of observation, depending on which kind of insights into, or from, the past, we are searching for. Episodes in the history of mathematics can e.g. be studied from the perspective of sub-disciplines within mathematics to understand if, and if so, how other fields in mathematics have influenced the emergence and/or the development of the episode under consideration. They can be studied from an applied point of view to understand e.g. dynamics between pure and applied mathematics, or the role of mathematical modelling in the production of mathematical and/or scientific knowledge. They can be studied from a sociological perspective to understand the institutionalization of mathematics, its funding etc. They can be studied from a gender perspective, from a philosophical perspective and so on.

4 Adaptation for mathematics education

In mathematics education the above approach can be implemented on a small scale, by focusing on a limited amount of perspectives that address the intended learning. In the present context the purpose is to use past mathematics and history of mathematics as a means for elucidating meta discursive rules and make them into explicit objects of students' reflections. Hence, students should study the sources to answer clearly formulated historical questions that concern the underlying meta rules of the mathematics in the source.

Theoretical constructs that have been developed by historians of mathematics and/or science to investigate the history of scientific practices can be used to "open" the sources. With respect to the purpose of the present paper of uses of history to reveal meta rules of a (past) mathematical discourse by studying the history of mathematics from its practice, the notions of epistemic objects and techniques are promising tools. The term epistemic object refers to mathematical objects that are treated in a source, i.e. the object about which mathematicians were searching for new knowledge or were trying to grasp. The term epistemic technique refers to the methods employed in the source by the mathematicians to investigate the epistemic objects.[4] These theoretical constructs can give insights into the dynamics of concrete productions of pieces of mathematical knowledge, since they are constructed to distinguish between elements of the source that provide answers and elements that generate mathematical questions.[5]

[4] These notions have been adapted into the historiography of mathematics by Epple (2004) from Rheinberger's (1997) study of experimental science.
[5] For examples of uses of this methodological tool see (Epple 2004) and (Kjeldsen 2009a).

Text 10: T. H. Kjeldsen (2011). "Does history have a significant role to play for the learning of mathematics? Multiple perspective approach to history, and the learning of meta level rules of mathematical discourse". In: *History and Epistemology in Mathematics Education. Proceedings of the Sixth European Summer University ESU 6*. Ed. by E. Barbin, M. Kronfellner, and C. Tzanakis. Vienna: Verlag Holzhausen GmbH, pp. 51–62.

The question is whether history dealt with in this way, where students study episodes from the history of mathematics from perspectives that pertain to meta rules of (past) discourses, ask historians' questions to the sources concerning the practice of mathematics, and answer them using theoretical constructs such as epistemic objects and techniques, can facilitate meta level learning in mathematics education. In the following section four texts from the 1600s will be analyzed to provide some answers to this question.

5 Analysis of four sources within the proposed methodology

Four texts from the 1600s will be used in the following; two by Pierre de Fermat (Fermat I and Fermat II) and two by Isaac Newton (Newton I and Newton II). Fermat I is Fermat's text on maxima and minima taken from Struik's (1969) *A Source Book in Mathematics, 1200-1800*, whereas Fermat II is called "A second method for finding maxima and minima", which is published in Fauvel's and Gray's (1988) reader in the history of mathematics. Newton I is Newton's demonstration of how he found a relation between the fluxions of some fluent quantities from a given relation between these. This text is the one prepared by Baron and Bos (1974), whereas Newton II is Newton's method of tangent taken from Whiteside's (1967) *The Mathematical Works of Isaac Newton*. The quality of these translations of sources can be criticised, and investigated for degrees of whiggism (Schubring 2008), but this will not be done in the present paper. In a teaching situation the students should work with the four texts, but in order to give the reader an impression of the texts, summaries of the four texts are inserted here:

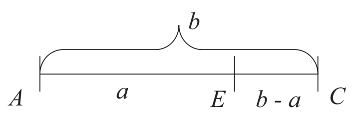
In Fermat I, Fermat stated a rule for the evaluation of maxima and minima and gave an example. The text is summarised below in Box 1.

Fermat I: On a method for the evaluation of max. and min.

Rule: let a be any unknown of the problem

- Indicate the max or min in terms of a
- Replace the unknown a by $a+e$ – express max./min. in terms of a and e
- "adequate" the two expressions for max./min. and remove common terms
- Both sides will contain terms with e – divide all terms by (powers of) e
- Suppress all terms in which e will still appear – and equate the others
- The solution of this equation will yield the value of a leading to max./min.

Example: To divide the segment AC at E so that AE x EC may be a maximum



Max: $a(b\text{-}a) = ab\text{-}aa$

$(a+e)b\text{-}(a+e)(a+e) = ab+eb\text{-}aa\text{-}2ae\text{-}ee$

$ab+eb\text{-}aa\text{-}2ae\text{-}ee \sim ab\text{-}aa$ *"adequate"*

$eb \sim 2ae + ee$ remove common terms

$b \sim 2a + e$; $b=2a$; $a=½b$; divide, suppress, solve

Box 1

If the above procedure is translated into modern mathematics using functions and the derivative it can be explained why Fermat reached the correct solution. But this does not explain how Fermat was thinking, since he knew neither our concept of a function nor our concept of derivatives. In Fermat II we can get a glimpse of how Fermat was thinking.

Text 10: T. H. Kjeldsen (2011). "Does history have a significant role to play for the learning of mathematics? Multiple perspective approach to history, and the learning of meta level rules of mathematical discourse". In: *History and Epistemology in Mathematics Education. Proceedings of the Sixth European Summer University ESU 6*. Ed. by E. Barbin, M. Kronfellner, and C. Tzanakis. Vienna: Verlag Holzhausen GmbH, pp. 51–62.

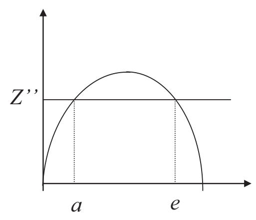
The text is summarised below in Box 2.

Fermat II: A second method for finding maxima and minima

- Here he explained why his "rule" leads to max./min.: correlative equations – Viete
- Resolving all the difficulties concerning limiting conditions

Example: To divide the line *b* such that the product of the segments shall be a max.

> If one proposes to divide the line *b* in such a way that the product of the segments [*a* and (*b-a*)] shall equal *z''* … there will be two points answering the question, and they will be found situated on one side and the other of the point corresponding to the max.



ba-aa = z'' and *be-ee = z''*

ba-aa = be-ee ; ba-be = aa-ee

Divide by *a-e*

b = a + e

At the point of maximum we will have *a = e*, then

b = a +a = 2*a*, hence as before *a*=½*b*.

If we call the roots *a* and *a*+*e* (instead of *a* and *e*) the procedure follows the rule from text I.

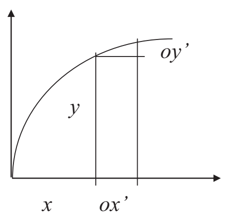
Box 2

In Newton I, Newton explained through an example, how, given a relation between fluent quantities, a relation between the fluxions of these quantities can be found. In Box 3 his procedure is summarised and illustrated with an example of a second degree equation instead of the third degree equation that Newton used in the text.

Newton I: Find relation between fluxions from fluents

Newton's fluxions and fluents

- Curves are trajectories (paths) for motions
- Variables are entities that change with time – fluents *x* , *y*
- The speed with which fluents change – fluxions *x'* , *y'* (*Newton: dots!*)
- Newton: All problems relating to curves can be reduced to two problems:
    1. Find the relation between the fluxions given the relation between the fluents.
    2. The opposite.



Example: *axx+bx+c-y*=0 substitute *x*, *y* with *x+x'o*, *y+y'o*

a(*x+x'o*)(*x+x'o*)+*b*(*x+x'o*)+*c-y-y'o*=0

axx+*a*2*xx'o*+*ax'x'oo*+*bx*+*bx'o*+*c-y*-*y'o*=0

*a*2*xx'o*+*ax'x'oo*+*bx'o-y'o*=0

*a*2*xx'*+~~*ax'x'o*~~+*bx'-y'*=0 divided by *o*; cast out terms with *o*

*a*2*xx'*+*bx'-y'*=0 hence *y'/x'*=2*ax*+*b*

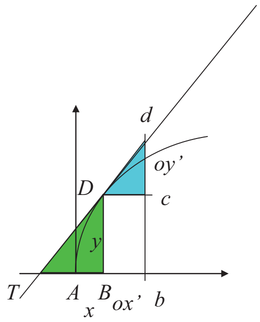
Box 3

In Newtons's terminology *o* denotes an infinitely small period of time, so *ox'* [Newton used a dot over *x* instead of *x'* to designate the fluxions] is the infinitely small addition by which *x* increases during the infinitely small interval of time.

Finally, in Newton II, Newton showed how to draw tangents to curves and illustrated it with the same example as he used in the first text. In Box 4 below the example is illustrated with reference to the example used in Box 3.

Newton II: To draw Tangents to Curves

Example:



Similar triangles: *dcD* and *DBT*

TB:BD = Dc:cd “infinitesimal triangle”

BT/y = x'o/y'o =x'/y'

x'/y' can be found by the method from Newton I

Box 4

The suggestion made in this paper is that these four sources can be used to exhibit changes in meta rules of mathematical discourse, if students read the sources from the perspective of rigor, and focus on entities and arguments. The following worksheet (Box 5) can be used to guide the students work. It consists of two sets of questions. The first set concerns questions that help the students to identify the epistemic objects and techniques of the two texts. The students are asked to compare and contrast the answers they get from studying Fermat, Newton, and their textbook, respectively.

Perspective

Rigor – entities, arguments

Worksheet: History from the practice of math. Compare/contrast Fermat and Newton

Questions:

- What mathematical objects are Fermat/Newton dealing with? Compare/contrast
- How do they perceive them? – compare with your textbook
- What are the problems they are trying to solve?
- What techniques are they using? – what do we do today?
- How do they argue for their claims? – how do we argue today?

Epistemic objects and techniques

- Can you find any changes in understandings of the involved mathematical concepts from Fermat over Newton to today? Explain
- Can you find any changes in the way of argumentation from Fermat over Newton to today? Explain
- What kind of objections do you think your math teacher would have to Fermat’s and Newton’s texts?

Meta-rules – explicit object of reflection
Opportunities provided by history

Box 5

The second set of questions refers directly to meta rules of the involved mathematical discourses.

Text 10: T. H. Kjeldsen (2011). "Does history have a significant role to play for the learning of mathematics? Multiple perspective approach to history, and the learning of meta level rules of mathematical discourse". In: *History and Epistemology in Mathematics Education. Proceedings of the Sixth European Summer University ESU 6*. Ed. by E. Barbin, M. Kronfellner, and C. Tzanakis. Vienna: Verlag Holzhausen GmbH, pp. 51–62.

Regarding the first set of questions, an analysis of the four texts and the comparison between the objects that Fermat and Newton investigated, how they perceived them, the problems they tried to solve, the techniques they used and the arguments they employed might be summarised in the following scheme (Box 6):

Fermat:	Newton:
Objects: curves - algebraic expressions; ex.: multiplication of line segments	**Objects:** any curve; variables that change in time
Perceive: Area; geometrical problems treated by algebraic methods	**Perceive:** trajectories for moving particles
Problem: evaluate max/min	**Problem:** relations between fluxions (velocities) given relations between the fluents
Techniques: equations, roots, algebraic mani.	**Techniques:** algebraic mani; physics, geometry
Argue: Text 1: shows the method works on an example; Text 2: heuristic arguments with roots in equations given by an example	**Argue:** Physical arguments about distance and velocity, algebraic arguments, infinitesimal triangle, *o*-infinitely small

Box 6

Regarding the second set of questions, which refers to meta rules of the discourse, the following changes can be discussed (se Box 7):

Changes in understanding:

Fermat: curves; algebraic expressions

Newton: curves, traced by a moving point, variables change in time

Today: functions, correspondence between variables in domains

Changes in the way of argumentation:

Fermat: ad hoc; "it works – its true"; heuristic argument, no infinitely small quantities

Newton: more general procedure, physical arguments, infinitesimal triangle, infinitely small quantities (*o*)

Today: limit, the real numbers, epsilon-delta proofs

Box 7

In Kjeldsen and Blomhøj (2011) we have analysed some student directed problem oriented project work conducted by students in a degree level university mathematics programme. Here we were able to demonstrate that history, used within the framework of a multiple perspective approach to the history of mathematics from its practice, can be used in mathematics education to give students insights into how meta rules of a mathematical discourse are established and why/how they change. These projects were made in a rather unique educational setting and the question is whether this methodology can be implemented in more traditional educational settings. The analyses of the sources guided by the worksheet (Box 5) and presented in Box 6 and Box 7 suggest that this approach can elucidate meta rules and turn them into explicit objects for students reflections. In the following section I present an outline for a so called matrix-organised design for how such a multiple perspective approach to history of mathematics from its practice might be implemented in upper secondary mathematics education.

6 Implementation in upper secondary school: A proposal

In the Danish upper secondary school system history of mathematics is part of the mathematics curriculum. The curriculum is comprised of a core curriculum which is mandatory and is tested in the national final, and a supplementary part, which should take up 1/3 of the teaching. History is mentioned explicitly in the supplementary part, which means that all upper secondary students should be taught some aspects of history of mathematics. The supplementary part of the curriculum is tested in an oral examination together with the core curriculum. In Box 8 below an outline is presented for a matrix organised design for how history could be (but has not yet been) implemented in a Danish upper secondary school for elucidating meta rules within the theoretical framework of section 2, 3 and 4, using the sources and the worksheet presented in section 5.

Implementation in a Danish high school: a proposal

Step 1: Six groups – basic groups (worksheets would have to be prepared for each group with respect to the intended learning)

1. The mathematical community in the 17th century
2. The standard history of analysis
3. Who were Fermat and Newton?
4. The two texts of Fermat - the questions of the worksheet of Box 5
5. The two texts of Newton - the questions of the worksheet of Box 5
6. Berkeley's critique of Newton

Step 2: Six groups – expert groups (each group consists of at least one member from each of the basic groups)
The experts teach the other group members of what they learned in their basic group. Each expert group write a common report/prepare an oral presentation of the collected work from all six basic groups as it was discussed in their expert groups

Step 3: A plenary discussion lead by the teacher focuses on methods of argumentation, the development/changes in the perception of objects and techniques, compared with the standards of today.

Box 8

This design follows a three step implementation. First six groups (so called basic groups) are formed who look into some aspects of the historical episode in question. In Box 8 it is suggested e.g. that group 1 investigates what the mathematical community of the 17th century looked like. Guided by a worksheet with questions relevant for the intended learning, the work in this group will provide the students with a sociological perspective on mathematics and its development. In step 2 new groups (so called expert groups) are formed. They consist of at least one member from each of the six basic groups. In this way each new group consists of individual experts. Each expert now teaches the other members of the new group what he/she learned in his/hers basic group, and based on their shared knowledge provided by the various experts they answer the second set of questions of the worksheet in Box 5. The design is referred to as being matrix organised because it can be illustrated with a matrix, where the members of basic group 1 is listed in column 1, the members of basic group in column 2, etc. In step 2 the expert groups are formed by taking the students in the rows, i.e. expert group 1 consists of the students listed in row 1; expert group 2 of the students listed in row 2, etc. In this way all expert groups consists of at least one member from each basic group. In such a set up it is possible to create complex teaching and learning situations where students work independently and autonomously in an inquire-like environment, developing general educational skills as well.[6]

7 Discussion and conclusion

The main question in the present paper is whether working with sources in the spirit of the worksheet of Box 5 within the methodology outlined in section 3 may give rise to situations where meta rules of (past) mathematical discourses are made into explicit objects of students' reflections, and whether this can assist the development of students' proper meta rules of mathematical discourse. As pointed out above, the analyses of the sources guided by the questions of the worksheet in Box 5, and the suggestions for answers outlined in Box 6 and 7, suggest that history and historical sources can be used within the methodological framework of section 2, 3 and 4 to elucidate meta rules and make them explicit objects for students reflections.

Regarding the second part of the question, whether such an approach to the use of history and historical sources in mathematics education also can assist the development of students' proper meta rules of our days mathematics is a complex question which is much more difficult to document. The framework and methodology outlined in this paper provide a theoretical argument for the claim that history has the potential for playing such a profound role for the learning of mathematics, but in order to realize this in practice more research needs to be done, and methodological tools for detecting students' meta rules and for monitoring any changes towards developing proper meta rules need to be developed.

REFERENCES

– Baron, M.E., Bos, H.J.M., 1974, "Newton and Leibniz", *History of Mathematics Origins and Development of the Calculus 3*, Milton Keynes: The Open University Press.

– Beckmann, A., 2009, "A Conceptual Framework for Cross-Curricular Teaching", *The Montana Mathematics Enthusiast* **6**, Supplement 1.

[6] Such a matrix organised design for using history in mathematics education to elucidate meta rules of past and present mathematics, using sources from the history of the development of the concept of a function, to have students reflect upon those, to develop students' mathematical competence, and general educational skills of independence and autonomy is being tried out in a pilot study in a Danish upper secondary class at the moment. Preliminary results from this study indicate that some of the students act according to meta discursive rules that coincide with Euler's; and that reading some of Dirichlet's text created obstacles for the students, that can be referenced to the differences in meta discursive rules. Results from the study will be published in forthcoming papers.

Text 10: T. H. Kjeldsen (2011). "Does history have a significant role to play for the learning of mathematics? Multiple perspective approach to history, and the learning of meta level rules of mathematical discourse". In: *History and Epistemology in Mathematics Education. Proceedings of the Sixth European Summer University ESU 6*. Ed. by E. Barbin, M. Kronfellner, and C. Tzanakis. Vienna: Verlag Holzhausen GmbH, pp. 51–62.

- Butterfield, H., 1931, *The whig interpretation of history*, London: G. Bell.
- Epple, M., 2004, "Knot Invariants in Vienna and Princeton during the 1920s: Epistemic Configurations of Mathematical Research", *Science in Context* **17**(1/2), 131-164.
- Fauvel, J., Gray, J., 1988, *The History of Mathematics: A Reader*, Milton Keynes: The Open University Press.
- Freudenthal, H., 1977, "What is Algebra and what has it been in History?" *Archive for History of Exact Sciences* **16**, 189-200.
- Grattan-Guiness, I., 2004, "The mathematics of the past: distinguishing its history from our heritage", *Historia Mathematica* **31**, 163-185.
- Jankvist, U.T., Kjeldsen, T.H., 2011, "New Avenues for History in Mathematics Education: Mathematical Competencies and Anchoring", *Science & Education*, **20**, 831-862.
- Jensen, B.E., 2003, *Historie – livsverden og fag*, Copenhagen: Gyldendal.
- Kjeldsen, T.H., 2009a, "Egg-forms and Measure-Bodies: Different Mathematical Practices in the Early History of the Modern Theory of Convexity", *Science in Context* **22**, 1-29.
- Kjeldsen, T. H., 2009b, "Abstraction and application: new contexts, new interpretations in twentieth-century mathematics", in The Oxford Handbook of the History of Mathematics, E. Robson & J. Stedall (eds.), New York: Oxford University Press, pp. 755–780.
- Kjeldsen, T. H., 2010, "History in mathematics education - why bother? Interdisciplinarity, mathematical competence and the learning of mathematics", in *Interdisciplinarity for the 21st Century: Proceedings of the 3rd International Symposium on Mathematics and its Connections to Arts and Sciences*, B. Sriraman & V. Freiman (Eds.), Charlotte, NC: Information Age Publishing, pp. 17-48.
- Kjeldsen, T. H., 2011, "History in a competency based mathematics education: a means for the learning of differential equations", in Recent developments on introducing a historical dimension in mathematics education, V. Katz & C. Tzanakis (eds.), Washington DC: The Mathematical Association of America, chapter 15.
- Kjeldsen, T. H., forthcoming, "Uses of History in Mathematics Education: Development of Learning Strategies and Historical Awareness", to be published in *CERME 7, Proceedings of the seventh Congress of the European Society for Research in Mathematics Education*
- Kjeldsen, T. H., Blomhøj, M., 2011, "Beyond Motivation – History as a method for the learning of meta-discursive rules in mathematics", *Educational Studies of Mathematics*, published on line first.
- Rheinberger, H-J., 1997, *Towards a History of Epistemic Things: Synthesizing Proteins in the Test Tube*, Standford: Standford University Press.
- Schubring, G., 2008, "The debate on a "geometric algebra" and methodological implications", Paper presented at the *HPM 2008 satellite meeting of ICME 11*, Mexico.
- Sfard, A., 2000, "On reform movement and the limits of mathematical discourse", *Mathematical Thinking and Learning* **2**(3), 157-189.
- Sfard, A., 2008, *Thinking as Communicating,* Cambridge: Cambridge University Press.
- Struik, D.J., 1969, *A Source Book in Mathematics, 1200-1800*, Cambridge, Massachusetts: Harvaard University Press.
- Unguru, S., 1975, "On the Need to Rewrite the History of Greek Mathematics", *Archive for History of Exact Sciences* **15**, 67-114.
- Unguru, S., Rowe, D., 1981/1982, "Does the Quadratic Equation Have Greek Roots? A Study of Geometric Algebra, Application of Areas, and Related Problems", *Libertas Mathematica* **1** 1981, **2** 1982.
- Van der Waerden, B.L., 1976, "Defence of a Shocking Point of View", *Archive for History of Exact Sciences* **15**, 199-210.
- Whiteside, D.T., 1967, *The Mathematical Works of Isaac Newton*, volume 1, New York: Cambridge University Press.
- Wilson, A., Ashplant, T.G., 1988, "Whig history and present-centred history" *Historical Journal* **31**, 1-16.

Text 11: Fermat on maxima and minima. From D. J. Struik (1969). *A Source Book in Mathematics. 1200–1800*. Cambridge (Mass.): Harvard University Press, pp. 222–225.

8 FERMAT. MAXIMA AND MINIMA

Modern textbooks on calculus take up first the differential and then the integral calculus. It may therefore come as a surprise to find that up to the middle of the seventeenth century the whole theory of infinitesimals concentrated on the computation of areas, volumes, and centers of gravity, that is, on what we now call the integral calculus. Tangent constructions were, until that period, based on the property that the tangent has only one point in common with the curve, as we can see in Euclid or Apollonius. Archimedes, in his book on spirals, found tangents by a method that seems to have been inspired by kinematic considerations. Even Torricelli, when determining the tangent at a point of the "hyperbola" $x^m y^n = k$, still used the ancient method (A. Agostini, "Il metodo delle tangenti fondato sopra la dottrina dei moti nelle opere di Torricelli," *Periodico di matematica* [*4*] *28* (1950), 141–158), and Descartes sought the normal prior to the tangent, and found it in some cases of algebraic curves by asking for double roots of a certain equation that expresses the abscissa of the intersections of the curve with a circle.

The beginning of the differential calculus, in which the tangent appears as the limit of a secant, can be studied in considerations concerning maxima and minima, as in Kepler's *Nova stereometria doliorum vinariorum* (Linz, 1615; see Selection IV.2). Here we read that "near a maximum the decrements on both sides are in the beginning only imperceptible" (decrementa habet insitio insensibilia; *Opere*, IV (1863), 612).

With Fermat we obtain an algorithm based on this fact. To understand his approach and its subsequent development into the method of the "characteristic triangle" (dx, dy, ds) we must take notice of the fact that Fermat and Descartes were among the first to apply the new algebra developed by Cardan, Bombelli, and Viète to the geometry of the ancients. This was, as we have seen, the beginning of the coordinate method. Descartes published his method in 1637, but Fermat's discovery was known only through his correspondence until 1679, the year of the publication of his works. Here is Fermat's approach, from his *Oeuvres*, III (1896), 121–123. It is followed by a paper in which he applied his method to the finding of a center of gravity (*Ibid.*, 124–126).

(1) ON A METHOD FOR THE EVALUATION OF MAXIMA AND MINIMA[1]

The whole theory of evaluation of maxima and minima presupposes two unknown quantities and the following rule:

Let a be any unknown of the problem (which is in one, two, or three dimensions, depending on the formulation of the problem). Let us indicate the maximum or minimum by a in terms which could be of any degree. We shall now replace the original unknown a by $a + e$ and we shall express thus the maximum or minimum quantity in terms of a and e involving any degree. We shall adequate [*adégaler*], to use Diophantus' term,[2] the two expressions of the maximum or minimum quantity and we shall take out their common terms. Now it turns out that both sides will contain terms in e or its powers. We shall divide all terms by e, or by a higher power of e, so that e will be completely removed from at least one of the terms. We suppress then all the terms in which e or one of its powers will still appear, and we shall equate the others; or, if one of the expressions vanishes, we shall equate, which is the same thing, the positive and negative terms. The solution of this last equation will yield the value of a, which will lead to the maximum or minimum, by using again the original expression.

Here is an example:

To divide the segment AC [Fig. 1] *at E so that AE × EC may be a maximum.*

Fig. 1 

We write $AC = b$; let a be one of the segments, so that the other will be $b - a$, and the product, the maximum of which is to be found, will be $ba - a^2$. Let now $a + e$ be the first segment of b; the second will be $b - a - e$, and the product of the segments, $ba - a^2 + be - 2ae - e^2$; this must be adequated with the preceding: $ba - a^2$. Suppressing common terms: $be \sim 2ae + e^2$. Suppressing e: $b = 2a$.[3] To solve the problem we must consequently take the half of b.

We can hardly expect a more general method.

ON THE TANGENTS OF CURVES

We use the preceding method in order to find the tangent at a given point of a curve.

Let us consider, for example, the parabola BDN [Fig. 2] with vertex D and of diameter DC; let B be a point on it at which the line BE is to be drawn tangent to the parabola and intersecting the diameter at E.

[1] This paper was sent by Fermat to Father Marin Mersenne, who forwarded it to Descartes. Descartes received it in January 1638. It became the subject of a polemic discussion between him and Fermat (*Oeuvres*, I, 133). On Mersenne, see Selection I.6, note 1.

[2] See Selection IV.7, note 5.

[3] Our notation is modern. For instance, where we have written (following the French translation in *Oeuvres*, III,122) $be \sim 2ae + e^2$, Fermat wrote: B in E adaequabitur A in E bis + Eq (Eq standing for E quadratum). The symbol $\sim$ is used for "adequates."

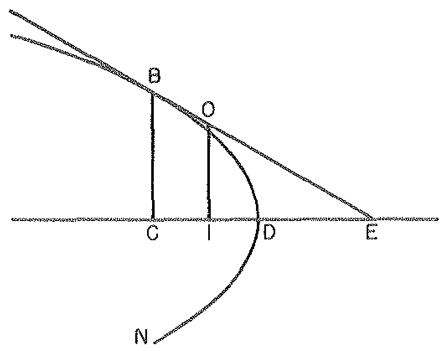


Fig. 2

We choose on the segment BE a point O at which we draw the ordinate OI; also we construct the ordinate BC of the point B. We have then: $CD/DI > BC^2/OI^2$, since the point O is exterior to the parabola. But $BC^2/OI^2 = CE^2/IE^2$, in view of the similarity of triangles. Hence $CD/DI > CE^2/IE^2$.

Now the point B is given, consequently the ordinate BC, consequently the point C, hence also CD. Let $CD = d$ be this given quantity. Put $CE = a$ and $CI = e$; we obtain

$$\frac{d}{d - e} > \frac{a^2}{a^2 + e^2 - 2ae}.$$ [4]

Removing the fractions:

$$da^2 + de^2 - 2dae > da^2 - a^2e.$$

Let us then adequate, following the preceding method; by taking out the common terms we find:

$$de^2 - 2dae \sim -a^2e,$$

or, which is the same,

$$de^2 + a^2e \sim 2dae.$$

Let us divide all terms by e:

$$de + a^2 \sim 2da.$$

On taking out de, there remains $a^2 = 2da$, consequently $a = 2d$.

Thus we have proved that CE is the double of CD—which is the result.

This method never fails and could be extended to a number of beautiful problems; with its aid, we have found the centers of gravity of figures bounded by straight lines or curves, as well as those of solids, and a number of other results which we may treat elsewhere if we have time to do so.

I have previously discussed at length with M. de Roberval[5] the quadrature of areas bounded by curves and straight lines as well as the ratio that the solids which they generate have to the cones of the same base and the same height.

---

[4] Fermat wrote: D ad $D - E$ habebit majorem proportionem quam $Aq.$ ad $Aq. + Eq. - A$ in E bis (D will have to $D - E$ a larger ratio than A^2 to $A^2 + E^2 - 2AE$).

[5] See the letters from Fermat to Roberval, written in 1636 (*Oeuvres*, III, 292–294, 296–297).

Now follows the second illustration of Fermat's "e-method," where Fermat's e = Newton's o = Leibniz' dx.[6]

[6] The gist of this method is that we change the variable x in $f(x)$ to $x + e$, e small. Since $f(x)$ is stationary near a maximum or minimum (Kepler's remark), $f(x + e) - f(x)$ goes to zero faster than e does. Hence, if we divide by e, we obtain an expression that yields the required values for x if we let e be zero. The legitimacy of this procedure remained, as we shall see, a subject of sharp controversy for many years. Now we see in it a first approach to the modern formula: $f'(x) = \lim_{e \to 0} \frac{f(x + e) - f(x)}{e}$, introduced by Cauchy (1820–21).

[7] This paper seems to have been sent in a letter to Mersenne written in April 1638, for transmission to Roberval. Mersenne reported its contents to Descartes. Fermat used the term "parabolic conoid" for what we call "paraboloid of revolution."

[8] "All parabolas" means "parabolas of higher order," $y = kx^n$, $n > 2$. The reference is to Archimedes' *On floating bodies*, II, Prop. 2 and following; see T. L. Heath, *The works of Archimedes* (Cambridge University Press, Cambridge, England, 1897; reprint, Dover, New York), 264ff.

Text 12: Fermat on maxima and minima. From J. Fauvel and J. Gray, eds. (1987). *The History of Mathematics: A Reader*. London: Macmillan Press Ltd., pp. 359–360.

11.C2 A second method for finding maxima and minima

In studying the method of syncriseos and anastrophe of Viète, and carefully following its application to the study of the nature of correlative equations, it occurred to me to derive a process for finding maxima and minima and thus for resolving easily all the

Text 12: Fermat on maxima and minima. From J. Fauvel and J. Gray, eds. (1987). *The History of Mathematics: A Reader*. London: Macmillan Press Ltd., pp. 359–360.

difficulties concerning limiting conditions which have caused so many problems for ancient and modern geometers.

Maxima and minima are in effect unique and singular, as Pappus said and as the ancients already knew, although Commandino claimed not to know what the term 'singular' signified in Pappus. It follows from this that on one side and the other of the point constituting the limit one can take an ambiguous equation, and that the two ambiguous equations thus obtained are accordingly correlative, equal and similar.

For example, let it be proposed to divide the line b in such a way that the product of the segments shall be a maximum. The point answering this question is evidently the middle of the given line, and the maximum product is equal to $b^2/4$; no other division of this line gives a product equal to $b^2/4$.

But if one proposes to divide the same line b in such a way that the product of the segments shall equal z'' (this area being besides supposed to be less than $b^2/4$) there will be two points answering the question, and they will be found situated on one side and the other of the point corresponding to the maximum product.

In fact let a be one of the segments of the line b, one will have $ba - a^2 = z''$; an ambiguous equation, since for the segment a one can take each of the two roots. Therefore let the correlative equation be $be - e^2 = z''$. Comparing the two equations according to the method of Viète:

$$ba - be = a^2 - e^2.$$

Dividing both sides by $a - e$, one obtains

$$b = a + e;$$

the lengths a and e will moreover be unequal.

If, in place of the area z'', one takes another greater value, although always less than $b^2/4$, the segments a and e will differ less from each other than the previous ones, the points of division approaching closer to the point constituting the maximum of the product. The more the product increases the more on the contrary diminishes the difference between a and e until it will vanish exactly at the division corresponding to the maximum product; in this case there will only be a unique and singular solution, the two quantities a and e becoming equal.

Now the method of Viète applied to the two correlative equations above leads to the equality $b = a + e$, therefore if $e = a$ (which will always happen at the point constituting the maximum or the minimum) one will have, in the case proposed, $b = 2a$, which is to say that if one takes the middle of the segment b, the product of the segments will be a maximum.

Let us take another example: to divide the segment b in such a way that the product of the square of one of the segments with the other shall be a maximum.

Let a be one of the segments; one must have $ba^2 - a^3$ maximum. The equal and similar correlative equation is $be^2 - e^3$. Comparing these two equations according to the method of Viète:

$$ba^2 - be^2 = a^3 - e^3;$$

dividing both sides by $a - e$ one obtains

$$ba + be = a^2 + ae + e^2,$$

which gives the form of the correlative equations.

13 WALLIS. COMPUTATION OF π BY SUCCESSIVE INTERPOLATIONS

After 1650, analytic methods began to receive more attention and to replace geometric methods based on the writings of the ancients. This was due partly to the acceptance into geometry of those algebraic methods that Descartes and Fermat had introduced, and partly to the still very active interest in numerical work—interpolation, approximation, logarithms—a heritage of the sixteenth and early seventeenth centuries. This tradition was strong in England, where Napier and Briggs had labored.

This analytic method advanced rapidly through the efforts of John Wallis (1616–1703), of Emmanuel College, Cambridge, who in 1649 became the Savilian professor of geometry at Oxford. He was one of the founders of the Royal Society and, through his work, influenced Newton, Gregory, and other mathematicians. In his *Arithmetica infinitorum* (Oxford, 1655), he led explorations into the realms of the infinite with daring analytic methods, using interpolation and extrapolation to obtain new results. The title of the book shows the difference between Wallis' method—he called it "arithmetica"; we would say (with Newton) "analysis"—and the geometric method of Cavalieri. First Wallis derived Cavalieri's integral in an original way. Thereupon, he plunged into a maelstrom of numerical work and, with fine mathematical intuition to guide him in his interpolations, arrived at the infinite product for π that bears his name. See J. F. Scott, *The mathematical work of John Wallis* (Taylor and Francis, Oxford, 1938); also A. Prag, "John Wallis," *Quellen und Studien zur Geschichte der Mathematik* (*B*) I (1931), 381–412.

Proposition 39.[1] Given a series of quantities that are the cubes of a series of numbers continuously increasing in arithmetic proportion (like the series of cubic numbers), which begin from a point or zero (say 0, 1, 8, 27, 64, . . .); we ask for the ratio of this series to the series of just as many numbers equal to the highest number of the first series.

[1] In previous propositions Wallis has derived the limit

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} i^k}{n^{k+1}} = \frac{1}{k+1}$$

for $k = 1, 2$. This Proposition 39 prepares for the case $k = 3$; it shows Wallis's typical inductive and analytic method.

The investigation is carried out by the inductive method, as before. We have

$$\frac{0 + 1 = 1}{1 + 1 = 2} = \frac{2}{4} = \frac{1}{4} + \frac{1}{4};$$

$$\frac{0 + 1 + 8 = 9}{8 + 8 + 8 = 24} = \frac{3}{8} = \frac{1}{4} + \frac{1}{8};$$

$$\frac{0 + 1 + 8 + 27 = 36}{27 + 27 + 27 + 27 = 108} = \frac{4}{12} = \frac{1}{4} + \frac{1}{12};$$

$$\frac{0 + 1 + 8 + 27 + 64 = 100}{64 + 64 + 64 + 64 + 64 = 320} = \frac{5}{16} = \frac{1}{4} + \frac{1}{16};$$

$$\frac{0 + 1 + \cdots + 125 = 225}{125 + \cdots + 125 = 750} = \frac{6}{20} = \frac{1}{4} + \frac{1}{20};$$

$$\frac{0 + \cdots + 125 + 216 = 441}{216 + \cdots + 216 = 1512} = \frac{7}{24} = \frac{1}{4} + \frac{1}{24};$$

and so forth.

The ratio obtained is always greater than one-fourth, or $\frac{1}{4}$. But the excess decreases constantly as the number of terms increases; it is $\frac{1}{4}, \frac{1}{8}, \frac{1}{12}, \frac{1}{16}, \frac{1}{20}, \frac{1}{24}, \cdots$ There is no doubt that the denominator of the fraction increases with every consecutive ratio by a multiple of 4, so that the excess of the resulting ratio over $\frac{1}{4}$ is the same as 1 : 4 times the number of terms after 0, etc.

Proposition 40. *Theorem.* Given a series of quantities that are the cubes of a series of numbers continuously increasing in arithmetic proportion beginning, for instance, with 0, then the ratio of this series to the series of just as many numbers equal to the highest number of the first series will be greater than $\frac{1}{4}$. The excess will be 1 divided by four times the number of terms after 0, or the cube root of the first term after 0 divided by four times the cube root of the highest term.

The sum of the series $0^3 + 1^3 + \cdots + l^3$ is $\frac{l+1}{4} l^3 + \frac{l+1}{4l} l^3$, or, if m is the number of terms, $\frac{m}{4} l^3 + \frac{m}{4l} l^3 = \frac{1}{4} ml^3 + \frac{1}{4} ml^2$. This is apparent from the previous reasoning.

If, with increasing number of terms, this excess over $\frac{1}{4}$ diminishes continuously, so that it becomes smaller than any given number (as it clearly does), when it goes to infinity, then it must finally vanish. Therefore:

Proposition 41. *Theorem.* If an infinite series of quantities which are the cubes of a series of continuously increasing numbers in arithmetic progression, beginning, say, with 0, is divided by the sum of numbers all equal to the highest and equal in number, then we obtain $\frac{1}{4}$. This follows from the preceding reasoning.

Proposition 42. *Corollary.* The complement AOT [Fig. 1] of half the area of the cubic parabola therefore is to the parallelogram TD over the same arbitrary base and altitude as 1 to 4.

Indeed, let AOD be the area of half the parabola AD (its diameter AD, and the corresponding ordinates DO, DO, etc.) and let AOT be its complement.

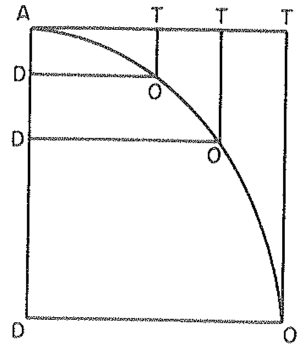


Fig. 1

Since the lines DO, DO, etc., or their equals AT, AT, etc. are the cube roots[2] of $AD, AD, \ldots$, or their equals $TO, TO, \ldots$, these TO, TO, etc. will be the cubes of the lines $AT, AT, \ldots$ The whole figure AOT therefore (consisting of the infinite number of lines TO, TO, etc., which are the cubes of the arithmetically progressing lines $AT, AT, \ldots$) will be to the parallelogram ATD (consisting of just as many lines, all equal to the greatest TO), as 1 to 4, according to our previous theorem. And the half-segment AOD of the parabola (the residuum of the parallelogram) is to the parallelogram itself as 3 is to 4.

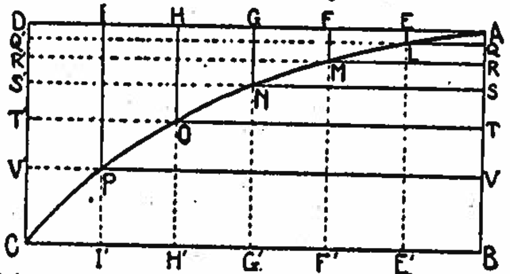
E. Walker, A Study of the Traité des Indivisibles of Gilles Persone de Roberval, New York 1932, pp. 181-182.

På et tidspunkt benytter han at for endelige n gælder:

$1+2^2+3^2+\ldots+n^2 = \frac{1}{3}n^3+\frac{1}{2}n^2+\frac{1}{6}n$

THE PARABOLA

Proposition 6. *The area of a parabola is equal to two-thirds of the area of the circumscribed rectangle.*



Let AC be a parabola with its vertex at A, and its diameter AB; and let BC be perpendicular to AB. Then the area $ABC = \frac{2}{3}$ of the area of the rectangle $ABCD$.

Proof. Draw AD tangent to the curve AC at A, and draw $CD \perp$ to BC. Divide AD into an infinite number of equal parts, and through the points of division draw lines parallel to AB, cutting the curve AC in the points $L, M, N, \ldots$. Through $L, M, N, \ldots$, draw ordinate parallel to CB. Now in the parabola we have the relation

$$\frac{EL}{FM} = \frac{AQ}{AR} = \frac{QL^2}{RM^2} = \frac{AE^2}{AF^2}, \quad \frac{FM}{GN} = \frac{AR}{AS} = \frac{RM^2}{SN^2} = \frac{AF^2}{AG^2},$$

. . . .

Let $AE = 1$, then $AF = 2$, $AG = 3, \ldots$

Then $AE^2 + AF^2 + AG^2 + \cdots =$

$QL^2 + RM^2 + SN^2 + \cdots = 1^2 + 2^2 + 3^2 + \cdots$

Hence

$$\frac{\text{the area } ADC}{\text{the area } ABCD} = \frac{AE(EL+FM+GN+\ldots)}{AD \cdot DC}$$

$$= \frac{AE(AE^2+AF^2+AG^2+\cdots+AD^2)}{AD \cdot AD^2}$$

$$= \frac{1^2+2^2+3^2+\cdots+AD^2}{AD \cdot AD^2} = \frac{1}{3}.$$

Therefore the area $ADC = \frac{1}{3}$ of the area $ABCD$, and therefore the area $ABC = \frac{2}{3}$ of the area $ABCD$.[1]

Q.E.D.

If, like M. Fermat, we should wish to investigate the parabola in which the parts of the diameter have the ratio of cubes, then, using the same diagram,

$$\frac{\text{the area } CAD}{\text{the area } ABCD} = \frac{1^3+2^3+3^3+\cdots+DC^3}{DC^4} = \frac{1}{4},$$

therefore the area $ABC = \frac{3}{4}$ of the area $ABCD$.

If the parts of the diameter have the ratio of fourth powers, then

$$\frac{\text{the area } CAD}{\text{the area } ABCD} = \frac{1^4+2^4+3^4+\cdots+DC^4}{DC^5} = \frac{1}{5},$$

therefore the area $ABC = \frac{4}{5}$ of the area $ABCD$.

This may be continued indefinitely through the higher powers.

(1)

Text 15: H. J. M. Bos (1980). "Newton, Leibniz and the Leibnizian Tradition". In: *From the Calculus to Set Theory, 1630–1910. An Introductory History*. Ed. by I. Grattan-Guinness. Princeton and Oxford: Princeton University Press. Chap. 2, pp. 49–93.

Chapter 2

Newton, Leibniz and the Leibnizian Tradition

H. J. M. Bos

2.1. *Introduction and biographical summary*

The starting-point of this chapter is the ' invention ', or rather ' inventions ', of the calculus. Both Newton (in 1664–1666) and Leibniz (in 1675) created, independently of each other, an infinitesimal calculus. Their inventions were very different in concepts and style, but each contains so much of what we now recognise as essential to the calculus that the expression ' invention of the calculus ' is justified in both cases. I go on to consider the subsequent development of the calculus till about 1780. In this development the Leibnizian type of calculus with differentials and integrals proved more successful than the Newtonian fluxional calculus ; therefore I concentrate on the former.

Many great and lesser mathematicians were involved in the development of the calculus in the period covered by this chapter. I shall restrict myself to those who played the prime roles in the story : Isaac Newton, Lucasian professor of mathematics at Cambridge and later Master of the Mint in London ; Gottfried Wilhelm Leibniz, historian and scientist at the ducal court of Hanover ; Jakob Bernoulli, professor of mathematics at Basle ; his brother Johann Bernoulli, younger by thirteen years, who after a professorate at Groningen succeeded Jakob in Basle in 1705 ; Guillaume François Marquis de l'Hôpital, a French nobleman living by private means, and an able mathematician eagerly interested in the new developments in infinitesimal methods ; and finally Leonhard Euler, who studied with Johann Bernoulli and then entered a career in the typically 18th-century scientific institutions, the academies. He was professor at the St. Petersburg (now Leningrad) Academy from 1730 to 1741 and from 1766 till his death ; in the intervening years he served the Berlin Academy as professor.

Many of the great ideas that were to make Isaac Newton famous in mathematics and natural science came to him in the years 1664–1666.

At that time he was a graduate student at Trinity College, Cambridge, but for some time during those two years he lived in Lincolnshire, staying away from Cambridge for fear of the Plague (compare Whiteside *1966a*). His ideas on gravity, which he was to work out later and present to the world in his famous *Principia* (*1687a*), date from that period, as well as his theory of colours, published in the treatise *Opticks* in 1704, the binomial series theorem and his fluxional calculus, which we shall discuss in more detail in section 2.2.

As with gravity and colours, publication of these mathematical ideas in print was long delayed. Newton did compose several accounts of his findings in infinitesimal calculus. In October 1666 he summarised the discoveries of the fruitful two years in a tract on fluxions (*1666a*) ; in 1669 he wrote a treatise on infinite series, the *De analysi* (*1669a*), which circulated in manuscript form among members of the Royal Society ; from 1671 dates a treatise on the method of fluxions and infinite series (*1671a*) ; and in about 1693 he composed a treatise on the quadrature of curves (*1693a*). However, the 1666 tract and the treatise on the method of fluxions were not published in his lifetime, the *De analysi* was published only in 1711, and the treatise on quadratures of curves in 1704. Meanwhile the *Principia* of 1687 had brought for the first time to the general public indications of his methods in infinitesimal calculus, but these were not enough to show the scope and power of his mathematical discoveries.

About the turn of the century a fair amount was published about Leibniz's calculus (as we shall see in sections 2.5–2.8 below), and sufficient information about Newton's calculus was available to show that both men had found new methods in essentially the same mathematical field. This caused a nasty quarrel over priority, in which feelings of personal and national pride combined with insufficient insight in the mathematics involved (at least in the case of the lesser participants in the debate) to create a distasteful muddle of misunderstandings and insinuations which has only been cleared up through patient historical research in the present century. The net result of the historical research is that Leibniz found his calculus later than Newton and independently of him, and that he published it earlier.

In 1669 Newton had succeeded Isaac Barrow as Lucasian professor, but in the 1690s he grew dissatisfied with his position at Cambridge. He visited London often, to attend meetings of the Royal Society, of which he was a fellow from 1672, and to be present at sessions of Parliament as a member for the Cambridge University constituency. He moved finally to London in 1696 when he was offered the office of Warden of the Mint. In 1703 he became president of the Royal Society, a post which he held till his death. His position as the most eminent British scientist was further emphasised by a knighthood in 1705.

By the 1710s so much on the fluxional calculus was in print that the method was taken up and applied by others. However, this further development of the Newtonian type of calculus remained restricted to Great Britain, and it did not achieve much. Reasons of the lack of success lie in the isolation from the Continental developments in analysis because of the priority dispute, in the lack of mathematicians in Britain of sufficient stature to really develop Newton's calculus, and in an over-stressed loyalty to Newton's conception of the calculus and to his notations, which were less versatile than Leibniz's.

On the Continent Leibniz's inventions gave rise to a much more intense development, to whose origins in the 1670s we now turn.

Before Leibniz entered the service of the house of Hanover in 1676 he had spent four years in Paris on a diplomatic mission, which left him ample time to pursue his interest in mathematics, the sciences, history, philosophy and many other things. He met many French philosophers and made two visits to London to the Royal Society. The Paris years were his formative period. When he arrived in 1672 his knowledge of mathematics was slight, despite the fact that he had published a small tract on combinatorics. He was trained in law at the university of his home town of Leipzig. In Paris Christiaan Huygens, who lived there at that time, recognised Leibniz's mathematical abilities and guided his first studies in the higher mathematics. Leibniz's 'growth to mathematical maturity' (see Hofmann *1949a*) was indeed impressive; it led to his discovery of the calculus in 1675, the elaboration of that calculus in the following years and its publication in 1684–1686. He contributed to other branches of mathematics as well, for instance to algebra (solvability of equations, determinants) and to nearly all other fields of human learning, including religion, politics, history, physics, mechanics, technology, mathematics, geology, linguistics and natural history. Many of his results were not immediately published and became known only gradually, through correspondence (from his comparative intellectual isolation in Hanover Leibniz corresponded with over a thousand scholars), through publication of short articles in journals (he was one of the founders of the first scientific journal in Germany, the *Acta eruditorum*), and later through the publication of his manuscripts, most of which he kept and which are now stored at the Leibniz archive in Hanover.

Leibniz's publication of his calculus in two articles in the *Acta* of 1684 and 1686 did not provoke great commotion in mathematical circles. The articles were rather short, and they were marred by misprints and in places deliberately obscure, so that it is in fact surprising that in the following decade they were understood at all.

Jakob and Johann Bernoulli studied the articles from 1687, and by 1690 they showed, in articles published in the *Acta*, that they had mastered the Leibnizian symbolism and its use. They both started a correspondence with Leibniz ; the contact between Johann and Leibniz was especially intensive and productive. After 1690 a stream of articles in the *Acta* and in other journals, written by the Bernoullis and Leibniz and later joined by l'Hôpital and others, showed the learned world that the new calculus was something to be reckoned with.

However, for people of lesser mathematical calibre than the Bernoullis, it would have been very difficult actually to learn the calculus from these articles. What was wanted was a proper textbook of the calculus. Such a textbook came, though only of the differential calculus, in 1696 with l'Hôpital's *Analyse des infiniment petits pour l'intelligence des lignes courbes* (' Analysis of infinitely small quantities for the understanding of curved lines ' : *1696a*).

The Marquis de l'Hôpital was introduced to the calculus by Johann Bernoulli, who, after finishing his medical studies in 1690, had travelled to Paris, where he impressed learned circles by a method to determine, by means of differentials, the curvature of arbitrary curves—a problem which by the methods of Cartesian analytic geometry was well nigh unsolvable. l'Hôpital was most impressed and asked Bernoulli to give him, for a good fee, lectures on the new method. Bernoulli accepted and the lectures were given, in Paris and at the country chateau of the Marquis. They were written out and both men kept copies. After about a year Bernoulli left Paris but agreed to continue instructing l'Hôpital by letter. In fact the agreement was that Bernoulli, for a handsome monthly salary, would answer all l'Hôpital's questions concerning mathematics, would send him all his mathematical discoveries and would give no one else access to these findings (see Bernoulli *Correspondence*, 144) ; a most curious and hardly honourable agreement which put Bernoulli's originality strictly in l'Hôpital's service. From the start Bernoulli did not quite keep to the letter of the contract, and l'Hôpital soon realised that he could not bind a brilliant mathematician in this way. But when in 1696 l'Hôpital published his textbook, and Bernoulli saw that most of its content was taken from his lectures with not more than a passing reference to the Marquis's indebtedness to Bernoulli, he could only be angry in silence, being bound by the contract.

Later, after l'Hôpital's death, Johann Bernoulli did try to get his part in the *Analyse* acknowledged, but by that time his credibility in priority questions had become very low because of open quarrels on such matters with his brother. Jakob Bernoulli was a rather introverted personality, but he was sensitive to praise from members of the mathematical community and he resented being overshadowed by his brilliant younger brother. Johann, on the other hand, liked his own success too much to spare his brother's feelings. So there appeared insinuating remarks in articles, and later a quarrel exploded and went on quite openly. Johann Bernoulli's claim to much of the content of the *Analyse* was found to be justified only when in 1921 the manuscript of his Paris lectures on the differential calculus was found (see Johann Bernoulli *1924a*).

However strained their mutual relations, through the writings of these men the Leibnizian calculus became known and proved its power. By the first decade of the 18th century other mathematicians devoted themselves to the new calculus, such as Jakob Hermann, Pierre Varignon, Niklaus Bernoulli (a nephew) and Daniel Bernoulli (son of Johann). The family Bernoulli continued to yield famous mathematicians throughout the 18th century.

In these early days the new calculus consisted mainly of rather loosely connected methods, and problems solved by these methods. The man who reshaped the Leibnizian calculus into a soundly organised body of mathematical knowledge was Leonhard Euler. Euler was the central figure of continental mathematics in the middle years of the 18th century. He published an enormous number of books and articles on mathematics, mechanics, optics, astronomy, navigation, hydrodynamics, technical matters such as artillery and shipbuilding, and very many other topics. He maintained this impressive productivity despite losing the sight of one eye in 1735 and becoming completely blind in 1766. His position at the academies involved him in many other tasks besides scientific research, such as advice on the performance of new inventions as fire-engines and pumps, and on technological enterprises like canal-building and the construction of water-works in the park of the royal palace *Sans Souci* of Prussia's Frederick the Great.

Euler's greatest influence on the calculus and on analysis in general was through his great textbooks, in which he gave analysis a definitive form, which it was to keep until well into the 19th century. These textbooks, written in Latin, were: *Introductio ad analysin infinitorum* (' Introduction to the analysis of infinites ' : *1748a*), *Institutiones calculi differentialis* (' Textbooks on the differential calculus ' : *1755b*), and *Institutiones calculi integralis* (' Textbooks on the integral calculus ' : *1768–1770a*).

These were the men who created the calculus and shaped the Leibnizian tradition in analysis. In sections 2.3–2.8 I shall describe the mathematics involved, but first I shall devote the next section to an overview of the Newtonian calculus.

2.2. *Newton's fluxional calculus*

As was mentioned above, Newton's main mathematical discoveries in the infinitesimal calculus date from 1664 to 1666. (For a detailed account of his achievements in this period, see Newton *Papers*, vol. 1, 145–154, and *Works*$_2$, vol. 1, viii–xiii.) Autodidactically he quickly acquired adequate knowledge of existing theories in the field, benefitting especially from reading Descartes's *La géométrie* in van Schooten's edition with commentaries, and from the works of Wallis. Starting from these studies he developed in these fruitful two years his *fluxional calculus*.

In Newton's discoveries, complex, deep and many-sided as they are, a number of central themes may be distinguished. These are: series expansions, algorithms, the inverse relationship of differentiation and integration, the conception of variables as moving in time, and the doctrine of prime and ultimate ratios. Although these themes are interconnectedly present in almost all of his studies in the infinitesimal calculus, I shall deal with them separately.

Newton valued *power-series expansions* very highly, because they provide a means to reduce the analytical formulae of curves to a form in which all terms simply consist of a constant times a power of the variable. Thus transcendental curves (admitting no algebraic equation), as well as algebraic curves with complicated equations, can be represented by much simpler equations (be it with an infinite number of terms). Newton saw that this has two great advantages. Firstly, series expansion makes it possible to apply rules and algorithms which are defined for simple equations only, to a much wider range of curves. In particular, the relation

$$\int x^n \, dx = \frac{1}{n+1} x^{n+1}, \qquad (2.2.1)$$

which was known in various forms by the 1660s (see sections 1.10 and 1.11) can be used, in combination with power-series expansions, to provide series expressions for the quadratures of almost all curves. Secondly, series expansion provides a ready means for the approximation and simplification of formulae through the discarding of higher-order terms—a feature which he used with virtuosity in his applications of his mathematical methods to physical problems.

Newton's most famous series expansion is the 'binomial theorem', which he found in the winter of 1664–1665 and which states that the well-known binomial expansion for integer powers n,

$$(a+x)^n = a^n + \frac{n}{1} a^{n-1}x + \frac{n(n-1)}{1\,.\,2} a^{n-2}x^2 + \ldots + x^n, \qquad (2.2.2)$$

can be generalised for fractional powers $\alpha = p/q$, in which case the right hand side of

$$(a+x)^\alpha = a^\alpha + \frac{\alpha}{1} a^{\alpha-1}x + \frac{\alpha(\alpha-1)}{1\,.\,2} a^{\alpha-2}x^2 + \ldots \qquad (2.2.3)$$

is an infinite series. He found the theorem in connection with the problem of squaring the circle $y=(1-x^2)^{1/2}$. He compared the formulae $(1-x^2)^0$, $(1-x^2)^{1/2}$, $(1-x^2)^{2/2}$, $(1-x^2)^{3/2}$, $(1-x^2)^{4/2}$, ... The first, third, fifth, ... formulae involve no root, and therefore the quadratures of the corresponding curves are easily found :

$$\left.\begin{array}{l} \text{quadrature of } y=(1-x^2)^0 \text{ is } x, \\ \text{quadrature of } y=(1-x^2)^{2/2} \text{ is } x-\tfrac{1}{3}x^3, \\ \text{quadrature of } y=(1-x^2)^{4/2} \text{ is } x-\tfrac{2}{3}x^3+\tfrac{1}{5}x^5. \end{array}\right\} \qquad (2.2.4)$$

On examining the coefficients in these expansions, Newton noted that the denominators are the odd numbers 1, 3, 5, 7, ... and that the numerators are, in the successive expansions, $\{1\}$, $\{1, 1\}$, $\{1, 2, 1\}$, $\{1, 3, 3, 1\}$, ..., that is, the numbers in the ' Pascal triangle ', which he knew could be expressed for successive integral values of n as

$$\left\{1,\ n,\ \frac{n(n-1)}{1\,.\,2},\ \frac{n(n-1)(n-2)}{1\,.\,2\,.\,3},\ \ldots\right\}.$$

He then guessed that, by analogy, the same expressions would apply for *fractional* values of n. When $n=\frac{1}{2}$ this yields :
quadrature of $y=(1-x^2)^{1/2}$ is

$$x - \frac{\frac{1}{2}x^3}{3} - \frac{\frac{1}{8}x^5}{5} - \frac{\frac{1}{16}x^7}{7} - \frac{\frac{5}{128}x^9}{9} - \ldots \qquad (2.2.5)$$

He then saw that this procedure of guessing, or ' interpolating ', expansions such as (2.2.5) from the scheme of the series (2.2.4) could be applied to the equations of the curves as well as to their quadratures, and in this way he found that

$$(1-x^2)^{1/2} = 1 - \tfrac{1}{2}x^2 - \tfrac{1}{8}x^5 - \tfrac{1}{16}x^6 - \tfrac{5}{128}x^8 - \ldots \qquad (2.2.6)$$

Not satisfied with the reliability of the interpolation procedure, he checked (2.2.6) in two ways. He showed that the product of the right hand side of (2.2.6) with itself yields $1-x^2$ (that is, all further coefficients in the product series are zero), and he saw that a common method of root extraction known as the ' galley method ', applied formally to $1-x^2$, yields the same series. In the same way as with root extraction, he used the algorithm of long division to obtain series expansions, for instance,

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \ldots, \qquad (2.2.7)$$

which provided the quadrature of the hyperbola $y = 1/(1+x)$. He also obtained (2.2.7) by assuming that the binomial expansion applied when $n = -1$.

In the *De analysi* (*1669a*), in which these methods of series expansions are explained and used, Newton also provides a general rule to compute, for a given polynomial equation

$$\sum a_{ij} x^i y^j = 0 \qquad (2.2.8)$$

between x and y, the first coefficients of the pertaining series

$$y = \sum b_i x^i \qquad (2.2.9)$$

(*Papers*, vol. 2, 222–247).

Both in the way that Newton found the binomial theorem and in the application of series expansions in general, the relation, which we now write as

$$\int x^n \, dx = \frac{1}{n+1} x^{n+1}, \qquad (2.2.10)$$

plays an important role. He mentioned this 'quadrature of simple curves' at the outset of his *De analysi*: 'RULE 1. If $ax^{m/n} = y$, then will $(na/(m+n))x^{(m+n)/n}$ equal the area ABD' (*ibid.*, 206–207; see figure 2.2.1). Later in that treatise he gave a general procedure (of which rule 1 is a direct consequence) for finding the relation between the quadrature of a curve (as AD in figure 2.2.1) and its ordinate. The procedure makes it clear that Newton recognised the *inverse relationship of integration and differentiation* (although, of course, he did not use these terms). He explains his method by means of an example, from which, however, the generality of the procedure is quite clear. He

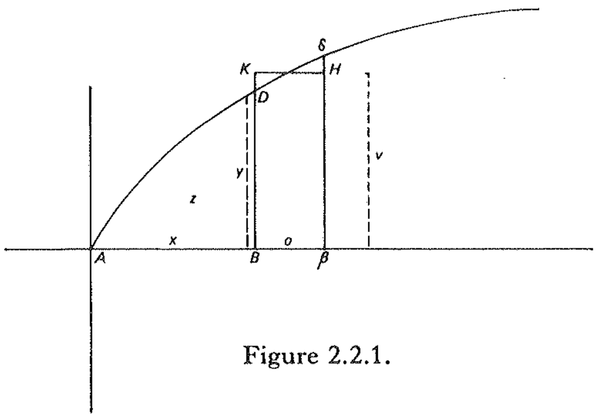


Figure 2.2.1.

proceeds as follows (*ibid.*, 242–245). In figure 2.2.1 let area $ABD = z$, $BD = y$ and $AB = x$; let further $B\beta = o$ and let $BK = v$ be chosen such that area $BD\delta\beta$ = area $BKHb = ov$. Consider, as example, the curve for which

$$z = \tfrac{2}{3}x^{3/2}, \tag{2.2.11}$$

that is (removing roots to get a polynomial equation),

$$z^2 = \tfrac{4}{9}x^3\,; \tag{2.2.12}$$

then also

$$(z + ov)^2 = \tfrac{4}{9}(x + o)^3, \tag{2.2.13}$$

from which

$$z^2 + 2zov + o^2v^2 = \tfrac{4}{9}(x^3 + 3x^2o + 3xo^2 + o^3). \tag{2.2.14}$$

Now by removing the terms without o, which are equal on both sides from (2.2.12), and dividing the remainder by o, we obtain

$$2zv + ov^2 = \tfrac{4}{9}(3x^2 + 3xo + o^2). \tag{2.2.15}$$

Now Newton takes $B\beta$ ' infinitely small ', in which case, as the figure suggests, $v = y$ and the terms with o vanish :

$$2zy = \tfrac{4}{3}x^2. \tag{2.2.16}$$

Inserting the value of z from (2.2.11), he obtains

$$y = x^{1/2}. \tag{2.2.17}$$

Clearly the procedure is applicable to all polynomial relations between x and z. It consists in essence of calculating the derivative (in this case the y) for any algebraic function z of x.

Newton saw clearly that the problem of quadratures was to be approached in this inverse way : by calculating y for all manner of algebraic z, he could find all manner of curves (y, x) which are quadrable. Indeed, he calculated many such quadrable curves, writing them together in extensive lists, which are thus nothing less than the first tables of integrals (compare *Papers*, vol. 1, 404–411).

The essential element in the foregoing procedure is the substitution of ' small ' corresponding increments o and ov for x and z in the equation. In studies on the determination of maxima and minima, tangents and curvature, Newton had extensively made use of this method, and he had worked out various ***algorithms*** for these problems, by which he could calculate the slope of the tangent or the curvature in any point of an algebraic curve. (In modern terms, he had developed algorithms to determine the derivative of any algebraic function.) Later he reformulated these algorithms and their proofs in terms of fluents and fluxions, and we shall come back to them after discussing these concepts.[1]

[1] Compare, for instance, Newton *1671a*, in *Papers*, vol. 3, 72–73.

The terms 'fluents' and 'fluxions' indicate Newton's conception of variable quantities in analytical geometry: he saw these as 'flowing quantities', that is, *quantities that change with respect to time*. Thus, when considering the curve of figure 2.2.1, he would conceive the point D as moving along the curve, while correspondingly the ordinate y, the abscissa x, the quadrature z or any other variable quantity connected with the curve would increase or decrease, or in general change or 'flow'. He called these flowing quantities 'fluents' (as opposed to the constant quantities occurring in the figure or in the problem at hand), and he called their rate of change with respect to time their 'fluxion'. In his earlier researches he indicated fluxions by separate letters; in *1671a* he introduced the dot-notation, where the fluxions of the fluents x, y, z are $\dot{x}$, $\dot{y}$, $\dot{z}$ respectively.

It should be remarked that the way in which the fluents vary with time is arbitrary. Newton often makes, for simplicity, an additional assumption about the movement of the variables, supposing that one of the variables, say x, moves uniformly, so that $\dot{x}=1$. Such assumptions can be made because the values of the fluxions themselves are not of interest but rather their ratio, such as $\dot{y}/\dot{x}$, which gives the slope of the tangent. By this conception of quantities moving in time Newton thought himself able to solve the foundational difficulties inherent in considering 'small' corresponding increments of variables, which are so small that we may discard them, and yet are not equal to zero, as we want to divide through by them. In his approach to this problem, his theory of *prime and ultimate ratios*, which we shall discuss in section 2.10, his conception of flowing quantities is essential; through this conception he comes very near to a use of limits as foundation of the calculus.

We now return to the *algorithms* mentioned above. The corresponding increments of variables, can be expressed in terms of fluxions: let o now be an infinitesimal element of time, then the corresponding increments of the fluents $x, y, z, \ldots$ are $\dot{x}o, \dot{y}o, \dot{z}o, \ldots$ respectively. The ratio of $\dot{y}$ to $\dot{x}$ can now be determined in a way which is evident in the following example, which Newton gives himself in *1671a* (*Papers*, vol. 3, 79–81). Let a curve be given with equation

$$x^3-ax^2+axy-y^3=0. \qquad (2.2.18)$$

Substituting $x+\dot{x}o$ and $y+\dot{y}o$ for x and y respectively yields

$$(x^3+3\dot{x}ox^2+3\dot{x}^2o^2x+\dot{x}^3o^3)-(ax^2+2a\dot{x}ox+a\dot{x}^2o^2) \\ +(axy+a\dot{x}oy+a\dot{y}ox+a\dot{x}\dot{y}o^2) \\ -(y^3+3\dot{y}oy^2+3\dot{y}^2o^2y+\dot{y}^3o^3)=0. \qquad (2.2.19)$$

Deleting $x^3-ax^2+axy-y^3$ as equal to zero from (2.2.18), dividing through by o and discarding the terms in which o is left, yields

$$3\dot{x}x^2-2a\dot{x}x+a\dot{x}y+a\dot{y}x-3\dot{y}y^2=0, \qquad (2.2.20)$$

from which the ratio of $\dot{y}$ and $\dot{x}$ is easily obtained :

$$\frac{\dot{y}}{\dot{x}}=\frac{3x^2-2ax+ay}{3y^2-ax}. \qquad (2.2.21)$$

We note that the numerator and the denominator in the result are (apart from a sign) the partial derivatives f_x and f_y of $f(x,y)=x^3-ax^2+axy-y^3$, the left hand side of the equation of the curve. Thus

$$\frac{\dot{y}}{\dot{x}}=-\frac{f_x}{f_y}. \qquad (2.2.22)$$

Indeed, this relation is implicit in the algorithms which, as we mentioned before, Newton worked out for problems of tangents, maxima and minima, and curvature. He even at one time introduced special notations in this connection (see *Papers*, vol. 1, 289–294), writing $\mathscr{X}$ for the left hand side of the equation of the curve (with the right hand side zero). He then wrote $\cdot\mathscr{X}$ and $\mathscr{X}\cdot$ for what we would write as xf_x and yf_y respectively (the so-called 'homogeneous partial derivatives'), using further symbols for homogeneous higher-order partial derivatives occurring in connection with curvature. However, the connection of Newton's $\cdot\mathscr{X}$ and $\mathscr{X}\cdot$ with modern partial derivatives should not be considered without some qualifications ; he defined them formally as modifications of the formula $\mathscr{X}$, and he did not explicitly view $\mathscr{X}$ as a function of two variables which assumes also other values than the zero in the equation.

With these algorithms, and further finesses which we cannot go into here, Newton was able to solve what he formulated as one of the two fundamental problems in infinitesimal calculus : given the fluents and their relations, to find the fluxions.

The second problem is the converse of the first : given the relation of the fluxions, to find the relation of the fluents. Transposed in modern terminology, this means : given a differential equation, to find its solution. This of course is a much harder problem than the first. Newton did more about the problem than formulate it ; his integral tables, already mentioned, form a means toward its solution, and he also studied various individual differential equations (or rather, fluxional equations).

As we have seen in the previous section, Newton's calculus was not to have the influence which Leibniz's achieved. Therefore, within the space and organisation of this chapter, we must leave it at this short summary of the fluxional calculus and some more remarks on its foundations in section 2.10, turning now our attention to the more successful rival, the Leibnizian calculus.

2.3. *The principal ideas in Leibniz's discovery*

One of the most precious documents of the Leibniz archive at Hanover is a set of mathematical manuscripts dated 25, 26 and 29 October, and 1 and 11 November, 1675.[1] On these sheets Leibniz wrote down his thoughts, more or less as they came to him, during a study of that most important problem of 17th-century mathematics : to find methods for the quadrature of curves. In the course of these studies he came to introduce the symbols '$\int$' and 'd', to explore the operational rules which they obey in formulas, and to apply them in translating many geometrical arguments about the quadrature of curves into symbols and formulas. In short, these manuscripts contain the record of Leibniz's 'invention' of the calculus. We will discuss them in more detail below, but first we will mention three principal ideas which guided him in those fateful studies in 1675.

The first principal idea was a philosophical one, namely Leibniz's idea of a *characteristica generalis*, a general symbolic language, through which all processes of reason and argument could be written down in symbols and formulas ; the symbols would obey certain rules of combination which would guarantee the correctness of the arguments. This idea guided him in much of his philosophical thinking ; it also explains his great interest in notation and symbols in mathematics and in general his endeavour to translate mathematical statements and methods into formulas and algorithms. Thus, in studying the geometry of curves, he was interested in methods rather than in results, and especially in ways to transform these methods into algorithms performable with formulas. In short, he was looking for a *calculus* for infinitesimal-geometrical problems.

The second principal idea concerned difference sequences. In studying sequences $a_1, a_2, a_3, \ldots$, and the pertaining difference sequences $b_1=a_1-a_2$, $b_2=a_2-a_3$, $b_3=a_3-a_4, \ldots$, Leibniz had noted that

$$b_1+b_2+\ldots+b_n=a_1-a_{n+1}. \qquad (2.3.1)$$

This means that difference sequences are easily summed, an insight which he put to good use in solving a problem which Huygens suggested

[1] They are discussed in Hofmann *1949a*, and an English translation is given in Child *1920a*.

to him in 1672: to sum the series $\frac{1}{1}+\frac{1}{3}+\frac{1}{6}+\frac{1}{10}+\frac{1}{15}+\ldots$, the denominators being the so-called 'triangular numbers' $r(r+1)/2$. He found that the terms can be written as differences,

$$\frac{2}{r(r+1)}=\frac{2}{r}-\frac{2}{r+1}, \tag{2.3.2}$$

and hence

$$\sum_{r=1}^{n}\frac{2}{r(r+1)}=2-\frac{2}{n+1}. \tag{2.3.3}$$

In particular, the series, when summed to infinity has sum 2. This result motivated him to study a whole scheme of related sum and difference sequences, which he put together in his so-called 'harmonic triangle' (figure 2.3.1), in which the oblique rows are successive difference sequences, so that their sums can be easily read off from the scheme (Leibniz *Writings*, vol. 5, 405: compare Hofmann *1949a*, 12; *1974a*, 20).

$$1$$
$$\tfrac{1}{2}\quad \tfrac{1}{2}$$
$$\tfrac{1}{3}\quad \tfrac{1}{6}\quad \tfrac{1}{3}$$
$$\tfrac{1}{4}\quad \tfrac{1}{12}\quad \tfrac{1}{12}\quad \tfrac{1}{4}$$
$$\tfrac{1}{5}\quad \tfrac{1}{20}\quad \tfrac{1}{30}\quad \tfrac{1}{20}\quad \tfrac{1}{5}$$
$$\tfrac{1}{6}\quad \tfrac{1}{30}\quad \tfrac{1}{60}\quad \tfrac{1}{60}\quad \tfrac{1}{30}\quad \tfrac{1}{6}$$
$$\tfrac{1}{7}\quad \tfrac{1}{42}\quad \tfrac{1}{105}\quad \tfrac{1}{140}\quad \tfrac{1}{105}\quad \tfrac{1}{42}\quad \tfrac{1}{7}$$
$$\cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot\quad \cdot$$

Figure 2.3.1.

Leibniz's 'harmonic triangle'. The numbers in the n-th row are

$$\left[(n-1)\binom{n}{k}\right]^{-1}.$$

Summations can be read off from the scheme as, for example:

$$\frac{1}{3}+\frac{1}{12}+\frac{1}{30}+\frac{1}{60}+\frac{1}{105}+\ldots=\frac{1}{2}.$$

These results were not exactly new, but they did make Leibniz aware that the forming of difference sequences and of sum sequences are mutually inverse operations. This principal idea became more significant when he transposed it to geometry. The curve in figure 2.3.2 defines a sequence of equidistant ordinates y. If their distance is 1, the sum of the y's is an approximation of the quadrature of the curve, and the difference of two successive y's yields approximately the slope of the pertaining tangent. Moreover, the smaller the unit 1 is chosen, the better the approximation. Leibniz concluded that if the unit could

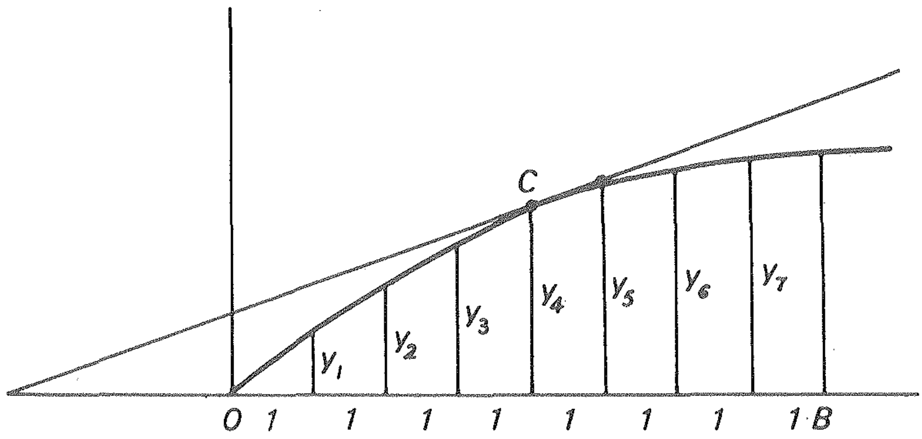


Figure 2.3.2.

be chosen *infinitely small*, the approximations would become exact: in that case the quadrature would be equal to the sum of the ordinates, and the slope of the tangent would be equal to the difference of the ordinates. In this way, he concluded from the reciprocity of summing and taking differences that the determination of quadratures and tangents are also mutually inverse operations.

Thus Leibniz's second principal idea, however vague as it was in about 1673, suggested already an infinitesimal calculus of sums and differences of ordinates by which quadratures and tangents could be determined, and in which these determinations would occur as inverse processes. The idea also made plausible that, just as in sequences the determination of differences is always possible but the determination of sums is not, so in the case of curves the tangents are always easily to be found, but not so the quadratures.

The third principal idea was the use of the 'characteristic triangle' in transformations of quadratures. In studying the work of Pascal, Leibniz noted the importance of the small triangle $cc'd$ along the curve in figure 2.3.3, for it was (approximately) similar to the triangles formed

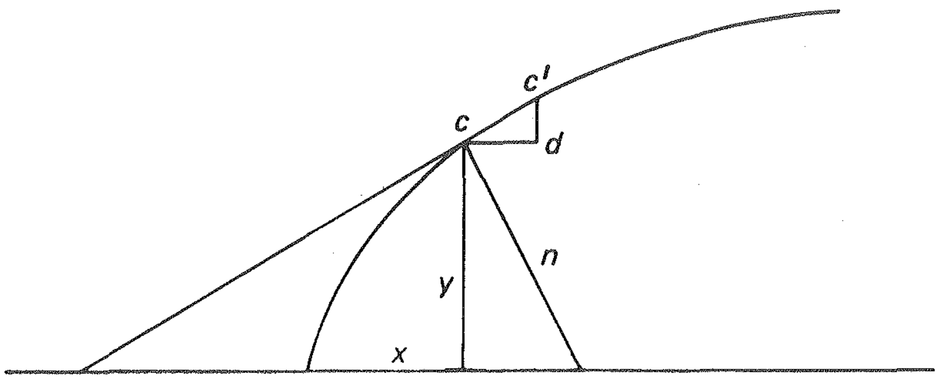


Figure 2.3.3.

by ordinate, tangent and sub-tangent, or ordinate, normal and subnormal. The configuration occurs in many 17th-century mathematical works ; Pascal's use of it concerned the circle. Leibniz saw its general use in finding relations between quadratures of curves and other quantities like moments and centres of gravity. For instance, the similarity of the triangles yields $cc' \times y = cd \times n$; hence

$$\sum cc' \times y = \sum cd \times n. \qquad (2.3.4)$$

The left hand side can be interpreted as the total moment of the curve arc with respect to the x-axis (the moment of a particle with respect to an axis is its weight multiplied by its vertical distance to the axis), whereas the right hand side can be interpreted as the area formed by plotting the normals along the x-axis.

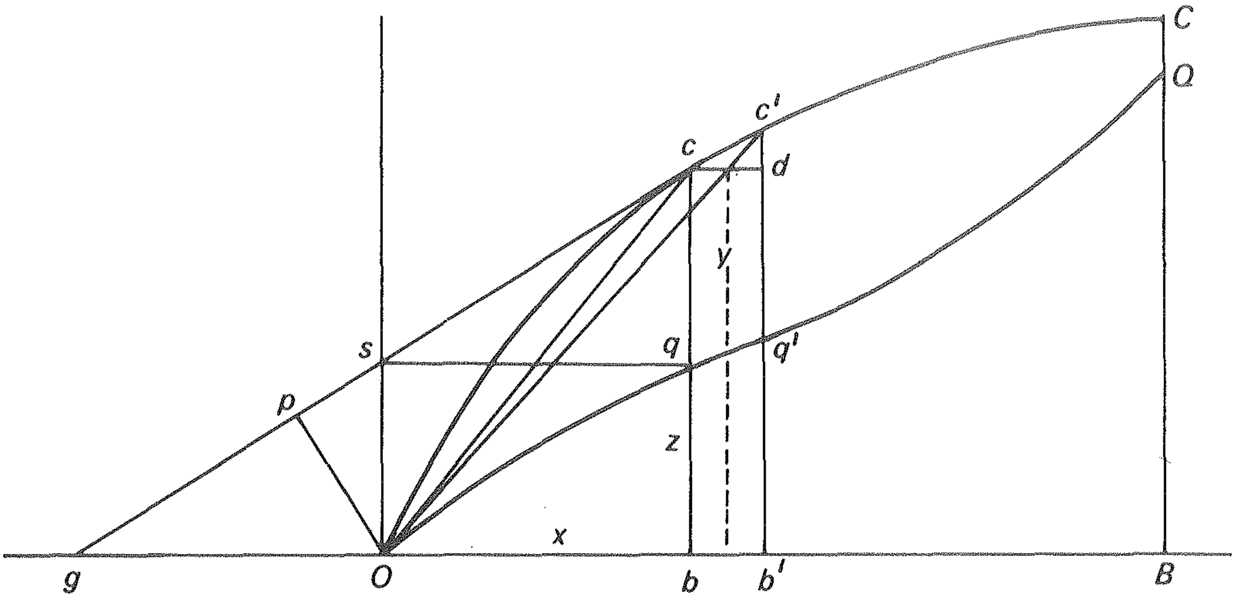


Figure 2.3.4.

As an example of Leibniz's use of the characteristic triangle, here is his derivation of a special transformation of quadratures which he called 'the transmutation' and which, for good reasons, he valued highly (compare Hofmann *1949a*, 32–35 (*1974a*, 54–60), and Leibniz *Writings*, vol. 5, 401–402). In figure 2.3.4 let the curve $Occ'C$ be given, with characteristic triangle cdc' at c. Its quadrature $\mathcal{Q} = \widehat{OCB}$, the sum of the strips $bcc'b'$, can also be considered as the sum of the triangles Occ' supplemented by the triangle OBC :

$$\mathcal{Q} = \sum \Delta Occ' + \Delta OBC. \qquad (2.3.5)$$

Now

$$\Delta Occ' = \tfrac{1}{2}cc' \times Op$$
$$= \tfrac{1}{2}cd \times Os$$

(since the characteristic triangle cdc' is similar to ΔOsp)

$$= \tfrac{1}{2}bqq'b'. \qquad (2.3.6)$$

Now for each c on $Occ'C$ we can find the corresponding q by drawing the tangent, determining s and taking $bq = Os$. Thus we form a new curve $Oqq'Q$, and we have from (2.3.5) :

$$\mathcal{Q} = \tfrac{1}{2}\text{ (quadrature } Oqq'Q) + \Delta OCB. \qquad (2.3.7)$$

This is Leibniz's transmutation rule which, through the use of the characteristic triangle, yields a transformation of the quadrature of a curve into the quadrature of another curve, related to the original curve through a process of taking tangents. It can be used in those cases where the quadrature of the new curve is already known, or bears a known relation to the original quadrature. Leibniz found this for instance to be the case with the general parabolas and hyperbolas (see section 1.3), for which the rule gives the quadratures very easily. He also applied his transmutation rule to the quadrature of the circle, in which investigation he found his famous arithmetical series for π :

$$\frac{\pi}{4} = 1 - \tfrac{1}{3} + \tfrac{1}{5} - \tfrac{1}{7} + \tfrac{1}{9} - \tfrac{1}{11} + \dots . \qquad (2.3.8)$$

The success of the transmutation rule also convinced him that the analytical calculus for problems of quadratures which he was looking for would have to cover transformations such as this one by appropriate symbols and rules.

The transmutation rule as Leibniz discovered it in 1673 belongs to the style of geometrical treatment of problems of quadrature which was common in the second half of the 17th century. Similar rules and methods can be found in the works of Huygens, Barrow, Gregory and others. Barrow's *Lectiones geometricae* (*1670a*), for instance, contain a great number of transformation rules for quadratures which, if translated from his purely geometrical presentation into the symbolism and notation of the calculus, appear as various standard alogrithms of the differential and integral calculus. This has even been used (by J. M. Child in his *1920a*) as an argument to give to Barrow, rather than Newton or Leibniz, the title of inventor of the calculus. However, this view can be sustained only when one disregards completely the effect of the translation of Barrow's geometrical text into analytical formulas. It is the very possibility of the analytical expression of methods, and hence the understanding of their logical coherence and generality, which was the great advantage of Newton's and Leibniz's discoveries.

It is appropriate to illustrate this advantage by an example. To do this, I shall give a translation, with comments, of Leibniz's transmutation rule into analytical formulas.

The ordinate z of the curve $Oqq'Q$ is, by construction,

$$z = y - x\frac{dy}{dx} \tag{2.3.9}$$

(note the use of the characteristic triangle). The transmutation rule states that, for $OB = x_0$,

$$\int_0^{x_0} y\,dx = \tfrac{1}{2}\int_0^{x_0} z\,dx + \tfrac{1}{2}x_0y_0. \tag{2.3.10}$$

Inserting z from (2.3.9), we find

$$\begin{aligned}\int_0^{x_0} y\,dx &= \tfrac{1}{2}\int_0^{x_0}\left(y - x\frac{dy}{dx}\right)dx + \tfrac{1}{2}x_0y_0\\ &= \tfrac{1}{2}\int_0^{x_0} y\,dx - \tfrac{1}{2}\int_0^{x_0} x\frac{dy}{dx}\,dx + \tfrac{1}{2}x_0y_0.\end{aligned}$$

Hence

$$\int_0^{x_0} y\,dx + \int_0^{x_0} x\frac{dy}{dx}\,dx = x_0y_0, \tag{2.3.11}$$

so that we recognise the rule as an instance of ' integration by parts '.

Apart from the indication of the limits of integration $(0, x_0)$ along the ∫-sign, the symbolism used above was found by Leibniz in 1675. The advantages of that symbolism over the geometrical deduction and statement of the rule are evident : the geometrical construction of the curve $Oqq'Q$ is described by a simple formula (2.3.9), and the formalism carries the proof of the rule with it, as it were. (2.3.11) follows immediately from the rule

$$d(x\,y) = x\,dy + y\,dx. \tag{2.3.12}$$

These advantages, manipulative ease and transparency through the rules of the symbolism, formed the main factors in the success of Leibniz's method over its geometrical predecessors.

But we have anticipated in our story. So we return to October 1675, when the transmutation rule was already found but not yet the new symbolism.

2.4. *Leibniz's creation of the calculus*

In the manuscripts of 25 October–11 November 1675 we have a close record of studies of Leibniz on the problem of quadratures. We find him attacking the problem from several angles, one of these being the use of the Cavalierian symbolism ' omn.' in finding, analytically (that is, by manipulation of formulas) all sorts of relations between quadratures. ' Omn. ' is the abbreviation of ' omnes lineae ', ' all lines ' ; in section 1.10 it was represented by the symbol ' $\mathcal{O}$ '.

A characteristic example of Leibniz's investigations here is the following. In a diagram such as figure 2.4.1 he conceived a sequence of

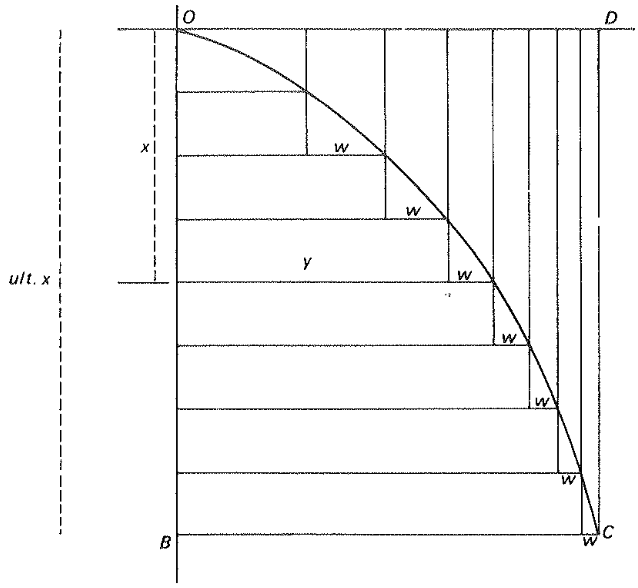


Figure 2.4.1.

ordinates y of the curve $\widehat{OC}$; the distance between successive ordinates is the (infinitely small) unit. The differences of the successive ordinates are called w. OBC is then equal to the sum of the ordinates y. The rectangles like $w \times x$ are interpreted as the moments of the differences w with respect to the axis OD (moment = weight × distance to axis). Hence the area OCD represents the total moment of the differences w. $\widehat{OCB}$ is the complement of $\widehat{OCD}$ within the rectangle $ODCB$, so that Leibniz finds that ' The moments of the differences about a straight line perpendicular to the axis are equal to the complement of the sum of the terms ' (Child *1920a*, 20). The ' terms ' are the y. Now w is the difference sequence of the sequence of ordinates y ; hence, conversely, y is the sum-sequence of the w's, so that we may eliminate y

and consider only the sequence w and its sum-sequences, which yields : ' and the moments of the terms are equal to the complement of the sum of the sums ' (*ibid.*). Here the ' terms ' are the w. Leibniz writes this result in a formula using the symbol ' Omn. ' for what he calls ' a sum '. We give the formula as he gave it, and we add an explanation under the accolades ; $\sqcap$ is his symbol for equality, ' ult. x ' stands for *ultimus* x, the last of the x, that is, OB, and he uses overlining and commas where we would use brackets (*ibid.*) :

$$\underbrace{\text{omn. }\overline{xw}}_{\substack{\text{moments of}\\ \text{the terms } w}} \sqcap \underbrace{\underbrace{\text{ult. } x, \overline{\text{omn. } w},}_{\text{total}} - \underbrace{\overline{\text{omn. omn. } w}}_{\substack{\text{sum of the sums}\\ \text{of the terms}}}}_{\substack{\text{complement of the sum of the sums}\\ \text{of the terms}}} \tag{2.4.1}$$

(Compare the form of (2.4.1) with that of (2.3.11).) Immediately he sees the possibility to obtain from this formula, by various substitutions, other relations between quadratures. For instance, substitution of $xw = a$, $w = a/x$ yields

$$\text{omn. } a \sqcap \text{ult. } x, \text{ omn. } \frac{a}{x} - \text{omn. omn. } \frac{a}{x}, \tag{2.4.2}$$

which he interprets as an expression of the ' sum of the logarithms in terms of the quadrature of the hyperbola ' (*ibid*,. 71). Indeed, omn. a/x is the quadrature of the hyperbola $y = a/x$, and this quadrature is a logarithm, so that omn. omn. a/x is the sum of the logarithms.

We see in these studies an endeavour to deal analytically with problems of quadrature through appropriate symbols and notations, as well as a clear recognition and use of the reciprocity relation between difference and sum sequences. In a manuscript of some days later, these insights are pushed to a further consequence. Leibniz starts here from the formula (2.4.1), now written as

$$\text{omn. } xl \sqcap x \text{ omn. } l - \text{omn. omn. } l. \tag{2.4.3}$$

He stresses the conception of the sequence of ordinates with infinitely small distance : ' . . . l is taken to be a term of the progression, and x is the number which expresses the position or order of the l corresponding to it ; or x is the ordinal number and l is the ordered thing ' (*ibid.*, 80). He now notes a rule concerning the dimensions in formulas like (2.4.3), namely that omn., prefixed to a line, such as l, yields an area (the quadrature) ; omn., prefixed to an area, like xl, yields a solid, and so on. Such a law of dimensional homogeneity was well-known from the Cartesian analysis of curves, in which the formulas must consist of terms all of the same dimension. (In (2.4.3) all terms are of three dimensions, in $x^2+y^2=a^2$ all terms are of two dimensions ; an expression like a^2+a is, if dimensionally interpreted, unacceptable, for it would express the sum of an area and a line.)

This consideration of dimensional homogeneity seems to have suggested to Leibniz to use a single letter instead of the symbol ' omn. ', for he goes on to write : ' It will be useful to write ∫ for omn, so that $\int l$ stands for omn. l or the sum of all l's ' (*ibid.*). Thus the ∫-sign is introduced. ' ∫ ' is one of the forms of the letter ' s ' as used in script (or italics print) in Leibniz's time : it is the first letter of the word *summa*, sum. He immediately writes (2.4.3) in the new formalism :

$$\int xl = x\int l - \int\int l\,; \qquad (2.4.4)$$

he notes that

$$\int x = x^2/2 \quad \text{and} \quad \int x^2 = x^3/3, \qquad (2.4.5)$$

and he stresses that these rules apply for ' series in which the differences of the terms bear to the terms themselves a ratio that is less than any assigned quantity ' (*ibid.*), that is, series whose differences are infinitely small.

Some lines further on we also find the introduction of the symbol ' d ' for differentiating. It occurs in a brilliant argument which may be rendered as follows : The problem of quadratures is a problem of summing sequences, for which we have introduced the symbol ' ∫ ' and for which we want to elaborate a *calculus*, a set of useful algorithms. Now summing sequences, that is, finding a general expression for $\int y$ for given y, is usually not possible, but it *is* always possible to find an expression for the differences of a given sequence. This finding of differences is the reciprocal calculus of the calculus of sums, and therefore we may hope to acquire insight in the calculus of sums by working out the reciprocal calculus of differences. To quote Leibniz's own words (*ibid.*, 82) :

> Given l, and its relation to x, to find $\int l$. This is to be obtained from the contrary calculus, that is to say, suppose that $\int l = ya$. Let $l = ya/d$; then just as ∫ will increase, so d will diminish the dimensions. But ∫ means a sum, and d a difference. From the given y, we can always find y/d or l, that is, the difference of the y's.

Thus the ' d '-symbol (or rather the symbol ' $1/d$ ') is introduced Because Leibniz interprets ∫ dimensionally, he has to write the ' d ' in the denominator : l is a line, $\int l$ is an area, say ya (note the role of ' a ' to make it an area), the differences must again be lines, so we must write ' ya/d '. In fact he soon becomes aware that this is a notational disadvantage which is not outweighed by the advantage of dimensional interpretability of $\int$ and d, so he soon writes '$d(ya)$' instead of 'ya/d' and henceforth re-interprets 'd' and '$\int$' as dimensionless symbols. Nevertheless, the consideration of dimension did guide the decisive steps of choosing the new symbolism.

In the remainder of the manuscript Leibniz explores his new symbolism, translates old results into it and investigates the operational rules for $\int$ and d. In these investigations he keeps for some time to the idea that $d(uv)$ must be equal to $du\,dv$, but finally he finds the correct rule

$$d(uv) = u\,dv + v\,du. \tag{2.4.6}$$

Another problem is that he still for a long time writes $\int x$, $\int x^2$, . . . for what he is later to write consistently as $\int x\,dx$, $\int x^2\,dx$, ...

A lot of this straightening out of the calculus was still to be done after 11 November 1675 ; it took Leibniz roughly two years to complete it. Nevertheless, the manuscripts which we discussed contain the essential features of the new, the Leibnizian, calculus : the concepts of the differential and the sum, the symbols d and $\int$, their inverse relation and most of the rules for their use in formulas.

Let us summarise shortly the main features of these Leibnizian concepts (compare Bos *1974a*, 12–35). The *differential* of a variable y is the infinitely small difference of two successive values of y. That is, Leibniz conceives corresponding sequences of variables such as y and x in figure 2.4.2. The successive terms of these sequences lie infinitely close. dy is the infinitely small difference of two successive ordinates y, dx is the infinitely small difference of two successive abscissae x, which, in this case, is equal to the infinitely small distance of two successive y's. A sum (later termed 'integral' by the Bernoullis) like $\int y\,dx$ is the sum of the infinitely small rectangles $y \times dx$. Hence the quadrature of the curve is equal to $\int y\,dx$.

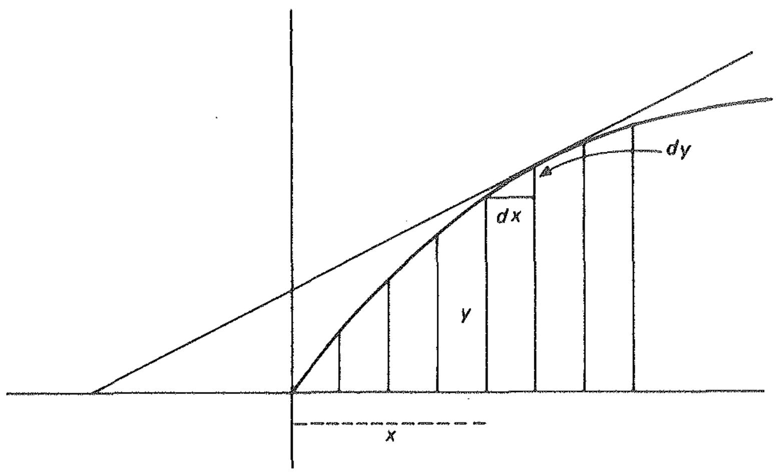


Figure 2.4.2.

Leibniz was rather reluctant to present his new calculus to the general mathematical public. When he eventually decided to do so, he faced the problem that his calculus involved infinitely small quantities, which were not rigorously defined and hence not quite acceptable in mathematics. He therefore made the radical but rather unfortunate decision to present a quite different concept of the differential which was not infinitely small but which satisfied the same rules. Thus in his first publication of the calculus, the article 'A new method for

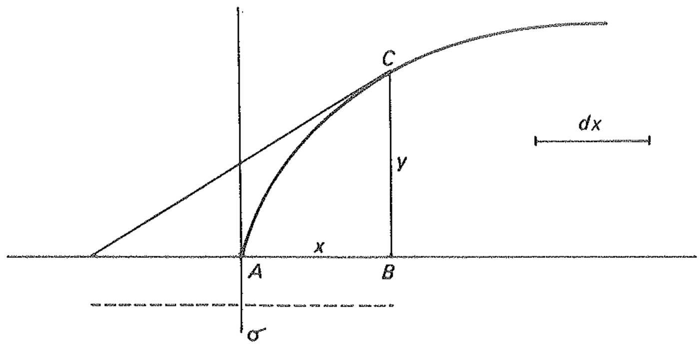


Figure 2.4.3.

maxima and minima as well as tangents' (*1684a*) in the issue for October 1684 of the *Acta*, he introduced a fixed finite line-segment (see figure 2.4.3) called dx, and he defined the dy at C as the line-segment satisfying the proportionality

$$y : \sigma = dy : dx, \tag{2.4.7}$$

σ being the length of the sub-tangent, or

$$dy = \frac{y}{\sigma}\, dx. \tag{2.4.8}$$

So defined, dy is also a finite line-segment. Leibniz presented the rules of the calculus for these differentials, and indicated some applications. In an article published two years later (*1686a*) he gave some indications about the meaning and use of the ∫-symbol. This way of publication of his new methods was not very favourable for a quick and fruitful reception in the mathematical community. Nevertheless, the calculus was accepted, as we shall see in the following sections.

2.5. *l'Hôpital's textbook version of the differential calculus*

Leibniz's publications did not offer an easy access to the art of his new calculus, and neither did the early articles of the Bernoullis. Still, a good introduction appeared surprisingly quickly, at least to the differential calculus, namely l'Hôpital's *Analyse* (*1696a*).

As a good textbook should, the *Analyse* starts with definitions, of variables and their differentials, and with postulates about these differentials. The definition of a differential is as follows: 'The infinitely small part whereby a variable quantity is continually increased or decreased, is called the differential of that quantity' (ch. 1). For further explanation l'Hôpital refers to a diagram (figure 2.5.1), in which,

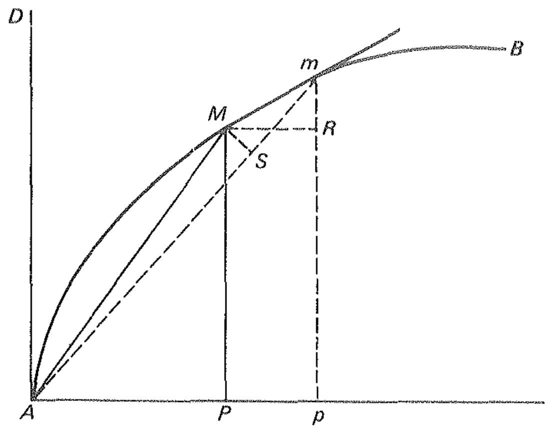


Figure 2.5.1.

with respect to a curve AMB, the following variables are indicated: abscissa $AP = x$, ordinate $PM = y$, chord $AM = z$, arc $\widehat{AM} = s$ and quadrature $\widehat{AMP} = \mathcal{Q}$. A second ordinate pm 'infinitely close' to PM is drawn, and the differentials of the variables are seen to be: $dx = Pp$, $dy = mR$, $dz = Sm$, $ds = Mm$ (the chord Mm and the arc Mm are taken to coincide) and $d\mathcal{Q} = MPpm$. l'Hôpital explains that the 'd' is a special symbol, used only to denote the differential of the variable written after it. The small lines Pp, mR, . . . in the figure have to be considered as 'infinitely small'. He does not enter into the question whether such quantities exist, but he specifies, in the two postulates, how they behave (*ibid.*):

> Postulate I. Grant that two quantities, whose difference is an infinitely small quantity, may be used indifferently for each other: or (which is the same thing) that a quantity, which is increased or decreased only by an infinitely smaller quantity, may be considered as remaining the same.

This means that AP may be considered equal to Ap (or $x = x + dx$), MP equal to mp ($y = y + dy$), and so on.

The second postulate claims that a curve may be considered as the assemblage of an infinite number of infinitely small straight lines, or equivalently as a polygon with an infinite number of sides. The first postulate enables l'Hôpital to derive the rules of the calculus, for instance :

$$\left.\begin{aligned} d(xy) &= (x+dx)(y+dy)-xy \\ &= x\,dy+y\,dx+dx\,dy \\ &= x\,dy+y\,dx \end{aligned}\right\} \quad (2.5.1)$$

' because $dx\,dy$ is a quantity infinitely small, in respect of the other terms $y\,dx$ and $x\,dy$: for if, for example, you divide $y\,dx$ and $dx\,dy$ by dx, we shall have the quotients y. and dy, the latter of which is infinitely less than the former ' (*ibid.*, ch. 1, para. 5). l'Hôpital's concept of differential differs somewhat from Leibniz's. Leibniz's differentials are infinitely small differences between successive values of a variable. l'Hôpital does not conceive variables as ranging over a sequence of infinitely close values, but rather as continually increasing or decreasing ; the differentials are the infinitely small parts by which they are increased or decreased.

In the further chapters l'Hôpital explains various uses of differentiation in the geometry of curves : determination of tangents, extreme values and radii of curvature, the study of caustics, envelopes and various kinds of singularities in curves. For the determination of tangents he remarks that postulate 2 implies that the infinitesimal part Mm of the curve in figure 2.5.2, when prolonged, gives the tangent.

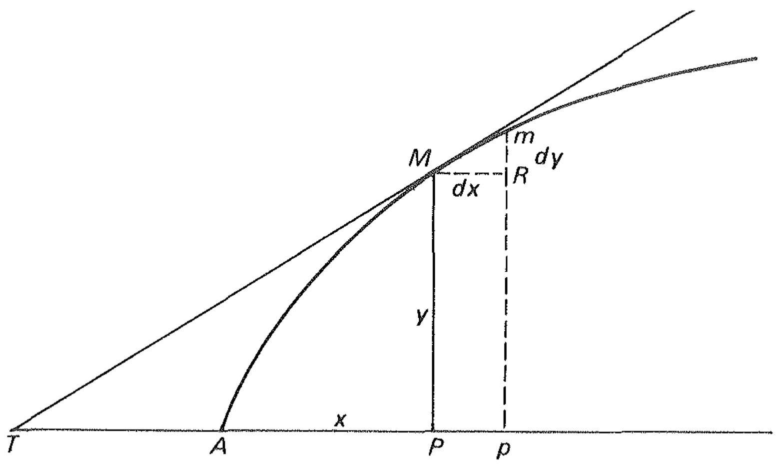


Figure 2.5.2.

Therefore $Rm : RM$, or $dy : dx$, is equal to $y : PT$, so that $PT = y(dx/dy)$, and the tangent can be constructed once we have determined $y\,dx/dy$ (*ibid.*, ch. 2, para. 9) :

> Now by means of the difference of the given equation you can obtain a value of dx in terms which all contain dy, and if you multiply by y and divide by dy you will obtain an expression for the sub-tangent PT entirely in terms of known quantities and free from differences, which will enable you to draw the required tangent MT.

To explain this, consider for example the curve $ay^2 = x^3$. The 'difference of the equation' is derived by taking differentials left and right:

$$2ay\, dy = 3x^2\, dx. \qquad (2.5.2)$$

dx can now be expressed in terms of dy:

$$dx = \frac{2ay}{3x^2}\, dy. \qquad (2.5.3)$$

Hence

$$PT = \frac{y\, dx}{dy} = y\, \frac{2ay}{3x^2} = \frac{2ay^2}{3x^2}, \qquad (2.5.4)$$

which provides the construction of the tangent.

The 'difference of the equation' is a true differential equation, namely an equation between differentials. l'Hôpital considers expressions like 'dy/dx' actually as quotients of differentials, not as single symbols for derivatives.

2.6. *Johann Bernoulli's lectures on integration*

In 1742, more than fifty years after they were written down, Johann Bernoulli published his lectures to l'Hôpital on 'the method of integrals' in his collected works (Bernoulli *1691a*), stating in a footnote that he omitted his lectures on differential calculus as their contents were now accessible to everyone in l'Hôpital's *Analyse*. His lectures may be considered as a good summary of the views on integrals and their use in solving problems which were current around 1700.

Bernoulli starts with defining the integral as the inverse of the differential: the integrals of differentials are those quantities from which these differentials originate by differentiation. This conception of the integral—the term, in fact, was introduced by the Bernoulli brothers—differs from Leibniz's, who considered it as a sum of infinitely small quantities. Thus, in Leibniz's view, $\int y\, dx = \mathcal{Q}$ means that the sum of the infinitely small rectangles $y \times dx$ equals $\mathcal{Q}$; for Bernoulli it means that $d\mathcal{Q} = y\, dx$.

Bernoulli states that the integral of $ax^p\, dx$ is $(a/(p+1))x^{p+1}$, and he gives various methods usable in finding integrals ; among them is the method of substitution, explained by several examples, such as the following (*1691a*, lecture 1) :

> Suppose that one is required to find the integral of
>
> $$(ax+xx)\,dx\,\sqrt{(a+x)}.$$
>
> Substituting $\sqrt{(a+x)}=y$ we shall obtain $x=yy-a$, and thus $dx=2y\,dy$, and the whole quantity
>
> $$(ax+xx)\,dx\,\sqrt{(a+x)}=2y^6\,dy-2ay^4\,dy.$$
>
> It is now easy and straightforward to integrate this expression ; its integral is $\frac{2}{7}y^7-\frac{2}{5}ay^5$ and, after substituting the value of y, we find the integral to be $\frac{2}{7}(x+a)^3\sqrt{(x+a)}-\frac{2}{5}a(x+a)^2\sqrt{(x+a)}$.

The principal use of the integral calculus, Bernoulli goes on to explain, is in the squaring of areas. For this the area has to be considered as divided up into infinitely small parts (strips, triangles, or quadrangles in general as in figure 2.6.1). These parts are the differentials of the areas ; one has to find an expression for them ' by means of determined letters and only one kind of indeterminate ' (*ibid.*, lecture 2), that is, an expression $f(u)\,du$ for some variable u. The required area is then equal to the integral $\int f(u)\,du$.

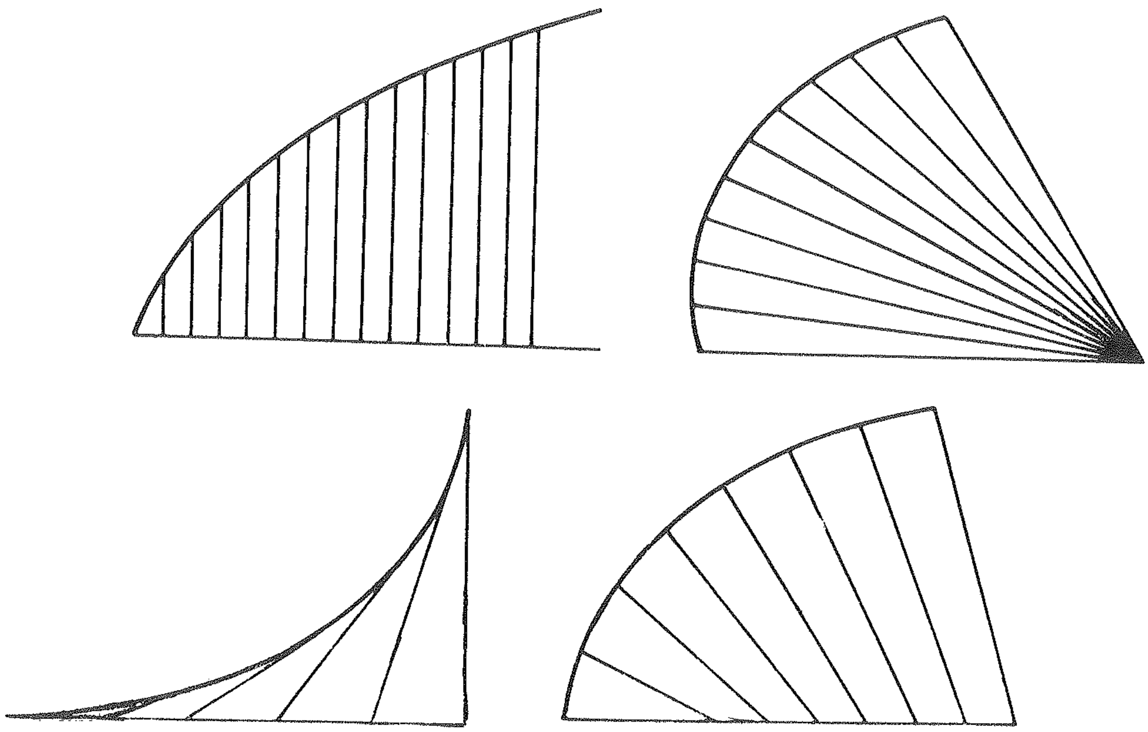


Figure 2.6.1.

The further use of the method of integrals is in the so-called ' inverse method of tangents ' (*ibid.*, lecture 8). The method, or rather the type of problem which Bernoulli has in mind here, originated in the 17th century ; it concerns the determination of a curve from a given property of its tangents. He teaches that the given property of the tangents has to be expressed as an equation involving differentials, that is, a differential equation. From this differential equation the equation of the curve itself has to be found by means of the method of integrals. His first example is (*ibid.*, lecture 8 ; see figure 2.6.2) :

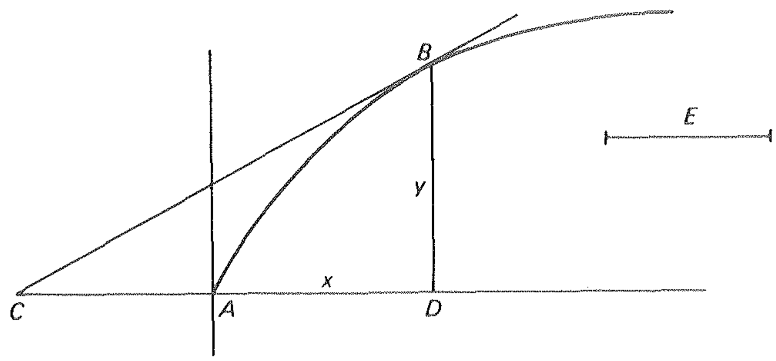


Figure 2.6.2.

It is asked what kind of curve AB it is whose ordinate BD is always the middle proportional between a given line E and the subtangent CD (that is, $E : BD = BD : CD$). Let $E=a$, $AD=x$, $DB=y$, then $CD=yy : a$. Now $dy : dx = y : CD = yy/a$ (that is, $CD=yy/a$) ; therefore we get the equation $y\,dx = yy\,dy : a$ or $a\,dx = y\,dy$; and after taking integrals on both sides, we get $ax = \frac{1}{2}yy$ or $2ax = yy$; which shows that the required curve AB is the parabola with parameter $=2a$.

In the further lectures Bernoulli solves many instances of inverse tangent problems. He devotes considerable attention to the question how to translate the geometrical or often mechanical data of the problem into a treatable differential equation. The problems treated in his lectures concern, among other things, the rectification (computation of the arc-length) of curves, cycloids, logarithmic spirals, caustics (linear foci occurring when light-rays reflect or refract on curved surfaces), the catenary (see section 2.8 below), and the form of sails blown by the wind.

2.7. *Euler's shaping of analysis*

In the (about) 50 years after the first articles on the calculus appeared, the Leibnizian calculus developed from a loose collection of methods for problems about curves into a coherent mathematical discipline: Analysis. Though many mathematicians, such as Jean le Rond d'Alembert, Alexis Clairaut, the younger generation of Bernoullis, and others, contributed to this development, it was in a large measure the work of one man: Leonhard Euler. Not only did Euler contribute many new discoveries and methods to analysis, but he also unified and codified the field by his three great textbooks mentioned already in section 2.1.

Shaping analysis into a coherent branch of mathematics meant first of all making clear what the subject was about. In the period of Leibniz, the elder Bernoullis and l'Hôpital, the calculus consisted of analytical methods for the solution of problems about curves; the principal objects were *variable geometrical quantities* as they occurred in such problems. However, as the problems became more complex and the manipulations with the formulas more intricate, the geometrical origin of the variables became more remote and the calculus changed into a discipline merely concerning formulas. Euler accentuated this transition by affirming explicitly that analysis is a branch of mathematics which deals with *analytical expressions*, and especially with *functions*, which he defined (following Johann Bernoulli) as follows: 'a function of a variable quantity is an analytical expression composed in whatever way of that variable and of numbers and constant quantities' (*1748a*, vol. 1, para. 4). Expressions like x^n, $(b+x)^2ax$ (with constants a and b) were functions of x. Algebraic expressions in general, and also infinite series, were considered as functions. The constants and the variable quantities could have imaginary or complex values.

Euler undertook the inventorisation and classification of that wide realm of functions in the first part of his *Introduction to the analysis of infinites* (*1748a*). The *Introduction* is meant as a survey of concepts and methods in analysis and analytical geometry preliminary to the study of the differential and integral calculus. He made of this survey a masterly exercise in introducing as much as possible of analysis without using differentiation or integration. In particular, he introduced the elementary transcendental functions, the logarithm, the exponential function, the trigonometric functions and their inverses without recourse to integral calculus—which was no mean feat, as the logarithm was traditionally linked to the quadrature of the hyperbola and the trigonometric functions to the arc-length of the circle.

Euler had to use some sort of infinitesimal process in the *Introduction*, namely, the expansion of functions in power-series (through long division, binomial expansion or other methods) and the substitution of infinitely large or infinitely small numbers in the formulas. A characteristic example of this approach is the deduction of the series expansion for a^z (*1748a*, vol. 1, paras. 114–116), where he proceeds as follows. Let $a > 1$, and let ω be an 'infinitely small number, or a fraction so small that it is just not equal to zero'. Then

$$a^{\omega} = 1 + \psi \tag{2.7.1}$$

for some infinitely small number ψ. Now put

$$\psi = k\omega \tag{2.7.2}$$

in which k depends only on a; then

$$a^{\omega} = 1 + k\omega \tag{2.7.3}$$

and

$$\omega = \log(1 + k\omega) \tag{2.7.4}$$

if the logarithm is taken to the base a.

Euler shows that for $a = 10$ the value of k can be found (approximately) from the common table of logarithms. He now writes

$$a^{i\omega} = (1 + k\omega)^i \tag{2.7.5}$$

for any (real) number i, so that by the binomial expansion

$$a^{i\omega} = 1 + \frac{i}{1} k\omega + \frac{i(i-1)}{1\,.\,2} k^2\omega^2 + \frac{i(i-1)(i-2)}{1\,.\,2\,.\,3} k^3\omega^3 + \ldots . \tag{2.7.6}$$

If z is any finite positive number, then $i = z/\omega$ is infinitely large, and by substituting $\omega = z/i$ in (2.7.6) we obtain

$$a^z = a^{i\omega} = 1 + \frac{1}{1} kz + \frac{1(i-1)}{1\,.\,2i} k^2z^2 + \frac{1(i-1)(i-2)}{1\,.\,2i\,.\,3i} k^3z^3 + \ldots . \tag{2.7.7}$$

But if i is infinitely large, $(i-1)/i = 1$, $(i-2)/i = 1$, and so on, and we arrive at

$$a^z = 1 + \frac{kz}{1} + \frac{k^2z^2}{1\,.\,2} + \frac{k^3z^3}{1\,.\,2\,.\,3} + \frac{k^4z^4}{1\,.\,2\,.\,3\,.\,4} + \ldots . \tag{2.7.8}$$

The *natural logarithms* arise if a is chosen such that $k = 1$. Euler gives that value of a up to 23 decimals, introduces the now familiar notation e for that number and writes (*ibid.*, para. 123):

$$e^z = 1 + \frac{z}{1} + \frac{z^2}{1\,.\,2} + \frac{z^3}{1\,.\,2\,.\,3} + \ldots . \tag{2.7.9}$$

In the next chapter Euler deals with trigonometric functions. He writes down the various sum-formulas and adds: 'Because $(\sin . z)^2 + (\cos . z)^2 = 1$, we have, by factorising, $(\cos . z + \sqrt{-1} . \sin . z)(\cos . z - \sqrt{-1} . \sin . z) = 1$, which factors, although imaginary, nevertheless are of immense use in comparing and multiplying arcs' (*ibid.*, para. 132).

He further finds that

$$(\cos y \pm \sqrt{-1}\sin y)(\cos z \pm \sqrt{-1}\sin z) = \cos(y+z) \pm \sqrt{-1}\sin(y+z), \qquad (2.7.10)$$

and hence

$$(\cos z \pm \sqrt{-1}\sin z)^n = \cos nz \pm \sqrt{-1}\sin nz, \qquad (2.7.11)$$

a relation usually called ' de Moivre's formula ' as it occurs already in the work of Abraham de Moivre (see Schneider *1968a*, 237–247).

By expanding (2.7.11) Euler obtains expressions for cos nz and sin nz. Now taking z to be infinitely small (so that $\sin z = z$ and $\cos z = 1$), $nz = v$ finite and hence n infinitely large, he arrives, by methods similar to those above, at

$$\cos v = 1 - \frac{v^2}{1.2} + \frac{v^4}{1.2.3.4} - \frac{v^6}{1.2.3.4.5.6} + \ldots, \qquad (2.7.12)$$

$$\sin v = v - \frac{v^3}{1.2.3} + \frac{v^5}{1.2.3.4.5} - \frac{v^7}{1.2.3.4.5.6.7} + \ldots \qquad (2.7.13)$$

(*ibid.*, para. 134). Some paragraphs later (art. 138) we find, derived by similar methods, the identities :

$$\exp(\pm v\sqrt{-1}) = \cos v + \sqrt{-1}\sin v, \qquad (2.7.14)$$

$$\cos v = \tfrac{1}{2}(\exp[v\sqrt{-1}] + \exp[-v\sqrt{-1}]), \qquad (2.7.15)$$

$$\sin v = \frac{1}{2\sqrt{-1}}(\exp[v\sqrt{-1}] - \exp[-v\sqrt{-1}]). \qquad (2.7.16)$$

Euler's *Textbooks on the differential calculus* (*1755b*) starts with two chapters on the calculus of finite differences and then introduces the differential calculus as a calculus of infinitely small differences, thus returning to a conception more akin to Leibniz's than to l'Hôpital's : ' The analysis of infinites . . . will be nothing else than a special case of the method of differences expounded in the first chapter, which occurs, when the differences, which previously were supposed finite, are taken infinitely small ' (*1755b*, para. 114). He considers infinitely small quantities as being in fact equal to zero, but capable of having finite ratios ; according to him, the equality $0 . n = 0$ implies that $0/0$ may in cases be equal to n. The differential calculus investigates the values of such ratios of zeros. Euler proceeds to discuss the differentiation of functions of one or several variables, higher-order differentiation and differential equations. He also obtains the equality

$$\frac{\partial^2 V}{\partial x \partial y} = \frac{\partial^2 V}{\partial y \partial x} \qquad (2.7.17)$$

for a function V of x and y (though not using this notation, and without obtaining a fully rigorous proof ; *1755b*, paras. 288 ff.).

In his discussion of higher-order differentiation Euler gives a prominent role to the *differential coefficients*, $p, q, r, \ldots$ defined, for a function $y = f(x)$, as follows :

$$dy = p\,dx \tag{2.7.18}$$

(where p is the coefficient with which to multiply the constant dx in order to obtain dy, so that p is again a function of x) ; and similarly,

$$dp = q\,dx \text{ (so that } ddy = q(dx)^2\text{)}, \tag{2.7.19}$$

$$dq = r\,dx \text{ (so that } dddy = r(dx)^3\text{)}, \ldots . \tag{2.7.20}$$

These differential coefficients are, though differently defined, equal to the first- and higher-order *derivatives* of the function f. In his textbook on the integral calculus he treats higher-order differential equations in terms of these differential coefficients, thus, in some measure, paving the way for the replacement of the differential by the derivative as fundamental concept of the calculus.

The three-volume *Textbooks on the integral calculus* (*1768–1770a*) give a magisterial close to the trilogy of textbooks. Here Euler gives a nearly complete discussion of the integration of functions in terms of algebraic and elementary transcendental functions, he discusses various definite integrals (including those now called the beta and gamma functions), and he gives a host of methods for the solution of ordinary and partial differential equations.

Apart from determining, through these textbooks, the scope and style of analysis for at least the next fifty years, Euler contributed to the infinitesimal calculus in many other ways. Two of these contributions are worth special emphasis. Firstly, he gave a thorough treatment of the calculus of variations, whose beginnings lie in the studies by the Bernoullis of the brachistochrone and of isoperimetric problems (see section 2.8 below). Secondly, he applied analysis, and indeed worked out many new analytical methods, in the context of studies in mechanics, celestial mechanics, hydrodynamics and many other branches of natural sciences, thus transforming these subjects into strongly mathematised form. In the next section I shall describe one example of each of these ways.

2.8. *Two famous problems : the catenary and the brachistochrone*

In writing the history of the calculus, it is customary to devote much attention to the fundamental concepts and methods. This tends to obscure the fact that most mathematicians spend most of their time not in contemplating these concepts and methods, but in using them to solve problems. Indeed, in the 18th century the term 'mathematics' comprised much more than the calculus and analysis, for it ranged from arithmetic, algebra and analysis through astronomy, optics, mechanics and hydrodynamics to such technological subjects as artillery, ship-building and navigation. In this section I discuss two famous problems whose solution was made possible by the new methods of the differential and integral calculus; in the next section I shall say something about what more was made possible through these methods.

The catenary problem

The catenary is the form of a hanging fully flexible rope or chain (the name comes from *catena*, which means 'chain'), suspended on two points (see figure 2.8.1). The interest in this curve originated with

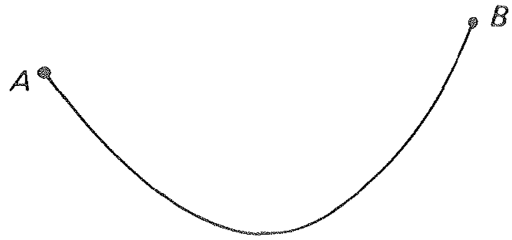


Figure 2.8.1.

Galileo, who thought that it was a parabola. Young Christiaan Huygens proved in 1646 that this cannot be the case. What the actual form was remained an open question till 1691, when Leibniz, Johann Bernoulli and the then much older Huygens sent solutions of the problem to the *Acta* (Jakob Bernoulli, *1690a*, Johann Bernoulli *1691b*, Huygens *1691a* and Leibniz *1691a*), in which the previous year Jakob Bernoulli had challenged mathematicians to solve it. As published, the solutions did not reveal the methods, but through later publications of manuscripts these methods have become known. Huygens applied with great virtuosity the by then classical methods of 17th-century infinitesimal mathematics, and he needed all his ingenuity to reach a satisfactory solution. Leibniz and Bernoulli, applying the new calculus, found the solutions in a much more direct way. In fact, the catenary was a test-case between the old and the new style in the study of curves, and only because the champion of the old style was a giant like Huygens, the test-case can formally be considered as ending in a draw.

A short summary of Johann Bernoulli's solution (he recapitulated it in his *1691a*, lectures 12 and 36), may provide an insight in how the

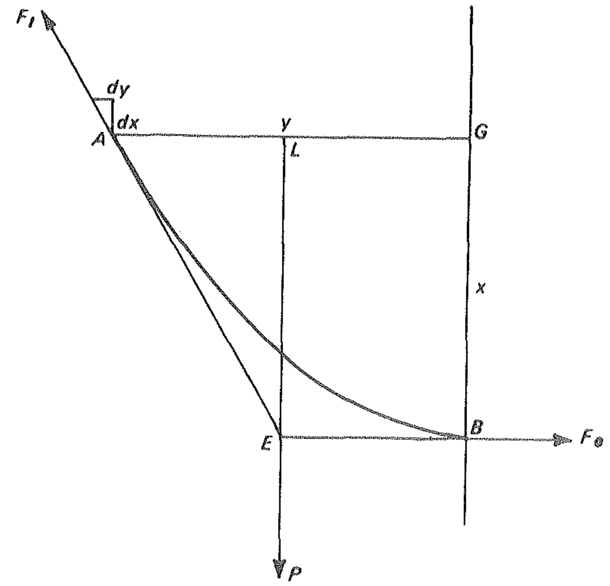


Figure 2.8.2.

new method was applied. In figure 2.8.2 let AB be part of the catenary. Using arguments from mechanics, he inferred that the forces F_0 and F_1, applicable in B and A to keep the part AB of the chain in position, are the same (in direction and quantity) as the forces required to keep the weight P of the chain AB in position, suspended as a mass at E on weightless cords AE and BE, which are tangent to the curve as in the figure. Moreover, the force F_0 at B does not depend on the choice of the position of A along the chain. P may be put equal to the length s of the chain from B to A; $F_0 = a$, a constant; and from composition of forces we have

$$P : F_0 = s : a = dx : dy. \qquad (2.8.1)$$

Hence

$$\frac{dy}{dx} = \frac{a}{s}. \qquad (2.8.2)$$

This is the differential equation of the curve, though in a rather intractable form as x and y occur implicitly in the arc-length s. Through skilful manipulation Bernoulli arrives at the equivalent differential equation

$$dy = \frac{a\,dx}{\sqrt{(x^2 - a^2)}}. \qquad (2.8.3)$$

I shall not follow his argument here in detail, but the equivalence can be seen by going backwards and calculating ds from (2.8.3):

$$ds = \sqrt{(dy^2 + dx^2)} = \sqrt{\left(\frac{a^2}{x^2 - a^2} + 1\right)}\,dx = \frac{x\,dx}{\sqrt{(x^2 - a^2)}}. \qquad (2.8.4)$$

Hence by integration

$$s=\int \frac{x\,dx}{\sqrt{(x^2-a^2)}}=\sqrt{(x^2-a^2)}=a\frac{dx}{dy}. \tag{2.8.5}$$

Through a substitution $x \to x+a$ Bernoulli reduces (2.8.3) to

$$dy=\frac{a\,dx}{\sqrt{(x^2+2ax)}}. \tag{2.8.6}$$

This substitution is needed to move the origin to B. In the differential equation (2.8.6) the variables are separated, so that the solution is

$$y=\int \frac{a\,dx}{\sqrt{(x^2+2ax)}}, \tag{2.8.7}$$

and the question is left to find out what the right hand side means. At that time, in the early 1690s, Bernoulli had not yet the analytical form of the logarithmic function at his disposal to express the integral as we would (namely, as $a \log (a+x+\sqrt{[x^2+2ax]})$). Instead he gave geometrical interpretations of the integral, namely, as quadratures of curves. He noted that the integral represents the area under the curve

$$z=\frac{a^2}{\sqrt{(x^2+2ax)}}. \tag{2.8.8}$$

But he also interpreted (through transformations which again we shall not present in detail) the integral as an area under a certain hyperbola and even as an arc-length of a parabola. By these last two interpretations, or ' constructions ' as this procedure of interpreting integrals was called, he proved that the form of the catenary ' depended on the quadrature of the hyperbola ' (we would say : involves only the transcendental function the logarithm) and with this proof the problem was, to the standards of the end of the 17th century, adequately solved.

The brachistochrone problem

If a body moves under influence of gravity, without friction or air resistance along a path γ (see figure 2.8.3), then it will take a certain time, say T_γ, to move to B starting from rest in A. T_γ depends on the form of γ. The *brachistochrone* (literally : shortest time) is the curve γ_0 from A to B for which T_γ is minimal. It can easily be seen that the fall along a straight line from A to B does not take the minimal time, so there is a problem : to determine the brachistochrone.

The problem was publicly proposed by Johann Bernoulli in the *Acta* of June 1696 (Bernoulli *1696a*) and later in a separate pamphlet. Several solutions reached the *Acta* and were published in May 1697 (Johann

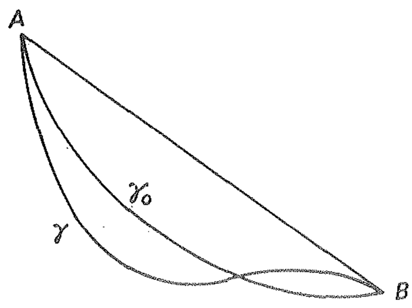


Figure 2.8.3.

Bernoulli *1697a*, l'Hôpital *1697a*, Leibniz *1697a* and Newton *1697a*; see Hofmann *1956a*, 35–36). Bernoulli's own solution used an analogy argument: he saw that the problem could be reduced to the problem of the refraction of a light-ray through a medium in which the density, and hence the refraction index, is a function of the height only. Leibniz and Jakob Bernoulli first considered the position of two consecutive straight line-segments (see figure 2.8.4) such that T_γ from P to Q is minimal. This is an extreme value problem depending on one variable and therefore solvable. Extending this to three consecutive straight segments and considering these as infinitely small, they arrived at a differential equation for the curve, which they solved. They found, as did Johann Bernoulli, that the brachistochrone is a *cycloid* (compare section 1.8) through A and B with vertical tangent at A. Newton had also reached this conclusion.

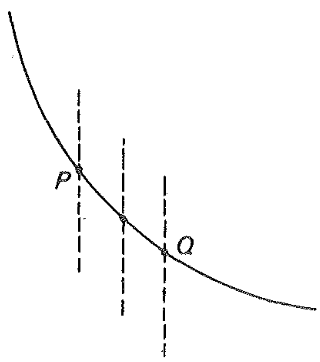


Figure 2.8.4.

The problem of the brachistochrone is very significant in the history of mathematics, as it is an instance of a problem belonging to the *calculus of variations*. It is an extreme value problem, but one in which the quantity (T_γ), whose extreme value is sought, does not depend on one or a finite number of independent variables but on the *form of a curve*.

Jakob Bernoulli proposed, as a sequel to his solution of the brachistochrone problem, further problems of this type, namely the so-called ***isoperimetric problems***. In the case of the brachistochrone, the class of curves considered consists of the curves passing through A and B. In isoperimetric problems one considers curves with prescribed length. For instance, it could be asked to find the curve through A and B with length l and comprising, together with the segment AB, the largest area (see figure 2.8.5). Jakob Bernoulli made much progress in finding methods to solve this type of problem. Euler unified and generalised these methods in his treatise *1744a*, thus shaping them into a separate branch of analysis. Lagrange contributed to the further development of the subject in his *1762a*, in which he introduced the concept of *variation* to which the subject owes its present name—the calculus of varia-t ons. On its history, see especially Woodhouse *1810a* and Todhunter *1861a*.

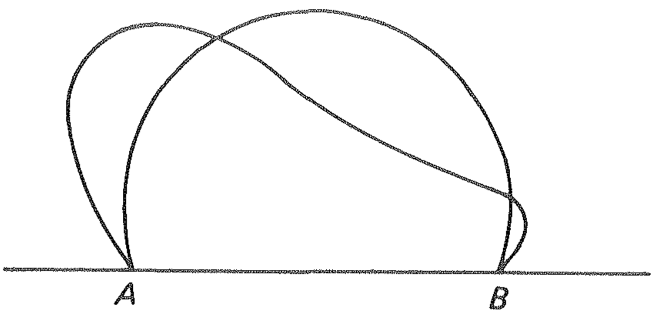


Figure 2.8.5.

2.9. *Rational mechanics*

The catenary and brachistochrone problems were two problems whose solution was made possible by the new methods. There were many more such problems, and their origins were diverse. The direct observation of simple mechanical processes suggested the problems of the form of an elastic beam under tension, the problem of the vibrating string (which Taylor, Daniel Bernoulli, d'Alembert, Euler and many others studied ; see section 3.3) and the problem of the form of a sail blown by the wind (discussed by the Bernoulli brothers in the early 1690s).

More technologically involved constructions suggested the study of pendulum motion (which Huygens initiated), the path of projectiles, and the flow of water through pipes. Astronomy and philosophy suggested the motion of heavenly bodies as a subject for mathematical treatment. Mathematics itself suggested problems too : special differential equations were generalised, types of integrals were classified (for example, elliptic integrals), and so on. Certain types of problems began rather quickly to form coherent fields with a unified mathematical approach: the calculus of variations, celestial mechanics, hydrodynamics, and mechanics in general. Somewhat later, probability (on which Jakob Bernoulli wrote a fundamental treatise *Ars conjectandi* (' The art of guessing '), which was published posthumously as *1713a*), joined this group of mathematicised sciences, or sub-fields of mathematics.

Something more should be said here about the new branches of mechanics (or ' rational mechanics ' as it was then called, to distinguish it from the study of machines), which acquired its now familiar mathematicised form in the 18th century. The basis for this mathematicisation was laid by Newton in his *Philosophiae naturalis principia mathematica* (*1687a*), in which he formulated the Newtonian laws of motion and showed that the supposition of a gravitational force inversely proportional to the square of the distance yields an appropriate description of the motion of planets as well as of the motion of falling and projected bodies here on earth. He gave here (among many other things) a full treatment of the motion of two bodies under influence of their mutual gravitational forces, several important results on the ' three-body problem ', and a theory of the motion of projectiles in a resisting medium. However, a great deal in the way of mathematicisation of these subjects still had to be done after the *Principia*. Though Newton made full use of his new infinitesimal methods in the *Principia*, he found and presented his results in a strongly geometrical style. Thus, although implicitly he set up and solved many differential equations, exactly or by approximation through series expansions, one rarely finds them written out in formulas in the *Principia*. Neither are his laws of motion expressed as fundamental differential equations to form the starting-point of studies in mechanics.

In the first half of the 18th century, through the efforts of men like Jakob, Johann and Daniel Bernoulli, d'Alembert, Clairaut and Euler, the style in this kind of study was further mathematicised—that is, the methods were transformed into the analytical methods—and they were unified through the formulation of basic laws expressed as mathematical formulas, differential equations in particular. Other fields were also tackled in this way, such as the mechanics of elastic bodies (on which Jakob Bernoulli published a fundamental article *1694a*) and hydrodynamics, on which father and son Johann and Daniel Bernoulli wrote early treatises (*1743a* and *1738a* respectively).

Great textbooks of analytic mechanics, such as Euler's *Mechanica* (*1736a*), d'Alembert's *1743a* and Lagrange's *1788a*, show a gradual process of mathematicisation of mechanics. Though Euler's *Mechanica* was strongly analytical, the formulation of Newton's laws in terms of differential equations (now termed 'Newton's equations') occurred for the first time only in a study of Euler published in 1752 (see Truesdell *1960a*). These branches of rational mechanics were very abstract fields in which highly simplified models of reality were studied. Therefore, the results were less often applicable than one might have hoped. These studies served to develop many new mathematical methods and theoretical frameworks for natural science which were to prove fruitful in a wider context only much later. Still, the interest in the problems treated was not entirely internally derived. Thus the projectiles of artillery suggested the study of motion in a resisting medium, while the three-body problem was studied by Newton, Euler and many others, especially in connection with the motion of the moon under the influence of the earth and the sun, a celestial phenomenon which was of the utmost importance for navigation as good moon tables would solve the problem of determining a ship's position at sea (the so-called 'longitudinal problem'). Indeed, Euler's theoretical studies of this problem, combined with the practical astronomical expertise of Johann Tobias Mayer, gave navigation, in the 1760s, the first moon tables accurate enough to yield a sufficiently reliable means for determining position at sea.

Central problems in hydrodynamics were the efflux of fluid from an opening in a vessel, and the problem of the shape of the earth. The latter problem was of philosophical as well as practical importance, because Cartesian philosophy predicted a form of the earth elongated along the axis, while Newtonian philosophy, considering the earth as a fluid mass under the influence of its own gravity and centrifugal forces through its rotation, concluded that the earth should be flattened at the poles. In practice, the deviation of the surface of the earth from the exact sphere form has to be known in order to calculate actual distances from astronomically determined geographical latitude and longitude. Several expeditions were held to measure one degree along a meridian in different parts of the earth, and the findings of these expeditions finally corroborated the Newtonian view.

2.10. *What was left unsolved : the foundational questions*

The problem that was left unsolved throughout the 18th century was that of the foundations of the calculus. That there *was* a problem was well-known, and that is hardly surprising when one considers how obviously self-contradictory properties were claimed for the fundamental concept of the calculus, the differential. According to l'Hôpital's first postulate, a differential can increase a quantity without increasing it. Nevertheless, this postulate is necessary for deriving the rules of the calculus, where higher-order differentials (or powers or products of differentials) have to be discarded with respect to ordinary differentials, and similarly ordinary differentials have to be discarded with respect to finite quantities (see (2.5.1)). Also, when Bernoulli takes the differential of the area $\mathcal{Q}$ to be equal to $y\,dx$ he discards the small triangle at the top of the strip (like MmR in figure 2.5.2) because it is infinitely small with respect to $y\,dx$. Thus the differentials have necessary but apparently self-contradictory properties. This leads to the foundational question of the calculus as many mathematicians since Leibniz saw it :

FQ 1 : *Do infinitely small quantities exist ?*

Most practitioners of the Leibnizian calculus convinced themselves in some way or other that the answer to FQ 1 is ' yes ', and thus they considered the rules of the calculus sufficiently proved. There is, however, a more sophisticated way of looking at the question, a way which for instance Leibniz himself adopted (see Bos *1974a*, 53–66). He had his doubts about the existence of infinitely small quantities, and he therefore tried to prove that by using the differentials as possibly meaningless symbols, and by applying the rules of the calculus, one would arrive at correct results. So his foundational question was :

FQ 2 : *Is the use of infinitely small quantities in the calculus reliable ?*
He did not obtain a satisfactory answer.

In Newton's fluxional calculus (see section 2.2) there also was a foundational problem. Newton claimed that his calculus was independent of infinitely small quantities. His fundamental concept was the *fluxion*, the velocity of change of a variable which may be considered to increase or decrease with time. In the actual use of the fluxional calculus, the fluxions themselves are not important (in fact they are undetermined), but their ratios are. Thus the tangent of a curve is found by the argument that the ratio of ordinate to sub-tangent is equal to the ratio of the fluxions of the ordinate and the abscissa respectively : $y/\sigma = \dot{y}/\dot{x}$ ($\dot{y}$ is the fluxion of y, $\dot{x}$ the fluxion of x ; see figure 2.10.1).

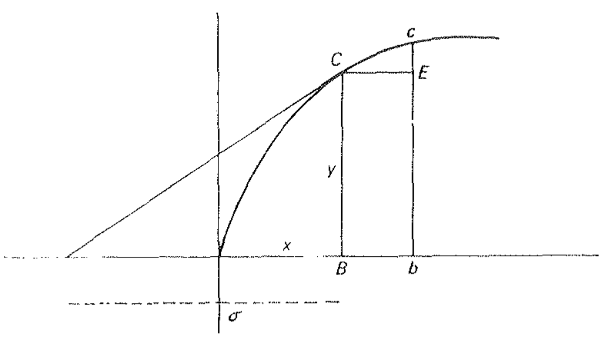


Figure 2.10.1.

He explains that the ratio of the fluxions $\dot{y}/\dot{x}$ is equal to the 'prime' or 'ultimate' ratio of the augments or decrements of y and x (see Newton *1693a* ; *Works*$_2$, vol. 1, 141). That is, he conceives corresponding increments Bb of x and Ec of y, and he considers the ratio Ec/CE for Ec and CE both decreasing towards 0 or both increasing from 0. In the first case he speaks of their *ultimate ratio* which they have just when they vanish into zero or nothingness ; in the latter case he speaks about their *prime ratio*, which they have when they come into being from zero or nothingness. The ratio $\dot{y}/\dot{x}$ is precisely equal to this ultimate ratio of evanescent augments, or equivalently to this prime ratio of 'nascent' augments.

Obviously there is a limit-concept implicit in this argument, but it is also clear that the formulation as it stands leaves room for doubt. For as long as the augments exist their ratio is not their ultimate ratio, and when they have ceased to exist they have no ratio. So here too is a foundational question, namely :

FQ 3 : *Do prime or ultimate ratios exist ?*

2.11. *Berkeley's fundamental critique of the calculus*

Most mathematicians who dealt with calculus techniques in the early 18th century did not worry overmuch about foundational questions. Indeed, it is significant that the first intensive discussion on the foundations of the calculus was not caused by difficulties encountered in working out or applying the new techniques, but by the critique of an outsider on the pretence of mathematicians that their science is based on secure foundations and therefore attains truth. The outsider was Bishop George Berkeley, the famous philosopher, and the target of his critique is made quite clear in the title of his tract *1734a* : 'The Analyst ; or a Discourse Addressed to an Infidel Mathematician Wherein It Is Examined Whether the Object, Principles, and Inferences of the Modern Analysis are More Distinctly Conceived, or More Evidently Deduced, than Religious Mysteries and Points of Faith '.

As we have seen, Berkeley indeed had a point. In sharp but captivating words he exposed the vagueness of infinitely small quantities, evanescent increments and their ratios, higher-order differentials and higher-order fluxions (*1734a*, para. 4) :

> Now as our Sense is strained and puzzled with the perception of Objects extremely minute, even so the Imagination, which Faculty derives from Sense, is very much strained and puzzled to frame clear Ideas of the least Particles of time, or the least Increments generated therein : and much more so to comprehend the Moments, or those

Increments of the flowing Quantities in *statu nascenti*, in their very first origin or beginning to exist, before they become finite Particles. And it seems still more difficult, to conceive the abstracted Velocities of such nascent imperfect Entities. But the Velocities of the Velocities, the second, third, fourth and fifth Velocities, &c. exceed, if I mistake not, all Humane Understanding. The further the Mind analyseth and pursueth these fugitive Ideas, the more it is lost and bewildered ; the Objects, at first fleeting and minute, soon vanishing out of sight. Certainly in any Sense a second or third Fluxion seems an obscure Mystery. The incipient Celerity of an incipient Celerity, the nascent Augment of a nascent Augment *i.e.* of a thing which hath no Magnitude : Take it in which light you please, the clear Conception of it will, if I mistake not, be found impossible, whether it be so or no I appeal to the trial of every thinking Reader. And if a second Fluxion be inconceivable, what are we to think of third, fourth, fifth Fluxions, and so onward without end ?

Further on comes the most famous quote from *The analyst* : 'And what are these Fluxions ? The Velocities of evanescent Increments ? And what are these same evanescent Increments ? They are neither finite Quantities, nor Quantities infinitely small nor yet nothing. May we not call them the Ghosts of departed Quantities ?' (para. 35). Berkeley also criticised the logical inconsistency of working with small increments which first are supposed unequal to zero in order to be able to divide by them, and finally are considered to be equal to zero in order to get rid of them.

Of course Berkeley knew that the calculus, notwithstanding the unclarities of its fundamental concepts, led, with great success, to correct conclusions. He explained this success—which led mathematicians to believe in the certainty of their science—by a *compensation of errors*, implicit in the application of the rules of the calculus. For instance, if one determines a tangent, one first supposes the characteristic triangle similar to the triangle of ordinate, sub-tangent and tangent, which involves an error because these triangles are only approximately similar. Subsequently one applies the rules of the calculus to find the ratio dy/dx, which again involves an error as the rules are derived by discarding higher-order differentials. These two errors compensate each other, and thus the mathematicians arrive 'though not at Science, yet at Truth, For Science it cannot be called, when you proceed blindfold, and arrive at the Truth not knowing how or by what means' (*1734a*, para. 22).

2.12. *Limits and other attempts to solve the foundational questions*

Berkeley's critique started a long-lasting debate on the foundations of the calculus. Before mentioning some arguments in this debate, it may be useful to recall how in modern differential calculus the foundational question is solved. Modern calculus concerns *functions* and relates to a function f its derivative f', which is again a function, defined by means of the concept of limit:

$$f'(x) \underset{\mathrm{Df}}{=} \lim_{h \to 0} \left(\frac{f(x+h)-f(x)}{h} \right). \qquad (2.12.1)$$

The preliminaries for this approach were worked out in the 18th and 19th centuries; they played different roles in the various approaches to the foundational questions which were adopted in that period. It is instructive to list the preliminaries. They are:

(1) the idea that the calculus concerns *functions* (rather than variables);

(2) the choice of the *derivative* as fundamental concept of the differential calculus (rather than the differential);

(3) the conception of the derivative as a function; and

(4) the concept of *limit*, in particular the limit of a function for explicitly indicated behaviour of the independent variable (thus explicitly $\lim_{h\to 0} (p(h))$, rather than merely the limit of the variable p).

Of the various approaches to the questions raised by Berkeley's critique, we have already seen the one adopted by Euler: he did conceive the calculus as concerning functions, but for him the prime concept was still the differential, which he considered as equal to zero but capable of having finite ratios to other differentials. Obviously this still leaves the foundational question QF 3 of section 2.10 unanswered. In fact, it does not seem that Euler was too much concerned about foundational questions.

Berkeley's idea of compensating errors was used by others to show that, rather than proceeding blindfold, the calculus precisely compensates equal errors and thus arrives at truth along a sure and well-balanced path. The idea was developed by Lazare Carnot among others. Another approach was due to Joseph Louis Lagrange, who supposed that for every function f and for every x one can expand $f(x+h)$ in a series

$$f(x+h) = f(x) + Ah + Bh^2 + Ch^3 + \dots . \qquad (2.12.2)$$

So Lagrange defined the 'derived function' $f'(x)$ as equal to the coefficient of h in this expansion. The idea, published first in *1772a*, became somewhat influential later through Lagrange's *Théorie des fonctions analytiques* (*Functions*). As a solution of the foundational questions the idea is unsound (not every $f(x+h)$ can be so expanded, and even so there would be the question of convergence), but in other ways this approach was quite fruitful ; it conceived the calculus as a theory about functions and their derived functions, which are themselves again functions. For more details on Carnot and Lagrange, see sections 3.3 and 3.4.

Eventually the most important approach towards solving the foundational questions was the use of limits. This was advocated with respect to the fluxional calculus by Benjamin Robins (see his *1761a*, vol. 2, 49), and with respect to the differential calculus by d'Alembert. Robins and d'Alembert considered limits of variables as the limiting value which these variables can approach as near as one wishes. Thus d'Alembert explains the concept in an article *1765a* on ' Limite ' in the *Encyclopédie* which he edited with D. Diderot : ' One magnitude is said to be the *limit* of another magnitude when the second may approach the first within any given magnitude, however small, though the first magnitude may never exceed the magnitude it approaches '.

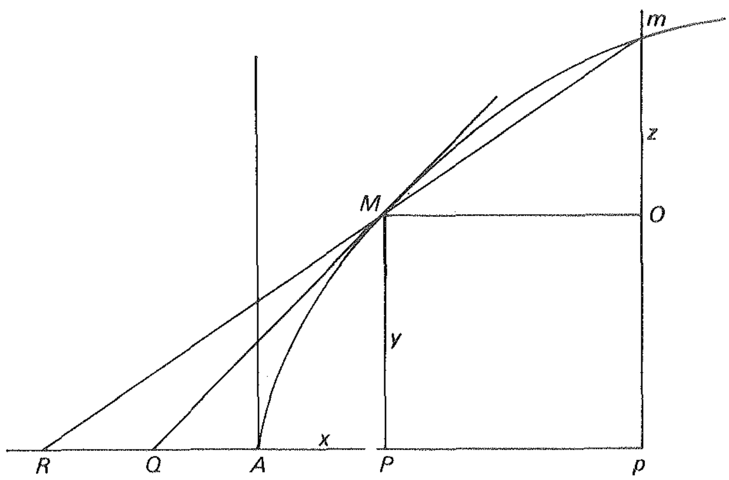


Figure 2.12.1.

In the *Encyclopédie* article ' Différentiel ' (*1764a*) d'Alembert gave a lengthy explanation, with the parabola $y^2=ax$ as example. His argument can be summarised as follows. From figure 2.12.1 it follows that MP/PQ is the limit of mO/OM. In formulae, $mO/OM=a/(2y+z)$, and algebraically the limit of $a/(2y+z)$ is easily seen to be $a/2y$. One variable can have only one limit, hence $MP/PQ=a/2y$. Furthermore, the rules of the calculus also give $dy/dx=a/2y$, so that we must conceive dy/dx not as a ratio of differentials or as $0/0$, but as the limit of the ratio of finite differences mO/OM.

Robins and d'Alembert were not the first to formulate the concept of limit ; in fact it occurs already implicitly in ancient Greek mathematics, and later Simon Stevin for instance came very close to formulating it (see his *Works*, vol. 1, 229–231). For a very long time after Robins and d'Alembert propagated the use of this concept to solve the foundational questions, the limit approach was just one among many approaches to the problem. The reason why it took so long until the value of the limit approach was recognised lay in the fact that Robins and d'Alembert considered limits of *variables*. In that way the concept still involves much unclarity (for details, see Baron and Bos *1976a*, unit 4) which could only be removed once the limit concept was applied to *functions* under explicitly specified behaviour of the independent variable.

2.13. *In conclusion*

In the century which followed Newton's and Leibniz's independent discoveries of the calculus, analysis developed in a most impressive way, despite its rather insecure foundations, thus making possible a mathematical treatment of large parts of natural science. During these developments analysis also underwent deep changes ; for Newton and Leibniz did not invent the modern calculus, nor did they invent the same calculus. It will be useful to recall, in conclusion, the main features of both systems, their mutual differences, and their differences from the forms of calculus to which we are now used (compare Baron and Bos *1976a*, unit 3, 55–57).

Both Newton's and Leibniz's calculi were concerned with *variable quantities*. However, Newton conceived these quantities as *changing in time*, whereas Leibniz rather saw them as *ranging over a sequence of infinitely close values*. This yielded a difference in the fundamental concepts of the two calculi ; Newton's fundamental concept was the *fluxion*, the finite velocity or rate of change (with respect to time) of the variable, while Leibniz's fundamental concept was the *differential*, the infinitely small difference between successive values in the sequence.

There was also a difference between the two calculi in the conception of the *integral*, and in the role of the *fundamental theorem*. For Newton integration was *finding the fluent quantity of a given fluxion* ; in his calculus, therefore, the fundamental theorem was implied in the definition of integration. Leibniz saw integration as *summation* ; hence for him the fundamental theorem was not implied in the definition of integration, but was a consequence of the inverse relationship between summing and taking differences. However, the Bernoullis re-interpreted the Leibnizian integral as the converse of differentiation, so that throughout the 18th century the fundamental theorem was implied in the definition of integration.

Both Newton and Leibniz worked with *infinitely small quantities* and were aware of the logical difficulties inherent in their use. Newton claimed that his calculus could be given a rigorous foundation by means of the concept of *prime and ultimate ratio*, a concept akin to (but certainly not the same as) the concept of limit.

Leibniz valued *notation* very much, and his choice of symbols for the calculus proved to be a happier one than Newton's. His use of separate letters, 'd' and '$\int$', indicated the role of differentiation and integration as operators ; moreover, his symbols were incorporated into complicated formulas much more easily than were Newton's. In general, Leibniz's calculus was the more analytical ; Newton's was nearer to the geometrical figures, with accompanying arguments in prose.

These are the principal differences between the two systems. If we compare them with the modern calculus, we note three further differences. Firstly, whereas Newton's and Leibniz's calculi were concerned with *variables*, the modern calculus deals with *functions*. Secondly, the operation of differentiation is defined in the modern calculus differently from in the 18th century ; it relates to a function a derived function, or derivative, defined by means of the concept of limit. Thirdly, unlike 18th-century calculus, modern analysis has a generally accepted approach to the problem of the foundation of the calculus namely, through a definition of real numbers (instead of the vague concept of quantity which had to serve as a basis for analysis before the 1870s) and through the use of a well-defined concept of limit. The next chapter describes much of this future progress.

Text 16: N. Guicciardini (2003). "Newton's Method and Leibniz's Calculus". In: *A History of Analysis*. Ed. by H. N. Jahnke. History of Mathematics 24. Providence (Rhode Island): American Mathematical Society. Chap. 3, pp. 73–103.

CHAPTER 3

Newton's Method and Leibniz's Calculus

Niccolò Guicciardini

3.1. Introduction

From the 1660s to the 1680s, Isaac Newton and Gottfried Wilhelm Leibniz created what we nowadays recognize as infinitesimal calculus. A study of their achievements reveals elements of continuity with previous work (see Chapter 2) as well as peculiarities which distinguish their methods and concepts from those which are accepted in present day mathematics. The statement itself that "Newton and Leibniz invented the calculus" is problematic. In the first place, they developed two different versions of calculus, and the problem of comparing the two, of establishing equivalences and differences, arises (see Chapter 3.5). In the second place, what do we mean by "inventing calculus" in this context?

The novelty of Newton's and Leibniz's contributions can be briefly characterized by pointing out three aspects of their mathematical work: problem-reduction, the calculation of areas by inversion of the process for calculating tangents, the creation of an algorithm. The "invention of calculus" can thus be conceived as consisting of these three contributions.

Newton and Leibniz realized that a whole variety of problems about the calculation of centres of gravity, areas, volumes, tangents, arclengths, radii of curvature, surfaces, etc., that had occupied mathematicians in the first half of the seventeenth century, were instances of two basic problems. Furthermore, they fully realized that these two problems were the inverse of each other (this is the "fundamental theorem" of calculus). They thus understood that the solution of the former, and easier, problem could be used to answer the latter. Last but not least, Newton and Leibniz developed two efficient algorithms that can be applied in a systematic and general way. It is thanks to these contributions that Newton and Leibniz transformed mathematics.

The peculiarity of Newton's and Leibniz's algorithms is a fact that the historian is sometimes led to forget. In fact, both, especially the latter, look very much the same as the one we employ nowadays. We can thus be tempted to modernize their calculi. As a matter of fact, their calculi are strongly embedded in the culture of their own times. We make two major points. Neither Newton's nor Leibniz's calculi are about "functions" (see (Bos 1980, 90).) The concept of function emerged only later (see Chapter 4). Newton and Leibniz talk in terms of "quantities" rather than "functions", and they refer to these quantities, their rates of change, their differences, etc., related to specific geometric entities (typically a given curve). Thus the reader will notice that in what follows I will always use the term "function" in "quotation marks". Furthermore, while we are used to referring to calculus as the continuum of the real numbers, the continuum to which Newton and Leibniz refer is geometrical or kinematical. It is by referring to an intuitive geometric or kinematic continuum that Newton and Leibniz develop their limit procedures (see 3.5.2).

3.2. Newton's method of series and fluxions

3.2.1. A mathematician working in isolation. Isaac Newton was born into a family of small landowners. After receiving an elementary education, he was sent to Cambridge, where he matriculated as a sub-sizar in 1661. "Sub-sizars" were poor students who worked as servants to the fellows and the rich students. Newton raised himself from this condition to become Lucasian Professor, Warden of the Mint, a member of Parliament and President of the Royal Society. His funeral was described by Voltaire as being as full of pomp as those of a king. His success in British society was determined by the high esteem which his published scientific discoveries aroused. In his secret, unpublished, studies Newton cultivated interests that would have ruined his public image. He was involved in alchemical studies, and his theological interests, inspired by deep religious feelings, gave him strongly critical attitude towards the established Church.

Some of Newton's greatest scientific discoveries were made during the years 1665–1667, when Cambridge university was closed because of the plague. During these *anni mirabiles* Newton performed experiments with prisms, convincing himself of the composite nature of white light, stated the binomial theorem for fractional powers, discovered the calculus of fluxions and speculated about the moon's motion. For complicated reasons, he did not immediately share his mathematical results with others. This is only explained in part by the cost of mathematical publications at that time. More decisive was his introverted character that led him to keep his thoughts to himself. Furthermore, he was not completely confident about the conceptual foundation of his calculus. To these causes which may have hindered Newton from publishing his discoveries on calculus, one can add that it was a practice of some seventeenth century mathematicians to keep their mathematical methods secret. The mathematical tools, which allowed the solution of problems, were considered private property, not to be shared too generously with others. Very much as painters kept the secrets for obtaining colours for themselves, the mathematicians often gave the solution without revealing the demonstration. In 1676 the secretary of the Royal Society, Henry Oldenburg, obtained from Newton two letters in which some of his mathematical results were summarized. These two letters were meant to inform a German philosopher, Gottfried Wilhelm Leibniz, about the scope of Newton's achievements. The *Philosophiae Naturalis Principia Mathematica* (1687), where Newton developed his theory of gravitation, also contained results connected with calculus. It was only in 1704 that Newton published a systematic treatise on calculus: the *De quadratura curvarum.* This was too late to prevent a priority dispute with Leibniz, who had already published his differential calculus in 1684. Leibniz was accused of plagiarism by Newton and by the British fellows of the Royal Society. Actually he had discovered differential and integral calculus in 1672–1676 independently. He therefore asked the Royal Society to withdraw the accusation of plagiarism that was circulating in several papers. A committee of the Royal Society, guided secretly by Newton, reported that Leibniz was guilty of plagiarism. The Newtonian and the Leibnizian schools differed strongly on a wide range of issues. They maintained different cosmologies, different views on the relationships between God and nature, different views on space and time and on the conservation laws basic in physics. The priority dispute divided them mathematically. This was a bitter outcome for Leibniz, who had always maintained that the demonstrative power of mathematics could end all disputes and promote a more harmonious world.

3.2.2. The binomial series (1664 to 1665). It appears that Newton's interest in mathematics began in 1664, when he read François Viète's works (1646), Descartes's *Géométrie* (1637) (the second Latin edition (1659–1661) with Frans van Schooten's commentaries and Hudde's rule), William Oughtred's *Clavis mathematicae* (1631), and Wallis's *Arithmetica Infinitorum* (1656). It was from reading this selected group of mathematical works in "modern analysis" that Newton learned about the most exciting discoveries on analytic geometry, algebra, tangent problems, quadratures and series. After a few months of self-instruction he was able, in the winter 1664–1665, to make his first mathematical discovery: the "binomial theorem" for fractional powers. In slightly modernized notation, he stated:

$$(a+x)^{m/n} = a^{m/n} + \frac{m}{n}a^{m/n-1}x + \frac{1}{1.2}\frac{m}{n}(\frac{m}{n}-1)a^{m/n-2}x^2 + \cdots . \tag{3.1}$$

Newton obtained this result generalising by Wallis's "inductive" method for squaring the unit circle. The process of interpolation with which Newton determined the binomial coefficients is too long to be described in detail here. A good presentation of Newton's guesswork can be found in (Edwards 1979, 178-187). Here it will suffice to say that Newton arrived at

$$1x - \frac{1}{2}\frac{1}{3}x^3 - \frac{1}{8}\frac{1}{5}x^5 - \frac{1}{16}\frac{1}{7}x^7 - \frac{5}{128}\frac{1}{9}x^9 \cdots . \tag{3.2}$$

as a series for the area under the curve $(1-x^2)^{1/2}$, a result which allows one to calculate the circle's area. He further noted that, since the area under x^n and over the interval $[0,x]$ is $x^{n+1}/(n+1)$, he could extend the result valid for the area to the curve itself to obtain

$$(1-x^2)^{1/2} = 1 - \frac{1}{2}x^2 - \frac{1}{8}x^4 - \frac{1}{16}x^6 - \frac{5}{128}x^8 \cdots . \tag{3.3}$$

By working through similar examples, Newton guessed the general law of formation of the binomial coefficients for fractional powers (see (3.1)). He further extrapolated (3.1) to negative powers. The case $n = -1$,

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - \cdots . \tag{3.4}$$

is particularly relevant. Since the proof of the binomial series rested on shaky "inductive" Wallisian procedures, Newton felt the need to verify the agreement of the series obtained by applying (3.1) by algebraical and numerical procedures. For instance, he applied standard techniques of root extraction to $(1-x^2)^{1/2}$ and standard techniques of "long division" to $(1+x)^{-1}$, and he was happy to see that he obtained the series (3.3) and (3.4).

He also knew that the area under $(1+x)^{-1}$ and over the interval $[0,x]$, or the negative of this area if $-1 < x < 0$, is $\ln(1+x)$. He could thus express $\ln(1+x)$ as a power series by termwise integration of (3.4):

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots . \tag{3.5}$$

Actually the order of Newton's reasoning is quite unexpected: He first obtained (3.5) via interpolation, and then he obtained (3.4) by differentiation. The series (3.5) allowed Newton to calculate $\ln(1+x)$, for $x \approx 0$. He carried out his numerical calculations up to more than fifty decimal places!

We note three aspects of Newton's work on the binomial series. First of all he introduced, following Wallis's suggestion, negative and fractional exponents. Without this innovative notation ($x^{a/b}$ for $\sqrt[b]{x^a}$) it would not have been possible to interpolate or extrapolate the binomial theorem from positive integers to the rationals. Secondly, Newton obtained a method for representing a large class of "curves" by a power series. For him curves are thus given not only by finite algebraical equations (as for Descartes) but also by infinite series (preferably power series) understood by Newton and by his contemporaries as *infinite equations.* In 1665 mathematicians had just begun to appreciate the usefulness of infinite series as representations of "difficult" curves. Transcendental curves, such as the logarithmic curve, can thus be given an "analytical" representation to which the rules of algebra can be applied. Before the advent of infinite series, such "functions" had no analytic representation, but they were generally defined in geometric terms. It should be noted that Newton had a rather intuitive concept of convergence. For instance he realized that the binomial series (3.1) can be applied when x is "small". Newton developed no rigorous treatment of convergence.

3.2.3. The fundamental theorem, 1665 to 1669. Newton's first systematic mathematical tract bears the title *De analysi per aequationes numero terminorum infinitas.* Newton began this short summary of his discoveries with the enunciation of three rules that can be rendered as follows (Newton 1669, 206 ff.):

Rule 1: If $y = ax^{m/n}$, then the area under y is $(an)/(n+m)x^{m/n+1}$.

Rule 2: If y is given by the sum of more terms (also an infinite number of terms), $y = y_1 + y_2 + \cdots$, then the area under y is given by the sum of the areas of the corresponding terms.

Rule 3: In order to calculate the area under a curve $f(x,y) = 0$, one must expand y as a sum of terms of the form $ax^{m/n}$ and apply Rule 1 and Rule 2. Rule 1 had been stated by Wallis. As we will see, Newton provided a proof of this rule based on the fundamental theorem. The binomial series proved to be an important tool implementing Rule 3. In several cases, however, the binomial series cannot be applied. In the years from 1669 to 1671 Newton devised several clever techniques for obtaining a series $z = \sum b_i x^i$, i rational, from an implicit "function" $f(x,z) = 0$. He also had a method for "reverting" series. That is, given $z = \sum b_i x^i$, he had a method of successive approximations which led to $x = \sum a_i z^i$. It is reverting the power series expansion of $z = \ln(1+x)$ (formula (3.5)) that he obtained the series for $x = e^z$ (see (Edwards 1979, 204–205) and Chapter 4).

The most general result concerning the squaring of curves (i.e., "integration") is the fundamental theorem of calculus which Newton discovered in 1665. Newton's reasoning, which resembles Barrow's (see 2.2.4), refers to two particular curves (see Fig. 3.1), $z = x^3/a$ and $y = 3x^2/a$, but it is completely general: y is equal to the slope of z and is defined as

$$bg = dh\frac{m\beta}{\Omega\beta}, \tag{3.6}$$

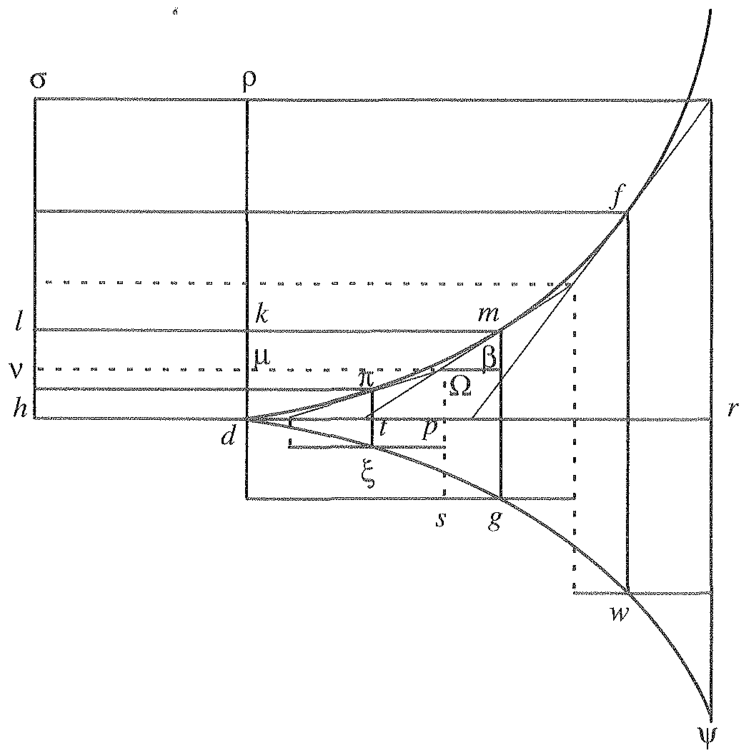


FIGURE 3.1

where bg is an ordinate of the curve y, and $m\beta$ and $\Omega\beta$ are infinitesimal increments of z and x, while dh is a unit length segment. It follows immediately that the area $bpsg$ ($= \Omega\beta \cdot bg$) and the area $\mu\kappa\lambda\nu$ ($= m\beta \cdot dh$) are equal. It was commonplace in seventeenth century mathematics to consider the area subtended by a curve to be equal to the sum of infinitely many infinitesimal strips such as $bpsg$. It follows that the curvilinear area subtended by y, e.g., $d\psi n$, is equal to the rectangular area $dh\sigma\rho$. A knowledge of z then allows us to "square" y, since "the area under y (the derivative curve) is proportional to the difference between corresponding ordinates of z" (Westfall 1980, 127). In Leibnizian terms, Newton proves that the integral of the derivative of z is equal to z (see (Newton 1665)).

A proof of the fact that the derivative of the integral of y is equal to y was given by Newton at the end of *De analysi* as a proof of Rule 1. He proceeded as follows.

Newton considered a curve $AD\delta$ (see Fig. 3.2), where $AB = x$, $BD = y$ and the area $ABD = z$. He defined $B\beta = o$ and $BK = v$ such that "the rectangle $B\beta HK$ ($= ov$) is equal to the space $B\beta\delta D$." Furthermore, Newton assumed that $B\beta$ is "infinitely small." With these definitions one has that $A\beta = x + o$ and the area $A\delta\beta$ is equal to $z + ov$. At this point Newton wrote: "from any arbitrarily assumed relationship between x and z I seek y." He noted that the increment of the area ov, divided by the increment of the abscissa o is equal to v. But since one can assume "$B\beta$ to be infinitely small, that is, o to be zero, v and y will be equal." Therefore, the rate of increase of the area is equal to the ordinate (Newton 1669, 242–244).

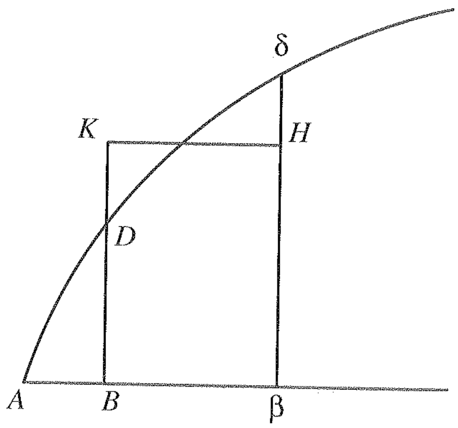


FIGURE 3.2

The fundamental theorem allowed Newton to reduce the probems of quadrature to the search for primitive "functions". He actually calculated the tangent for a great variety of "curves", so compiling what he called "tables of fluents" (in Leibnizian terms "table of integrals"). We will see in the next section how he deployed the fundamental theorem in order to square curves.

3.2.4. The method of fluents, fluxions and moments (1670 to 1671). While the *De analysi* was devoted mainly to series expansions and the use of series in quadratures, the *De methodis serierum et fluxionum* written in 1670–1671 was mainly devoted to the use of an algorithm that Newton had developed in the years from 1665 to 1666. The objects to which this algorithm is applied are quantities which "flow" in time. For instance the motion of a point generates a line and the motion of a line generates a surface. The quantities generated by a "flow" are called "fluents". Their instantaneous speeds are called "fluxions". The "moments" of the fluent quantities are "the infinitely small additions by which those quantities increase during each infinitely small interval of time" (Newton 1670–1671, 80). Consider a point which flows with variable speed along a straight line. The distance covered at time t is the fluent, the instantaneous speed is the fluxion, and the "infinitely" (or "indefinitely") small increment acquired after an indefinitely small period of time is the moment. Newton further observed that the moments "are as their speeds of flow", i.e., as the fluxions) (Newton 1670–1671, 78). His reasoning is based on the idea that during an "infinitely small period of time" the fluxion remains constant and so the moment is proportional to the fluxion. Newton warns the reader not to identify the "time" of the fluxional method with real time. Any fluent quantity whose fluxion is assumed constant plays the role of fluxional "time".

Newton did not develop a particularly handy notation in this context. He employed a, b, c, d for constants, v, x, y, z for the fluents and l, m, n, r for the respective fluxions, so that, e.g., m is the fluxion of x. The "indefinitely" (or "infinitely") small interval of time was denoted by o. Thus the moment of y is no. It was only in the 1690s that Newton introduced the now standard notation where the fluxion of x is denoted by $\dot{x}$ and the moment of x by $\dot{x}o$. The fluxions themselves can be considered as fluent quantities so that one can seek for the fluxion of n/m. In the 1690s Newton denoted the "second" fluxion of x by $\ddot{x}$.

Text 16: N. Guicciardini (2003). "Newton's Method and Leibniz's Calculus". In: *A History of Analysis*. Ed. by H. N. Jahnke. History of Mathematics 24. Providence (Rhode Island): American Mathematical Society. Chap. 3, pp. 73–103.

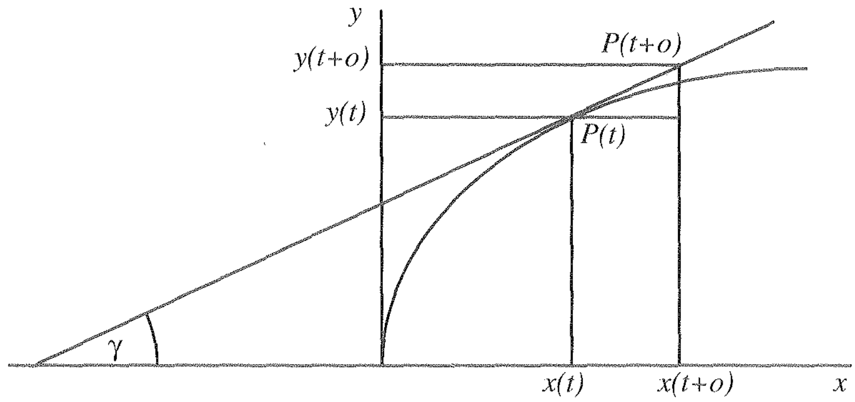


FIGURE 3.3

Newton did not use a single notation for the area under a curve. Generally he put words such as "the area of" or a capital Q before the analytical expression of the curve. In some cases he used "$\boxed{a/x^2}$" for "the area under the curve of equation $y = a/x^2$" (in Leibnizian terms this would be $\int(a/x^2)dx$). As we will see (3.2.6) Newton also employed $\acute{x}$ to denote a fluent quantity whose fluxion is x. The limits of integration were either understood by the context or explained by words.

In the *De methodis* Newton gives the solution of a series of problems. The main problems are to find maxima and minima, tangents, curvatures, areas and arclengths. The representation of quantities as generated by continuous flow allows all these problems to be reduced to the following Problems 1 and 2:

> 1) Given the length of the space continuously (that is, at every time), find the speed of motion at any time proposed.
> 2) Given the speed of motion continuously, find the length of the space described at any time proposed.
> (Newton 1670–1671, 70–71)

The problems of finding tangents, extremal points and curvatures are related to the former, and the problems of finding areas and arclengths are related to the latter.

Imagine a plane curve $f(x, y) = 0$ to be generated by the continuous flow of a point $P(t)$. If (x, y) are the Cartesian coordinates of the curve, $\dot{y}/\dot{x}$ will be equal to $\tan\gamma$, where γ is the angle formed by the tangent in $P(t)$ with the x-axis (see Fig. 3.3). According to Newton's conception, the point will move during the "indefinitely small period of time" with uniform rectilinear motion from $P(t)$ to $P(t + o)$. The infinitesimal triangle indicated in Fig. 3.3 has sides equal to $\dot{y}o$ and $\dot{x}o$ and so $\tan\gamma = \dot{y}o/\dot{x}o = \dot{y}/\dot{x}$. An extremal point will have $\dot{y}/\dot{x} = \tan\gamma = 0$. Newton showed that the radius of curvature is given by $\rho = (1 + (\dot{y}/\dot{x})^2)^{3/2}/(\ddot{y}/\dot{x}^2)$.

The fact that the finding of areas can be reduced to Problem 2 is a consequence of the fundamental theorem. Let z be the area generated by continuous uniform flow ($\dot{x} = 1$) of ordinate y (see Fig. 3.2). The speed of motion is given continuously, i.e., it is given by $\dot{z}$. By the fundamental theorem $y = \dot{z}$. In order to find the area, a method is required for obtaining z from $y = \dot{z}$. This is Problem 2. It should be stressed how the conception of quantities as generated by continuous flow allowed to Newton to conceive the problem of determining the area under a curve as a special example of Problem 2. The reduction of arclength problems to Problem 2 depends on the application of Pythagoras's theorem to the moment of arclength s: $\dot{s}o = \sqrt{(\dot{x}o)^2 + (\dot{y}o)^2}$ (see Fig. 3.3). Therefore $s = \boxed{\sqrt{\dot{x}^2 + \dot{y}^2}}$.

The basic algorithm for Problem 1 is given by Newton with an example (Newton 1670–1671, 78–81). He considered the equation $x^3 - ax^2 + axy - y^3 = 0$. He substituted $x + \dot{x}o$ in place of x and $y + \dot{y}o$ in place of y. Deleting $x^3 - ax^2 + axy - y^3$ as equal to zero and then dividing by o, he obtained an equation from which he cancelled the terms which had o as a factor. These terms have the property that they "will be equivalent to nothing in respect to the others", since "o is supposed to be infinitely small." At last Newton arrived at

$$3\dot{x}x^2 - 2a\dot{x}x + a\dot{x}y + a\dot{y}x - 3\dot{y}y^2 = 0. \tag{3.7}$$

This result is achieved by employing a rule of cancellation of higher-order infinitesimals (equivalent to Leibniz's $x + dx = x$), according to which, if x is finite and o is an infinitesimal interval of time, then

$$x + \dot{x}o = x. \tag{3.8}$$

Notice that the above example also contains the rules for the fluxions of a product xy and of x^n, respectively: $x\dot{y} + y\dot{x}$ and $nx^{n-1}\dot{x}$.

Newton dealt with irrational "functions" as follows. He considered $y^2 - a^2 - x\sqrt{a^2 - x^2} = 0$. He set $z = x\sqrt{a^2 - x^2}$ and so obtained $y^2 - a^2 - z = 0$ and $a^2x^2 - x^4 - z^2 = 0$. Applying the direct algorithm, he determined $2\dot{y}y - \dot{z} = 0$ and $2a^2\dot{x}x - 4\dot{x}x^3 - 2\dot{z}z = 0$. He then eliminated $\dot{z}$, restored $z = x\sqrt{a^2 - x^2}$, and thus arrived at

$$2\dot{y}y + (-a^2\dot{x} + 2\dot{x}x^2)/\sqrt{a^2 - x^2} = 0$$

as the relation sought between $\dot{y}$ and $\dot{x}$.

Even though Newton presents his "direct" algorithm by applying it to particular cases, his procedure can be generalized. Given a curve expressed by a function in parametric form, $f(x(t), y(t)) = 0$, the relation between the fluxions $\dot{x}$ and $\dot{y}$ is obtained by application of the equation

$$f(x + \dot{x}o, y + \dot{y}o) = \frac{\partial f}{\partial x}\dot{x}o + \frac{\partial f}{\partial y}\dot{y}o + o^2(\cdots) = 0.$$

After division by o, the remaining terms in o are cancelled. Such a modern reconstruction clearly says more than what Newton could express. I used concepts and notation, not available to Newton, for a function $f(x(t), y(t))$ and for partial derivatives. However, with due caution, it can be used to highlight the following points.

1) Newton assumes that, during the infinitesimal interval of time o, the motion is uniform, so that when x flows to $x + \dot{x}o$, y flows to $y + \dot{y}o$. Therefore, $f(x, y) = f(x + \dot{x}o, y + \dot{y}o)$.

2) Newton applies the principle of cancellation of infinitesimals, so in the last step the terms in o are dropped.

Newton's justification for his algorithmic procedure is not much more rigorous than those in the works of Pierre Fermat or Hudde. As we will see in the next subsection, he was soon to face serious foundational questions.

Problem 2 is, of course, much more difficult. Given a "fluxional equation" $f(x, y, \dot{x}, \dot{y}) = 0$, Newton seeks a relation $g(x, y, c) = 0$ (c constant) such that the

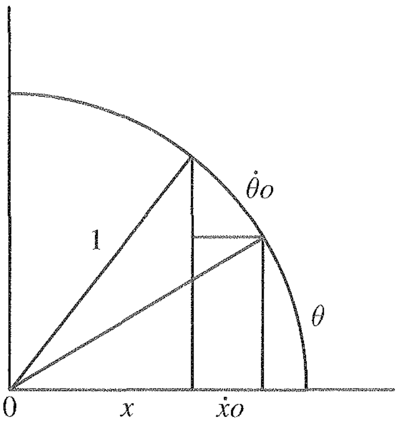


FIGURE 3.4

application of the direct algorithm yields $f(x, y, \dot{x}, \dot{y}) = 0$. In Leibnizian terms, he poses the problem of integrating differential equations.

Newton has a very general strategy which allows him to solve a great variety of such "inverse problems". His strategy is twofold. 1) Either he changes variable in order to reduce to a known table of fluents (in Leibnizian terms, a "table of integrals") or 2) he deploys series expansion techniques (termwise integration). His strategy is a great improvement on the geometrical quadrature techniques of, e.g., Huygens, or the techniques of direct summation of, e.g., Wallis (see Chapter 2).

We can give some of the flavour of Newton's first strategy by looking again at the quadrature of the cissoid which had occupied Huygens and Wallis in the late 1650s (see Chapter 2 and (van Maanen 1991)). Newton used $y = x^2/\sqrt{ax - x^2}$ as the equation for the cissoid (see Fig. 2.21). Problem 2 is solved by the determination of a z such that $\dot{z}/\dot{x} = x^2/\sqrt{ax - x^2}$. For $k = x^{3/2}\sqrt{a - x}$,

$$\frac{\dot{k}}{\dot{x}} = \frac{3}{2}\sqrt{ax - x^2} - \frac{1}{2}\frac{x^2}{\sqrt{ax - x^2}}. \tag{3.9}$$

Rearranging, we get

$$\frac{\dot{z}}{\dot{x}} = 3\sqrt{ax - x^2} - 2\dot{k}/\dot{x}. \tag{3.10}$$

In Leibnizian terms, $z = \int_0^a 3\sqrt{ax - x^2}dx - 2[k(x)]_0^a$. The area under the cissoid and over the interval $[0, a]$ is therefore three times the area under the semicircle with equation $y = \sqrt{ax - x^2}$. Notice that the second term on the right of (3.10) vanishes when "integrated" over $[0, a]$.

When the first strategy failed, Newton tried the second. He generally reduced the quadrature to the area under the graph of a circular or a hyperbolic "function", such as $(a^2 - x^2)^{\pm 1/2}$ or $a/(b + cx)$. These he could evaluate by binomial expansion and termwise "integration". An example follows.

Consider a circle with unit length radius (see Fig. 3.4): The moment of the arc $\dot{\theta}o$ is to the moment of the abscissa $\dot{x}o$ as 1 to $\sqrt{1 - x^2}$. Applying the binomial theorem to $(1 - x^2)^{-1/2}$ and "integrating" termwise, Newton obtained the arcsin series

$$\theta = x + \frac{x^3}{6} + \frac{3x^5}{40} + \cdots. \tag{3.11}$$

"Reverting" the above series by a process of successive approximations, he obtained the power series for sin.

Newton was able to solve the inverse problem for a large class of fluxional equations. Had he published his tract in 1671, he would have aroused awe in all the corners of Europe.

3.2.5. The geometry of prime and ultimate ratios (1671 to 1704). As we have seen, Newton employed methods characteristic of the seventeenth-century "new analysis" in his early writings. He used series and infinitesimal quantities. Infinitesimals entered mainly as moments, momentaneous increments of a "flowing" variable quantity. The kinematical approach to the calculus was therefore prevalent in Newton's work from the very beginning. For him, reference to our intuition of continuous "flow" provided a means to "define" the reference objects of the calculus: fluents, fluxions and moments (see 3.5.2).

Up to the composition of the *De methodis*, Newton described himself with pride as a promoter of the seventeenth-century "new analysis". However, in the 1670s he abandoned the calculus of fluxions in favour of a geometry of fluxions where infinitesimal quantities were not employed. He labelled this new method the "synthetical method of fluxions" as opposed to his earlier "analytical method of fluxions" (Newton 1967–1981, 8, 454–455). Some of the results on the synthetical method were summarized in Section 1, Book 1 of *Principia Mathematica* entitled "The method of prime and ultimate ratios". He wrote:

> whenever in what follows I consider quantities as consisting of particles or whenever I use curved line-elements [or minute curved lines] in place of straight lines, I wish it always to be understood that I have in mind not indivisibles but evanescent divisibles, and not sums and ratios of definite parts but the limits of such sums and ratios, and that the force of such proofs always rests on the method of the preceding lemmas. (Newton 1687/1999, 441–442)

He also pointed out that the method of prime and ultimate ratios rested on the following Lemma 1:

> Quantities, and also ratios of quantities, which in any finite time constantly tend to equality, and which before the end of that time approach so close to one another that their difference is less than any given quantity, become ultimately equal. (Newton 1687/1999, 433)

Newton's *ad absurdum* proof runs as follows:

> If you deny this, let them become ultimately unequal, and let their ultimate difference be D. Then they cannot approach so close to equality that their difference is less than the given difference D, contrary to the hypothesis. (Newton 1687/1999, 433)

This principle might be regarded as an anticipation of Cauchy's theory of limits (see Chapter 6), but this would certainly be a mistake, since Newton's theory of limits is referred to as a geometrical rather than a numerical model.

The objects to which Newton applies his "synthetical method of fluxions" or "method of prime and ultimate ratios" are geometrical quantities generated by continuous flow (i.e., "fluents"). While in his early writings Newton represented

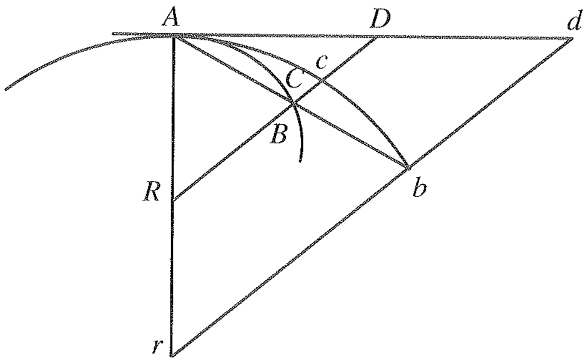


FIGURE 3.5

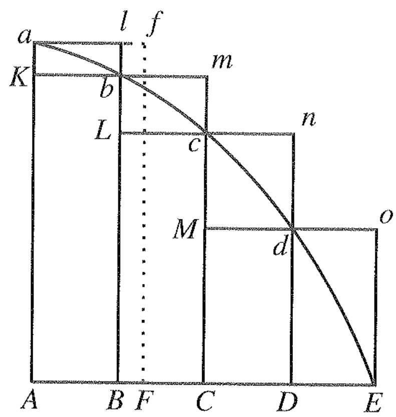


FIGURE 3.6

the fluents with algebraical symbols, in this new approach he referred directly to geometrical figures. These figures, however, are not static, as in classic geometry: they must be conceived as "in motion".

A typical problem is the study of the limit to which the ratio of two geometrical fluents tends when they vanish simultaneously (Newton used the expression of the "limit of the ratio of two vanishing quantities"). For instance, in Lemma 7 Newton shows that given a curveACB (see Fig. 3.5):

> the ultimate ratio of the arc, the chord, and the tangent to one another is a ratio of equality. (Newton 1687/1999, 436)

The proof, which rests on Lemma 1, is based on the fact that a difference between the arc $\overset{\frown}{ACB}$ and the tangent AD, or the arc $\overset{\frown}{ACB}$ and the chord AB, can be made less than any assignable magnitude by taking B sufficiently close to A.

In Lemma 2 Newton shows that a curvilinear area $AabcdE$ (see Fig. 3.6) can be approached as the limit of the inscribed $AKbLcMdD$ or the circumscribed $AalbmcndoE$ rectilinear areas. The proof is magisterial in its simplicity. Its structure is still retained in present day calculus textbooks in the definition of the definite integral. It consists in showing that the difference between the areas of the circumscribed and the inscribed figures tends to zero, as the number of parallelograms tends to infinity. In fact this difference is equal to the area of parallelogram $ABla$:

"but this rectangle, because its width AB is diminished indefinitely, becomes less than any given rectangle"(Newton 1687/1999, 433).

Notice how in Lemma 2 and Lemma 7 Newton gives a proof of two assumptions that were made in the seventeenth-century "new analysis". The "new analysts" (Newton himself in his early writings!) had assumed that a curve can be conceived as a polygonal of infinitely many infinitesimal sides and that a curvilinear area can be conceived as an infinite summation of infinitesimal strip (see Chapter 2). In the *Geometria curvilinea* and in *Principia*, curves are smooth and curvilinear areas are not resolved into infinitesimal elements. In the synthetical method of fluxions one always works with finite quantities and limits of ratios and sums of finite quantities.

In *De quadratura curvarum* Newton presented a calculus version of the method of prime and ultimate ratios (see (Newton 1691–1692) and (Newton 1704)). However, he made it clear that such symbolical demonstrations were safely grounded in geometry (see 3.5.4). Newton began working on this treatise devoted to "integration" in the early 1690s. It is opened by the declaration that calculus is referred to as only finite flowing quantities: "Mathematical quantities I here consider not as consisting of least possible parts, but as described by a continuous motion. [...] These geneses take place in the reality of physical nature and are daily witnessed in the motion of bodies" (Newton 1704, 122).

For instance, in order to find the fluxion of $y = x^n$ by the method of prime and ultimate ratios, Newton proceeded as follows:

> *Let the quantity x flow uniformly and the fluxion of the quantity x^n needs to be found.* In the time that the quantity x comes in its flux to be $x + o$, the quantity x^n will come to be $(x + o)^n$, that is [when expanded] by the method of infinite series
>
> $$x^n + nox^{n-1} + \frac{1}{2}(n^2 - n)o^2x^{n-2} + \cdots; \tag{3.12}$$
>
> and so the augments o and $nox^{n-1} + \frac{1}{2}(n^2 - n)o^2x^{n-2} + \cdots$ are one to the other as 1 and $nx^{n-1} + \frac{1}{2}(n^2 - n)ox^{n-2} + \cdots$. Now let those augments come to vanish and their last ratio will be 1 to nx^{n-1}; consequently the fluxion of the quantity x is to the fluxion of the quantity x^n as 1 to nx^{n-1}. (Newton 1704, 126–128)

Notice that the increment o is finite and that the calculation aims at determining the limit of the ratio $[(x + o)^n - x^n]/o$ as o tends to zero.

3.2.6. Higher-order fluxions and the Taylor series (1687 to 1692). In the 1690s Newton introduced a notation for fluxions and higher-order fluxions. He wrote $\dot{x}$, $\ddot{x}$, $\dddot{x}$, etc., for first, second, third, etc., fluxions. He also used the notation $\acute{x}$ for the fluent of x. Dots and accents could be repeated to generate higher-order fluxions and higher-order fluents. Newton also employed overindexes in order to avoid the multiplication of dots and accents: so he wrote $\overset{n}{y}$ for the nth fluxion of y (Newton 1967–1981, 7, 17–18 and 162).

In discussing higher-order fluxions, Newton stated that every ordinate y of a curve in the x-y plane can be expressed, assuming $\dot{x} = 1$, as a power series whose nth term is equal to the nth fluxion of y, i.e., $\overset{n}{y}$, divided by $n!$ (see (Newton 1691–1692, 7, 96–98)). He probably arrived at this statement by generalizing his experience with power series (see some examples in 3.2): For all of them this property holds. On the other hand, if we assume that y is expressible as a power series such as $y = a + bx + cx^2 + dx^3 + ex^4 + \cdots$, one gets immediately that $y(0) = a$, $\dot{y}(0) = b$, $\ddot{y}(0) = 2c$, etc.

Newton thus stated a theorem, nowadays called the Taylor theorem, which was to play an important role in the development of eighteenth-century calculus (see Chapter 4).

It should be noted that already in the *Principia* (e.g., Scholium to Proposition 93, Book 1, and Proposition 10, Book 2) Newton had come close to stating that the nth term of a power series expansion is proportional to the nth fluxion. He had actually stated that the first term represents the ordinate, the second the tangent (or the velocity), the third the curvature (or the acceleration), and so on. In Book 3 he had also solved the problem of determining "a parabolic curve that will pass through any number of given points" by a procedure which is equivalent to the so-called Gregory-Newton interpolation formula (a version of which he discovered in about 1676). It is indeed remarkable to see how important power series were in the work of Newton. From his early research on tangents and quadratures to his mature development of a theory of higher-order fluxions he used power series as a major analytical tool.

3.3. Leibniz's differential and integral calculus

3.3.1. A mathematician and diplomat. Gottfried Wilhelm Leibniz was born in Leipzig in 1646 from a Protestant family of distant Slavonic origins. His father, a professor at Leipzig University, died in 1652, leaving a rich library, where the young Gottfried began his scholarly life. He studied philosophy and law in the Universities of Leipzig, Jena and Altdorf. He also received some elementary education in arithmetic and algebra. Early on he formulated a project for the construction of a mathematical language with which deductive reasoning could be conducted. His manuscripts related to symbolical reasoning reveal anticipation of the nineteenth-century algebra of logic. Leibniz never abandoned his programme of devising a "characteristica universalis". As we will see, he conceived his mathematical research as part of this ambitious project. More specifically, his interest in number sequences played a role in the discovery of differential and integral calculus. After receiving his doctorate in 1666 from the University of Altdorf, he entered into the service of the Elector of Mainz. From 1672 to 1676 he was in Paris on a diplomatic mission. Here he met several distinguished scholars, most notably Christiaan Huygens, who belonged to the recently established *Académie Royale des Sciences.* It was in Paris, following Huygens's counsel, that Leibniz learned mathematics. In a few months he had digested all the relevant contemporary literature and was able to contribute original research. His discovery of calculus dates from the years 1675–1677. He published the rules of differential calculus in 1684 in the *Acta eruditorum*, a scientific journal that he had helped to found in 1682. In 1676 his seminal period of study in Paris came to an end. After 1676 Leibniz worked in the service of the Court of Hanover. He embarked on political projects, the most ambitious of which was the reunification of the Christian churches. Leibniz was very good in divulging his mathematical discoveries through scientific journals and learned correspondence. While Newton kept his method secret, Leibniz made great efforts to promote the use of calculus. In Basel, Paris and Italy several mathematicians, such as the Bernoulli brothers, l'Hôpital, Varignon, Manfredi, and Riccati began to use and defend the new calculus of sums and differences. A notable advance occurred at the turn of the century when Jakob and Johann Bernoulli extended integral calculus and applied it to dynamics.

Leibniz died in 1716. His funeral was attended only by his relatives and by his secretary. Leibniz's intellectual interests spanned from technology to mathematics, from physics to logic, from politics to religion. He is remembered as one of the profoundest philosophers and one of the most creative mathematicians of all ages.

3.3.2. Infinite series (1672 to 1673). Leibniz's interests in combinatorics led him to consider finite numerical sequences of differences such as

$$b_1 = a_1 - a_2, \quad b_2 = a_2 - a_3, \quad b_3 = a_3 - a_4, \dots \tag{3.13}$$

He noted that it is possible to obtain the sum $b_1 + b_2 + \cdots + b_n$ as a difference, $a_1 - a_{n+1}$. When extrapolated to the infinite, this simple law led to interesting results with infinite series. For instance, in order to find the sum of the series of reciprocals of the triangular numbers

$$\sum_{n=1}^{\infty} \frac{2}{n(n+1)} = \sum_{n=1}^{\infty} b_n, \tag{3.14}$$

Leibniz noted that the terms of this series may be expressed using a difference sequence by setting

$$b_n = \frac{2}{n} - \frac{2}{n+1} = a_n - a_{n+1}. \tag{3.15}$$

Therefore

$$\sum_{n=1}^{s} b_n = a_1 - a_{s+1} = 2 - \frac{2}{s+1}. \tag{3.16}$$

So, if we "sum" all the terms, we obtain 2.

Leibniz applied this procedure successfully to several other examples. For instance he considered the "harmonic triangle" (see Fig. 3.7). In the harmonic triangle the nth oblique row is the difference sequence of the $(n+1)$th oblique row. It follows, for instance, that

$$\frac{1}{4} + \frac{1}{20} + \frac{1}{60} + \frac{1}{140} + \cdots = \frac{1}{3}. \tag{3.17}$$

This research on infinite series implies an idea that played a central role in Leibnizian calculus (see (Bos 1980, 61)). The sum of an infinite number of terms b_n can be achieved via the difference sequence a_n.

3.3.3. The geometry of infinitesimals (1673 to 1674). In 1673 Leibniz met with the idea of the so-called "characteristic triangle". He was reading Pascal's *Lettres de "A. Dettonville"* (1659). Pascal, in dealing with quadrature problems, had associated a point on a circumference with a triangle with infinitesimal sides. Leibniz generalized this idea. Given any curve (see Fig. 3.8) he associated an infinitesimal triangle to an arbitrary point P. One can think of the curve as a polygonal constituted by infinitely many infinitesimal sides. The prolongation of one of the sides gives the tangent to the curve. A line at right angles with one of the sides is the normal. Call t and n the length of the tangent and the normal, respectively, intercepted between P and the x-axis. From the similarity of the three

$$1$$
$$\frac{1}{2} \quad \frac{1}{2}$$
$$\frac{1}{3} \quad \frac{1}{6} \quad \frac{1}{3}$$
$$\frac{1}{4} \quad \frac{1}{12} \quad \frac{1}{12} \quad \frac{1}{4}$$
$$\frac{1}{5} \quad \frac{1}{20} \quad \frac{1}{30} \quad \frac{1}{20} \quad \frac{1}{5}$$
$$\frac{1}{6} \quad \frac{1}{30} \quad \frac{1}{60} \quad \frac{1}{60} \quad \frac{1}{30} \quad \frac{1}{6}$$
$$\frac{1}{7} \quad \frac{1}{42} \quad \frac{1}{105} \quad \frac{1}{140} \quad \frac{1}{105} \quad \frac{1}{42} \quad \frac{1}{7}$$

FIGURE 3.7

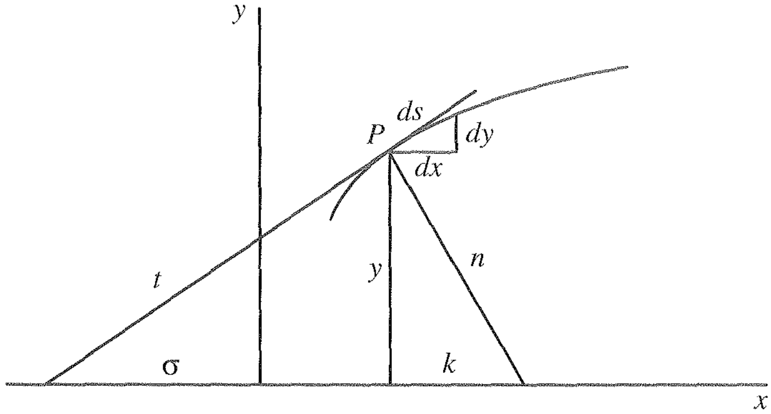


FIGURE 3.8

triangles shown in Fig. 3.8, Leibniz obtained several geometrical transformations which allowed him to transform a problem of quadrature into another problem. He stated equivalences which he would later write as $\int kdx = \int ydy$, $\int ydx = \int \sigma dy$, $\int yds = \int tdy$, $\int yds = \int ndx$ (here n is the normal, t is the tangent, k is the subnormal and σ is the subtangent). The most useful transformation obtained by Leibniz in 1673–1674, i.e., the years immediately preceding the invention of the algorithm of calculus, is the "transmutation theorem" ((Hofmann 1949, 32–35) and (Bos 1980, 62–64)).

Leibniz considered a smooth convex curve OAB (see Fig. 3.9). The problem is to determine the area $OABG$. Let PQN be the characteristic triangle associated to the point P. The area $OABG$ can be seen either as the sum of infinitely many strips $RPQS$ or as the sum of the triangle OBG plus the sum of infinitely many triangles OPQ. We can write

$$OABG = \sum RPQS = \frac{1}{2} OG \cdot GB + \sum OPQ. \tag{3.18}$$

Let the prolongation of PQ (i.e., the tangent in P) meet the y-axis in T and let OW be normal to the tangent. Triangle OTW is thus similar to the characteristic triangle PQN; therefore,

$$\frac{PN}{OW} = \frac{PQ}{OT}. \tag{3.19}$$

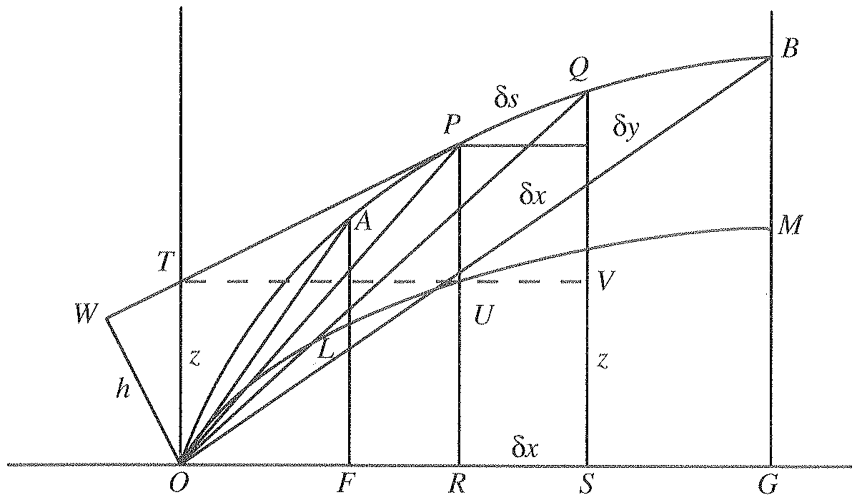


FIGURE 3.9

The area of the infinitesimal triangle OPQ is thus

$$OPQ = \frac{1}{2}OW \cdot PQ = \frac{1}{2}OT \cdot PN. \tag{3.20}$$

Leibniz defines a new curve OLM, related to the curve OAB through the process of taking the tangent. The new curve has an ordinate in R equal to OT. Geometrically the construction is obtained by drawing the tangent in P and determining the intersection T between the tangent and y-axis. In symbols not yet available to Leibniz, the ordinate z of the new curve OLM is $z = y - xdy/dx$.

Leibniz has thus shown that

$$\begin{aligned} OABG &= \frac{1}{2}OG \cdot GB + \Sigma OPQ \\ &= \frac{1}{2}OG \cdot GB + \Sigma\frac{1}{2}OT \cdot PN \\ &= \frac{1}{2}OG \cdot GB + \frac{1}{2}OLMG, \end{aligned} \tag{3.21}$$

where $OLMG$ is the area subtended by the new curve. In modern symbols, setting y as the ordinate of the curve OAB (see (Bos 1980, 65)),

$$\int_0^{x_0} ydx = \frac{1}{2}x_0y_0 + \frac{1}{2}\int_0^{x_0} zdx = \frac{1}{2}x_0y_0 + \frac{1}{2}\int_0^{x_0} ydx - \frac{1}{2}\int_0^{x_0} x\frac{dy}{dx}dx. \tag{3.22}$$

Leibniz's geometrical "transmutation" is thus equivalent to integration by parts. He was later (see, e.g., (Leibniz 1714, 408)) to express it as

$$\int ydx = xy - \int xdy. \tag{3.23}$$

Leibniz thus achieved, through the geometry of the infinitesimal characteristic triangle, a reduction formula for integration. The integration of curve OAB is reduced to the calculation of the area subtended to an auxiliary curve OLM related to OAB through the process of taking the tangent. The relation of the tangent and quadrature problem began thus to emerge in Leibniz's mind. This work with the characteristic triangle also made him aware of the fruitfulness of dealing with infinitesimal quantities.

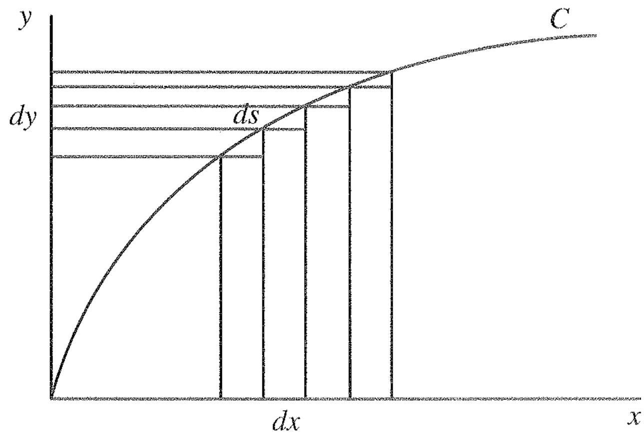


FIGURE 3.10

3.3.4. The calculus of infinitesimals (1675 to 1686). During 1675 Leibniz made the crucial steps which led him to forge the algorithm which is still utilized, though in a revised form and in a different conceptual context. He began considering two geometric constructions which had played a relevant role in seventeenth-century infinitesimal techniques: viz., the characteristic triangle and the area subtended to a curve as the sum of infinitesimal strips.

Let us consider a curve C (see Fig. 3.10) in a Cartesian coordinate system. Leibniz imagines a subdivision of the x-axis into infinitely many infinitesimal intervals with extremes x_1, x_2, x_3, etc. He further defines the differential $dx = x_{n+1} - x_n$. On the curve and on the y-axis one has the corresponding successions s_1, s_2, s_3, etc., and y_1, y_2, y_3, etc. Therefore $ds = s_{n+1} - s_n$ and $dy = y_{n+1} - y_n$. The characteristic triangle has sides dx, ds, dy. The tangent to the curve C forms an angle γ with the x-axis such that $\tan\gamma = dy/dx$. The area subtended to the curve is equal to the sum of infinitely many strips ydx. Leibniz initially employed Cavalieri's symbol "omn.", but he soon replaced this notation with the now familiar $\int ydx$, where $\int$ is a long "s" for "sum of". The first published occurrence of the d-sign was in (Leibniz 1684), while the integral appeared in (Leibniz 1686). Three aspects of Leibniz's representation of the curve C in terms of differentials should be noted.

1) The symbols d and $\int$ applied to a finite quantity x generate an infinitely little and an infinitely great quantity, respectively. So, if x is a finite angle or a finite line, dx and $\int x$ are, respectively, an infinitely little and an infinitely great angle or line. Thus the two symbols d and $\int$ change the order of infinity but preserve the geometrical dimensions. Notice that Newton's dot symbol does not do that. If x is a finite flowing line, $\dot{x}$ is a finite velocity.

2) Since geometrical dimension is preserved, the symbols d and $\int$ can be iterated to obtain higher-order infinitesimals and higher-order infinites. So ddx is infinitely little compared to dx, and $\int\int x$ is infinitely great compared to $\int x$. A hierarchy of infinitesimals and infinites is thus obtained. Higher-order differentials were denoted by repeating the symbol d. It became usual, from the mid-1690s, to abbreviate $dd \ldots d$ (n times) by d^n, so that the nth differential of x is $d^n x$.

3) The representation of the curve C in terms of differentials can be achieved in a variety of ways. One can chose the progressions of x_n, y_n and s_n so that dx

is constant or dy is constant or ds is constant. Or one can choose the three above-mentioned progressions such that dx, dy and ds are all variable. For instance, the choice of dx constant (i.e., the x_n equidistant) generates successions of y_n and s_n where ds and dy are not (generally) constant. As Bos has shown in (Bos 1974) the choice of dx constant is equivalent to selecting x as the independent variable and s and y as dependent variables. (The Newtonian equivalent is to choose $\dot{x}$ constant, i.e., x flowing with uniform velocity.)

Bos stresses, moreover, that the Leibnizian calculus is not concerned with "functions" and "derivatives" but with progressions of variable quantities and their differences. Therefore we should not read, for instance, dy/dx as the derivative of $y(x)$ as a function of x but as a ratio between two differential quantities, dy and dx. The conception of dy/dx as a ratio renders the algebraical manipulation of differentials quite "natural". For instance, the chain rule is nothing more than a compound ratio:

$$\frac{dy}{dx} = \frac{dy}{dw}\frac{dw}{dx}. \tag{3.24}$$

Selecting a variable x so that dx is constant simplifies the calculations since $ddx = 0$ and higher-order differentials of x are cancelled. There is another way for cancelling higher-order differentials. When one has a sum $A + \alpha$ and α is infinitely little in comparison to A, it can be stated that $A + \alpha = A$. This rule of cancellation for higher-order infinitesimals can be stated as follows:

$$d^n x + d^{n+1} x = d^n x. \tag{3.25}$$

Leibniz calculated the differential of xy and x^n as follows:

$$d(xy) = (x + dx)(y + dy) - xy = xdy + ydx + dxdy = xdy + ydx,$$

while

$$dx^n = (x + dx)^n - x^n = nx^{n-1}dx + dx^2(\cdots) = nx^{n-1}dx.$$

In fact, he assumed that $dxdy$ cancels against $xdy + ydx$ and that dx^2 cancels against dx (see 3.5.2 for Leibniz's attempts to justify this procedure).

Differentials of roots such as $y = \sqrt[b]{x^a}$ can be achieved by rewriting $y^b = x^a$, taking the differentials, $by^{b-1}dy = ax^{a-1}dx$, and rearranging so that $d\sqrt[b]{x^a} = (a/b)dx\sqrt[b]{x^{a-b}}$. A similar reasoning leads to $d(1/x^a) = -adx/x^{a+1}$.

Leibniz was clearly proud of the extension of his calculus. In the predifferentiation period (see 2.2) roots and fractions were difficult to handle. Leibniz published the rules for differential calculus in 1684 in a short and difficult paper which bears a title with the English translation *A new method for maxima and minima as well as tangents, which is neither impeded by fractional nor irrational quantities, and a remarkable type of calculus for them.*

Leibniz generally performed integration by reductions of $\int ydx$ through methods of variable substitution or integration by parts. These methods could be worked out in a purely analytical way. Instead of requiring complex geometrical constructions of auxiliary curves (as in the method of transmutation), the new notation allowed algebraical manipulations.

The most powerful method for performing integrations came from the understanding of the fundamental theorem of calculus. The notation d and $\int$, for difference and sum, immediately suggests the inverse relationship of differentiation and integration. Leibniz conceived $\int ydx$ as the "sum" of an infinite sequence of strips

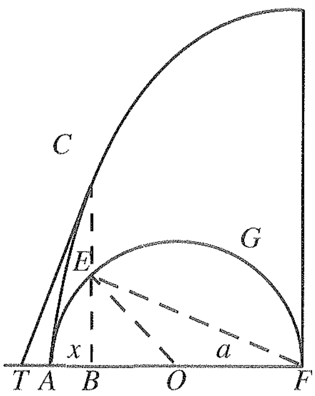


FIGURE 3.11

ydx. From his research on infinite series he knew that a sum of an infinite sequence can be obtained from the difference sequence (see 3.3.2). In order to reduce $\int ydx$ to a sum of differences, one must find a z such that $dz = ydx$. Thus, at once,

$$\int ydx = \int dz = d\int z = z. \tag{3.26}$$

Once the inverse relation of differentiation and integration is understood, several techniques of integration follow. For instance the rule of transmutation (integration by parts) comes by inverting $d(xy) = xdy + ydx$. We thus obtain $xy = \int d(xy) = \int xdx + \int ydx$.

As an example of Leibniz's inverse algorithm we can consider the applicaton of the transmutation theorem to the quadrature of the cycloid generated by a circle of radius a rolling along the vertical line $x = 2a$ (see Fig. 3.11). The ordinate BC is equal to $BE + EC = BE + AE$, where AE is the length s of the circular arc. Since $ds/a = dx/\sqrt{2ax - x^2}$, it follows that $s = \int_0^x adu/\sqrt{2au - u^2}$. (Nowadays we have notation for the elementary transcendental functions and we would write $s = a \cdot \arccos((a - x)/a)$.) Thus the equation of the cycloid is

$$y = \sqrt{2ax - x^2} + \int_0^x adu/\sqrt{2au - u^2}. \tag{3.27}$$

Since $dy/dx = (2a - x)/\sqrt{2ax - x^2}$, from (3.22),

$$\int_0^{x_0} ydx = x_0y_0 - \int_0^{x_0} \sqrt{2ax - x^2}dx. \tag{3.28}$$

If we take $x_0 = 2a$ and $y_0 = \pi a$, formula (3.28) gives $3\pi a^2/2$ for the area subtended under the half-arch (see (Dupont and Roero 1991, 118–119)).

Leibniz was greatly interested in the applications of his calculus to geometry and dynamics. In this applied context he wrote and solved several differential equations. This very important subject entered into the world of continental mathematics thanks to Leibniz's development of integration techniques (see 11.2.2).

3.4. Mathematizing force

The publication in 1687 of Newton's *Principia* was perhaps the major event of seventeenth-century natural philosophy. The reaction of Leibniz to the *Principia* is too complex a subject to be tackled here. To mention just a few points, Leibniz disagreed with Newton's cosmology of universal gravitation, with his conceptions of absolute time and space, with his dynamical principles, and with his theological views (see (Bertoloni Meli 1993a)). It is of interest for us that Leibniz and his school were critical of Newton's mathematical methods in dynamics.

Even though Newton was one of the discoverers of calculus, he made explicit use of it in only a few isolated propositions in the *Principia.* Instead he employed the synthetical method of fluxions, i.e., the method of prime and ultimate ratios (3.2.5). Limits of ratios and limits of sums, as well as infinitesimals of various orders, occur very often in his geometrical dynamics. A "translation" into the language of calculus thus might appear trivial. However, the mathematicians who, at the beginning of the eighteenth century, set themselves the task of applying the calculus to Newton's dynamics (most notably Pierre Varignon, Jakob Hermann, and Johann Bernoulli) had difficult problems to surmount. In some cases, the geometrical demonstrations of the *Principia* can be translated almost at once into calculus concepts; in other cases, this translation is complicated, unnatural, or even problematic.

Today, we take it for granted that calculus is a better suited tool than geometry for dealing with dynamics. But at the beginning of the eighteenth century, the choice of mathematical methods to be applied to dynamics was problematic. Newton's mathematization of dynamics was mainly, even though not exclusively, geometrical and several members of the Newtonian school, up to Colin Maclaurin and Matthew Stewart at the middle of the eighteenth century followed Newton from this point of view (see (Guicciardini 1989)).

Before writing the *Principia*, Newton had already turned his attention toward geometrical methods. In the 1670s he was led to distance himself from his early highly analytical mathematical research. Newton began to criticize modern mathematicians: He stressed the mechanical character of modern algebraical methods, their utility only as heuristic tools and not as demonstrative techniques, and the lack of referential clarity of the concepts employed. By contrast, he characterized the "geometry of the Ancients" as simple, elegant, concise, adherent to the problem posed, and always interpretable in terms of existing objects. Needless to say, notwithstanding Newton's rhetorical declaration of continuity between his methods and the methods of the "Ancients," his geometrical dynamics is a wholly seventeenth-century affair.

The reasons that induced this champion of analytics, series, infinitesimals and algebra to spurn his analytical research are complex. They have to do with foundational worries about the nature of infinitesimal quantities as well as with his desire to find in geometry a unifying principle of techniques which grew wildly in his early writings. They also have to do with his dislike of Descartes, towards anything Cartesian, and with his admiration for the geometrical methods of Huygens (see (Westfall 1980, 377–381)).

But other factors combined to give to the *Principia* the geometrical form we know. A sixteenth-century approach to natural philosophy, exemplified in the works of Johannes Kepler and Galileo Galilei, saw the *Book of Nature* as written in circles and triangles, not in equations. Furthermore, the community of natural philosophers to which Newton addressed the *Principia* was trained in geometry, certainly not in calculus: In 1687 almost a still unpublished discovery. It would have been

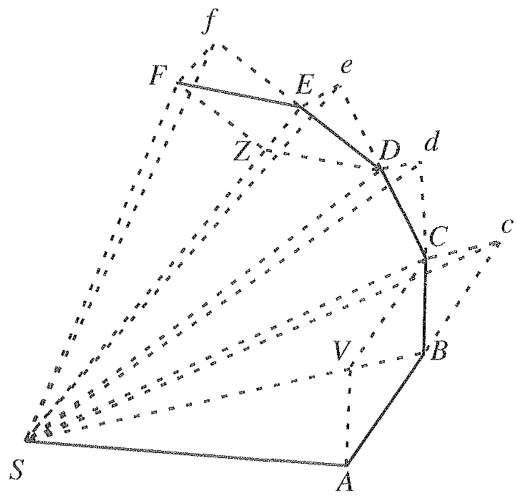


FIGURE 3.12

hopelessly difficult for them to understand a completely new dynamics expressed into a completely new language.

Another important factor that led Newton to use geometry in dynamics has to do with the relative weakness of calculus in 1687. Newton knew how to apply calculus to the simplest problems. We have manuscripts in which he writes fluxional (i.e., differential) equations of motion for the one-body problem ((Newton 1691–1692, 122–129) and (Guicciardini 1999)). However, universal gravitation allows perturbed motions in planetary orbits. The possibility of mathematizing fine details of planetary motions (such as the precession of equinoxes) or planetary shapes and tides was crucial for Newton and his followers. The calculus was not yet powerful enough to allow such dynamical studies. Geometry on the other hand offered a means to tackle these problems, at least at a qualitative level (see (Greenberg 1995)).

Employing the geometry of prime and ultimate ratios, refusing the new analysis in favour of the synthetical method of fluxions, was not therefore a defensive, backward move, but rather it was seen by Newton as a progressive move, a choice of a more powerful method. Newton believed this method was better, both from a foundational point of view and from a demonstrative point of view.

Let us consider, as an example of Newton's geometrical techniques in dynamics, the treatment of Kepler's area law of planetary motions, i.e., Proposition 1 of Book 1 of the *Principia*. This proposition states that Kepler's area law holds for any central force. Newton's geometric proof is based on an intuitive theory of limits. In the *Principia* we read:

> The areas which bodies made to move in orbits describe by radii drawn to an unmoving centre of forces lie in unmoving planes and are proportional to the times. (Newton 1687/1999, 444)

Newton's proof is as follows. Divide the time into equal and finite intervals, Δt_1, Δt_2, Δt_3, etc. At the end of each interval the force acts on the body "with a single but great impulse" (ibid.) and the velocity of the body changes instantaneously. The resulting trajectory (see Fig. 3.12) is a polygonal $ABCDEF$. The areas SAB, SBC, SCD, etc., are swept by the radius vector in equal times. Applying the first two laws of motion, it is possible to show that they are equal. In fact, if at the end of Δt_1, when the body is at B, the centripetal force did not act, the body would continue in a straight line with uniform velocity (because of the first law of motion). This means that the body would reach c at the end of Δt_2 such that $AB = Bc$. Triangles SAB and SBc have equal areas. However, we know that at the end of Δt_1, when the body is at B, the centripetal force acts. Where is the body at the end Δt_2? In order to answer this question, one has to consider how Newton, in Corollary 1 to the laws, defines the mode of action of two forces acting "simultaneously": "A body, acted on by two forces simultaneously, will describe the diagonal of a parallelogram in the same time as it would describe the sides by those forces separately" (ibid., 417). Invoking the above corollary, Newton deduces that the body will move along the diagonal of parallelogram $BcCV$ and reaching C at the end of Δt_2. Cc is parallel to VB, so that triangles SBc and SBC have equal areas. It follows that triangles SAB and SBC have equal areas. One can iterate this reasoning and construct points C, D, E, F. They all lie on a plane, since the force is directed towards S, and the areas of triangles SCD, SDE, SEF, etc., are equal to the area of triangle SAB. The body therefore describes a polygonal trajectory which lies on a plane, and the radius vector SP sweeps equal areas SAB, SBC, SCD, etc., in equal times. Newton passes from the polygonal to the smooth trajectory by a limit procedure based on the method of prime and ultimate ratios. He writes:

> Now let the number of triangles be increased and their width decreased indefinitely, and their ultimate perimeter ADF will [...] be a curved line; and thus the centripetal force by which the body is continually drawn back from the tangent of this curve will act continually, while any areas described, $SADS$ and $SAFS$, which are always proportional to the times of the description, will be proportional to those times in this case. (Ibid., 445)

That is to say, since Kepler's area law always holds for any discrete model (polygonal trajectory generated by an impulsive force) and since the continuous model (smooth trajectory generated by a continuous force) is the limit of the discrete models for $\Delta t \to 0$, then the area law holds for the continuous model. The area swept by SP is proportional to time.

The Leibnizians proceeded in a completely different way. They tackled Kepler's area law from an analytical point of view. After partial results obtained by Jakob Hermann in 1716 (see (Guicciardini 1999)), they obtained the following analytical representation for centripetal force.

The most natural choice is to use polar coordinates (r, θ) so that the origin coincides with the centre of force. The radial and transversal acceleration are thus expressed by the following two formulae:

$$a_r = \frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 \tag{3.29}$$

and

$$a_t = \frac{rd^2\theta}{dt^2} + 2\frac{dr}{dt}\cdot\frac{d\theta}{dt}. \tag{3.30}$$

Let A be the area swept out by the radius vector. Then $2dA/dt = r^2 d\theta/dt$ and $2d^2A/dt^2 = r^2 d^2\theta/dt^2 + 2r(dr/dt)(d\theta/dt) = ra_t$. For a central force, a_t is equal to zero. By integrating (3.30), we obtain $dA/dt = k$ (i.e., the areal velocity is equal to a constant k). Inversely, if $dA/dt = k$, it follows by differentiation that a_t is zero (i.e., the force is central). Proposition 1 and its inverse are thus embedded in the analytical formulation of transversal and radial acceleration.

The above demonstration is quite straightforward: Mathematically speaking, it requires only elementary calculus and the use of polar coordinates. However, such a demonstration was only worked out in the 1740s in the works of Daniel Bernoulli, Leonhard Euler and Alexis Claude Clairaut on constrained and planetary motion (see (Bertoloni Meli 1993b)).

This example shows how different the approach of the Leibnizian school was to the mathematization of dynamics (see (Whiteside 1970)). In the Leibnizian approach the geometry of infinitesimals is the model from which one can work out differential equations. The trajectory is represented locally in terms of differentials. The study of the geometrical and dynamical relationships of infinitesimals leads to differential equations which can be manipulated algebraically until the result sought is achieved. During the algebraical manipulation the geometrical interpretability of the symbols is not at issue. On the other hand, Newton adheres to geometry: The symbols he employs are always interpreted in geometric terms, and they are actually exhibited in the geometrical model, whose geometrical and dynamical properties are central to the demonstration.

3.5. Newton versus Leibniz

3.5.1. "Not-equivalent in practice". It is not easy to establish a comparison between Leibniz's and Newton's calculi because Leibniz and Newton presented several versions of their calculi. Leibniz never published a systematic treatise but rather divulged the differential and integral calculus in a series of papers and letters. He changed his mind quite often especially on foundational questions. Newton abandoned his earlier version of calculus based on moments and opted for the method of prime and ultimate ratios.

In my opinion, Leibniz's and Newton's calculi have sometimes been contrasted too sharply. For instance, it has been said that in the Newtonian version variable quantities are seen as varying continuously in time, while in the Leibnizian version they are conceived as ranging over a sequence of infinitely close values (Bos 1980, 92). It has also been said that in the fluxional calculus, "time", and in general kinematical concepts such as "fluent" and "velocity", play a role which is not accorded to them in differential calculus. It is often said that geometrical quantities are seen in a different way by Leibniz and Newton. For instance, for Leibniz a curve is conceived as polygonal—with an infinite number of infinitesimal sides—while for Newton curves are smooth (Bertoloni Meli 1993a, 61-73).

These sharp distinctions, which certainly help us to capture part of the truth, are made possible only by simplifying the two calculi. As a matter of fact, they are more applicable to a comparison between the simplified versions of the Leibnizian and the Newtonian calculi codified in textbooks such as l'Hôpital's *Analyse des infiniment petits* (1696) and Simpson's *The Doctrine and Application of Fluxions* (1750) rather than to a comparison between Newton and Leibniz. It seems to me that important aspects of their mathematics are ignored in these historical interpretations. For instance, one should not ignore Leibniz's highly skeptical attitude towards the existence of infinitesimals: He would have agreed with Newton that variables vary continuously and that curves are smooth. Leibniz explicitly employed infinitesimals as heuristic devices. In much the same way Newton conceived "moments" as useful abbreviations which can be eliminated by translating infinitesimalist proofs into rigorous limit-based proofs. Furthermore, Newton's conception of "time" as used in the fluxional calculus is highly abstract: He was quite careful to avoid any identification of "fluxional time" with "real time". "Fluxional time" is just a variable fluent with constant fluxion. So the fluxional calculus is not simply founded on kinematics but rather of the abstract concept of continuous variation.

The differences between the Leibnizian and the Newtonian calculi should not be overstressed. In particular, as I shall argue in this section, the differences *should not be looked for at the syntactic or at the semantic level but rather at the pragmatic level.* After all, the two calculi shared a great deal in common both at the syntactic level of the algorithm and at the semantic level of the interpretation of the algorithm's symbols and the justification of the algorithm's rules. It is possible to translate between the fluxional and the differential calculus (through correspondences between $\dot{x}o$ and dx). The Leibnizian and the Newtonian mathematicians made such translations: They were aware that there is not a single theorem which can be proved in one of the two calculi and which cannot have a counterpart in the other. It was exactly this "equivalence" which gave rise to the quarrel over priority.

In discussing the question of equivalence, A. R. Hall writes quite appropriately:

> Did Newton and Leibniz discover the same thing? Obviously, in a straightforward mathematical sense they did: [Leibniz's] calculus and [Newton's] fluxions are not identical, but they are certainly equivalent. [...] Yet one wonders whether some more subtle element may not remain, concealed, for example, in that word "equivalent". I hazard the guess that unless we obliterate the distinction between "identity" and "equivalence", then if two sets of propositions are logically equivalent, but not identical, there must be some distinction between them of a more than trivial symbolic character. (Hall 1980, 257–258)

In order to explore this more subtle and concealed level, where a comparison between Newton's and Leibniz's calculi can be established, S. Sigurdsson has proposed to use the category "not-equivalent in practice". Despite the equivalence of the two calculi,

> [this] equivalence breaks down once it is realized that competing formalisms suggest separate directions for research and therefore generate different kinds of knowledge. (Sigurdsson 1992, 110)

Similarly I. Schneider has remarked that "the starting point, the main emphasis and the expectations of the two pioneers were not at all identical" (Schneider 1988, 142). D. Bertoloni Meli has drawn a comparison between a Newtonian and a Leibnizian mathematician and two programmers who use different computer languages:

> Even if the two programmes are designed to perform the same operations, the skills required to manipulate them may differ considerably. Thus subsequent modifications and developments

> may follow different routes, and this is precisely what happened in Britain and on the Continent in the eighteenth century: despite the initial "equivalence" of fluxions and differentials. (Bertoloni Meli 1993a, 202)

I agree with the approach of the above-mentioned scholars. Rather than looking for sharp distinctions between the two calculi, we should look for subtler, less evident aspects. Newton and Leibniz had two "mathematically equivalent" symbolisms. At the syntactical level they could translate each other's results and, at the semantical level, they agreed on important foundational questions. Nonetheless, at the pragmatic level, they oriented their research in different directions. Belonging to the Newtonian or to the Leibnizian school meant having different skills and different expectations. It meant stressing different lines of research and different values. After all, it often happens in history of mathematics that the difference between two schools does not lay in logical or conceptual incommensurabilities but rather in more pragmatic aspects: such as the teaching methods, the formation of mathematicians, the expectations for future research, the system of values which support the view that a method of proof is preferable to another, etc.

In the following three sections, I will look for such a comparison between the two schools focusing on three aspects: the conceptual foundations, the algorithms and the role of geometry.

3.5.2. The problem of foundations. The problem of foundations did not exists in the seventeenth century in the form which it took in the early nineteenth century (see Chapter 6). One of the most important foundational questions faced by seventeenth- and eighteenth-century mathematicians was a question concerning the referential content of mathematical symbols (typically "do infinitesimals exists?"). This "ontological" question was followed by a "logical" question about the legitimacy of the rules of demonstration of the new analysis (typically "is $x+dx = x$ legitimate?"). To these two questions Newton and Leibniz gave similar answers.

They both stated that (a) *actual infinitesimals do not exist; they are useful fictions employed to abbreviate proofs*, (b) *infinitesimals should be defined rather as varying quantities in a state of approaching zero*, (c) *infinitesimals can be completely avoided by limit-based proofs, which constitute the rigorous formulation of calculus*, (d) *limit-based proofs are a direct version of and are thus equivalent to the indirect,* ad absurdum *Archimedean method of exhaustion.*

Once the calculus had been reduced to limit-based proofs, the logical question took the form: "Are limit-based proofs legitimate?" In order to answer this question, both Newton and Leibniz used the concept of continuity. However, the former legitimated limits in terms of our intuition of continuous flow, while the latter referred to a philosophical "principle of continuity".

To the question, "Do differentials exist?", many Leibnizians answered in the affirmative. Leibniz did not. From his very early manuscripts (see (Leibniz 1993)) to his mature works, it is possible to infer that for him actual differentials were just "fictions", symbols without referential content (see (Knobloch 1994)).

Nonetheless the use of these symbols was justified, according to Leibniz, since correct results could be derived by employing the algorithm of differentials. As Leibniz said, differentials are "fictions", but "well-founded fictions". Why "well-founded"? Leibniz seems to have had the following answer. He denies the actual infinite and actual infinitesimal and conceives the differentials as "incomparable quantities": varying quantities which tend to zero. In his writings of the 1690s Leibniz describes these "incomparables" as magnitudes in a fluid state which is different from zero but which is not finite. These quantities would give a meaning to dy/dx as a ratio between two quantities. In fact, if dy and dx are zero, we have the problem of giving a value to 0/0, but if they are finite, they cannot be neglected (thus $x + dx = x$ would be invalid).

However, in other later writings Leibniz stated that differentials are well-founded, since they are symbolic abbreviations for limit-procedures. From this viewpoint, the calculus of differentials is a shorthand for a calculus of finite quantities and limits, equivalent to Archimedean exhaustion. He wrote:

> In fact, instead of the infinite or the infinitely small, one can take magnitudes that are so large or so small that the error will be less than the given error, so that one differs from the style of Archimedes only in the expressions, which are, in our method, more direct and more apt to the art of discovery. (Leibniz 1701, 350)

Newton's approach to the question of the existence of infinitesimals is similar. Newton also spoke of infinitesimals ("moments" or "indefinitely little quantities") as a shorthand for longer and more rigourous proof given in terms of limits. He also speaks of infinitesimals as "vanishing quantities" in such a way that they seem to be defined as something in between zero and finite, as quantities in the state of disappearing, or coming to existence, in a fuzzy realm in between nothing and finite. More often he makes clear that infinitesimals can be replaced by using limits.

There is not, therefore, a strong conceptual opposition between Leibniz and Newton but rather a different attitude. Both agreed that limits provide a rigorous foundation for the calculus, but for Leibniz this was more a rhetorical move in defence of the legitimacy of the differential algorithm, while for Newton this was a programme that should be implemented. While Newton explicitly developed a theory of limits (see 3.2.5), Leibniz simply alluded to the possibility of building the calculus based on such a theory. Leibniz could live with the infinitesimal quantities; Newton made a serious effort in the *Principia* and *De quadratura* to eliminate them (see (Lai 1975)), (Kitcher 1973) and (Guicciardini 1999)).

Leibniz often refers to the heuristic character of calculus in order to justify the use of differentials. For him "metaphysical" questions on the foundations should not interfere with the acceptance of calculus. Calculus, according to Leibniz, should be seen also as an *ars inveniendi*: As such it should be valued by its fruitfulness, more than by its referential content. According to Leibniz, we can calculate with symbols devoid of referential content (for instance, with $\sqrt{-1}$) provided the calculus is structured in such a way as to lead to correct results. Newton could not agree: For him mathematics devoid of referential content could not be acceptable.

The argument of continuity with the "geometry of the Ancients" also played a different role in Newton's and in Leibniz's conceptions. For Newton, showing a continuity between his method and the methods of Archimedes was a crucial step in guaranteeing the acceptability of the "new analysis". Leibniz stressed this continuity only in passing references deviced to reassure the dubious or to reply to critics. He preferred to stress the novelty and revolutionary character of his calculus.

The next foundational question concerns the legitimacy of proofs based on limits. Newton in the *Principia* considers the objection that "there is no such thing as an ultimate proportion of vanishing quantities, inasmuch as before vanishing the proportion is not ultimate, and after vanishing it does not exist at all." However, he observes that

> by the same argument it could equally be contended that there is no ultimate velocity of a body reaching a certain place at which the motion ceases; for before the body arrives at this place, the velocity is not the ultimate velocity, and when it arrives there, there is no velocity at all. But the answer is easy; to understand the ultimate velocity as that with which a body is moving, neither before it arrives at its ultimate place and the motion ceases, nor after it has arrived there, but at the very instant when it arrives, that is, the very velocity with which the body arrives at its ultimate place and with which the motion ceases. (Newton 1687/1999, 442)

In order to demonstrate the existence of limits, Newton thus referred to the intuition of continuous motion: We know by intuition that natural systems evolve by continuous motion and that in every instant of time there *is* a velocity of flow.

Leibniz, to the contrary, in order to justify the limiting procedures referred to a metaphysical principle of continuity which he expressed in several forms and contexts (see (Breger 1990).) The "law of continuity" pervades Leibniz's thought. He made use of it in cosmology, in physics and in logic. Thus, invoking the law of continuity, he affirmed that rest can be conceived as an infinitely little velocity or that equality can be conceived as an infinitely little inequality. In 1687 he stated this principle as follows in his difficult philosophical prose:

> When the difference between two instances in a given series or that which is presupposed [*in datis*] can be diminished until it becomes smaller than any given quantity whatever, the corresponding difference in what is sought [*in quaesitis*] (Leibniz 1687, 52)

In order to explain the meaning of this general principle, Leibniz refers to the geometry of conic sections. An ellipse, he says, may approach a parabola as closely as one pleases, so that the difference between the ellipse and the parabola (the difference between what "results") may become "less than any given difference", provided that one of the foci (what is "posed") is removed far enough away from the other. Consequently, the theorems valid for the ellipse can be extrapolated to the parabola "considering the parabola as an ellipse when one of the foci is infinitely distant, or (in order to avoid this expression) as a figure which differs from a certain ellipse less than any given difference" (ibid.). It is the continuous dependence between what is "posed" and what "results" that justifies limit-based reasonings in which one extrapolates to the parabola what has been proved of the ellipses: "In continuous magnitudes the exclusive *extremum* can be treated as inclusive" (Leibniz 1713, 385).

3.5.3. The two algorithms: Method versus calculus. Leibniz's and Newton's algorithms are related through correspondences between $\dot{x}o$ and dx. The two schools could easily translate each other's results. The main advantage of Leibniz's algorithm concerns the integral sign. With Leibniz's $\int ydx$ the integration-variable x is explicitly indicated. Newton's $\boxed{y}$, Qy and $\acute{y}$ need to be accompanied by verbal statements. This has effects on integration techniques. In the Leibnizian calculus, integration by substitutions and by parts can be performed in a more mechanical way. This advantage was recognized by the Newtonians, who often employed hybrid notations: E.g., Maclaurin wrote $F, y\dot{x}$ in (Maclaurin 1742, 665 ff.).

I. Schneider remarks (Schneider 1988, 143) that in Leibniz's calculus the fundamental theorem is somehow "built into" the notation itself. Indeed, Leibniz's symbols d and $\int$ suggest that differentiation and integration are operations and that they are the inverses of each other.

As Scriba has observed (Scriba 1963), Newton emphasized the use of infinite series. He expanded fluents into infinite series and "integrated" termwise. Leibniz also employed this technique. However, Leibniz preferred integration in "closed" form: He looked for quadratures expressed not by infinite series but by a finite combination of "functions". Newton also obtained "closed" integrations, but it is certainly true that for him infinite series played a more prominent role than for Leibniz. This "contrast" is thus a matter of emphasis; i.e., it is a contrast which relates to the values which direct research along different lines.

Leibniz and Newton had equivalent symbolism but different approaches to notation. The former attached great importance to the construction of an efficient algorithm and chose symbols carefully. The latter was not particularly concerned with notation. Leibniz thought of his calculus as part of a general programme leading to the creation of a *mathesis universalis*, a language in which all reasoning could be framed. He often insisted on the advantages of symbolical reasoning as a method of discovery. Nobody, according to Leibniz, could follow a long reasoning without freeing the mind from the "effort of imagination". The calculus was deviced to favour this "blind reasoning" (*cogitatio caeca*) (see (Pasini 1993, 205)).

Newton, on the other hand, did not value mechanical algorithmic reasoning. He always spoke of the geometrical demonstrations of Huygens in the highest terms and contrasted the elegant geometrical methods of the "Ancients" with the mechanical algebraic methods of Descartes (which "provoked to him nausea" (Newton 1967–1981, 4, 277). He made clear that the symbols of the "analytical method of fluxions" had to be interpreted in terms of the "synthetical method". It is this interplay between algorithm and geometry that characterizes Newton's *method*.

Leibniz's concern with symbolism led him to develop an algebra of differentials (see 3.3.4). His main target was the construction of a set of algorithmic rules: a *calculus*. The rules of calculus are instructions on how to manipulate the d's and the $\int$'s, and they allow algorithmic procedures which are as much as possible independent of the initial geometrical context. Leibniz even considered $d^\alpha x$ for a fractional α. We note that the chain rule in Leibnizian terms takes a form (see formula (3.24)) which is suggested by the notation itself. Everything can be done, of course, also in Newton's notation. Newton, however, preferred to give examples which show the rule rather than give the rule itself. For instance, he would introduce the chain rule with an example, as a set of instructions applied to the solution of a particular problem.

3.5.4. The role of geometry. Newton valued geometrical thinking very highly. As we have seen in 3.2.5, he developed a geometrical version of the method of fluxion in the 1670s: He called it the "synthetical method of fluxions" in opposition to the "analytical method". Newton employed the synthetical method especially in dynamics (see 3.4). He often affirmed that the synthetical method was more rigorous and that it actually founded and justified the procedures employed in the analytical method. This foundation and justification depended on two factors.

First of all the geometrical method of fluxions offered a model in which the analytical method could be interpreted. In the geometrical method the fluents and fluxions were exhibited to the eye, their existence in "rerum natura" proved ostensibly. In the second place, Newton conceived his geometrical method of fluxions as a generalization of the method of exhaustion of the "Ancient Geometers".

The role given to geometry by Newton led him to underestimate the importance of notation. If a demonstration is legitimated when each step of it is interpretable in geometric terms, there is no motivation to develop the algorithm independently from geometry.

The complexity of the relationship between calculus and geometry should be stressed here. Newton's method was concerned with "fluxions *and* series". His treatment of series expansions remained a highly analytical aspect in Newtonian fluxional works, even when the interpretation of power series as Taylor expansions paved the way for a geometrical, or kinematical, interpretation of the successive terms (e.g., as position, velocity, acceleration, variation of acceleration, etc.).

On the other hand, Leibniz, notwithstanding his declarations in favour of a calculus as "blind reasoning", always embedded his algorithm in a geometrical interpretation. Leibniz's differentials and integrals, as much as Newton's fluents and fluxions, were referred to as geometrical objects. It is revealing that Leibniz always paid attention to the geometrical dimensions of the combination of symbols occurring in a differential equation. It was by studying the geometry of differentials (e.g., the characteristic triangle) that Leibniz and his immediate followers could extract differential equations. Once a differential equation was obtained, it was, however, handled as much as possible as an algebraic object. From time to time, it was necessary to use geometric thinking to interpret the model under study (see 4.2). Leibnizians had to do so since the rules of the calculus did not yet allow the solution of the problems in geometry and dynamics that they faced (especially when transcendental "functions" occurred). A complete algebraization of calculus came only in the late eighteenth century. The calculus as "blind reasoning" was thus more a *desideratum* than a reality. Reinterpretation of the symbolism in the geometric model was possible, and in many cases necessary, but, contrary to Newton's approach, this reinterpretation was not seen as a value, as a strategy to be pursued.

The stress on algorithmic improvements and on the idea that progress could be obtained by symbolical manipulations had momentous consequences in the Leibnizian school. Continental mathematicians felt that the differential and integral calculus opened new field of research. In this field many new results could easily be obtained by following as a guideline the analogies suggested by the calculus's notation. New generalizations, new relations and formulas could be found. The mechanization and standardization of mathematical research rendered possible by the stress over the algorithm rendered the Leibnizian school much more active and open to innovation.

Leibniz and the Leibnizian mathematicians looked at the geometrical proofs of Newton's *Principia* with suspicion. One of their aims was to translate Newton's geometrical proofs into the language of the differential and integral calculus. Indeed mechanics proved to be a great source of inspiration for Leibnizians. It is by trying to develop new mathematical tools for the mechanics of extended bodies (rigid, elastic and fluid) that mathematicians such as Varignon, Johann and Daniel Bernoulli, Clairaut, Euler, d'Alembert, and Lagrange enriched calculus by developing new concepts and techniques (see (Truesdell 1968)). Such important results of eighteenth-century calculus as trigonometric series, partial differential equations, and the calculus of variations were to a great extent motivated by the analytical approach to dynamics that Leibniz had sought to promote (see Chapters 4, 11, and 12). The eighteenth century was thus characterized by the analytical programme emphasized by the Leibnizian school, while the role attributed to geometry by Newton and his followers faded away.

Bibliography

Bertoloni Meli, D. 1993a. *Equivalence and priority: Newton versus Leibniz.* Oxford: Clarendon Press.

Bertoloni Meli, D. 1993b. The Emergence of Reference Frames and the Transformation of Mechanics in the Enlightenment. *Historical Studies in the Physical and Biological Sciences*, **23** (2), 301–335.

Bos, H. J. M. 1974. Differentials, Higher-Order Differentials and the Derivative in the Leibnizian Calculus. *Archive for History of Exact Sciences*, **14**, 1–90.

Bos, H. J. M. 1980. Newton, Leibniz and the Leibnizian tradition, in: Grattan-Guinness 1980, 49–93.

Breger, H. 1990. Das Kontinuum bei Leibniz, in: A. Lamarra (ed.), *L'infinito in Leibniz, problemi e terminologia*, Roma: Edizioni dell'Ateneo, 53–67.

Descartes, R. 1637, *Géométrie.* Leyden.

Dupont, P.; Roero, C. S. 1991. *Leibniz 84: il decollo enigmatico del calcolo differenziale.* Rende: Mediterranean Press.

Edwards, C. H. 1979. *The Historical Development of the Calculus.* New York: Springer-Verlag.

Greenberg, J. L. 1995. *The problem of the Earth's shape from Newton to Clairaut.* New York: Cambridge University Press.

Guicciardini, N. 1989. *The development of Newtonian calculus in Britain, 1700-1800.* Cambridge: Cambridge University Press.

Guicciardini, N. 1999. *Reading the Principia: The Debate on Newton's Mathematical Methods for Natural Philosophy from 1687 to 1736.* Cambridge: Cambridge University Press.

Hall, A.R. 1980. *Philosophers at war: the quarrel between Newton and Leibniz.* Cambridge: Cambridge University Press.

Hofmann, J.E. 1949. *Die Entwicklungsgeschichte der Leibnizschen Mathematik whrend des Aufenthaltes in Paris (1672-1676).* München: Oldenbourg.

de l'Hospital, G. F. A. 1696. *Analyse des infiniment petits.* Paris: Imprimerie royale.

Kitcher, P. 1973. Fluxions, limits and infinite littlenesse: a study of Newton's presentation of the calculus. *Isis*, **64**, 33–49.

Knobloch, E. 1994. The infinite in Leibniz's mathematics, the historiographical method of comprehension in context, in: K. Gavroglu (ed.), *Trends in the historiography of science*, Dordrecht/ Boston/ London: Kluwer, 265–278.

Lai T. 1975. Did Newton renounce infinitesimals? *Historia Mathmatica*, **2**, 127–136.

Leibniz, G.W. 1684. Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractas nec irrationales quantitates moratur et singulare pro illis calculi genus. *Acta Eruditorum*, 467–473, or in: *Mathematische Schriften*, **5**, 220–226.

Leibniz, G.W. 1686. De geometria recondita et analysi indivisibilium atque infinitorum. *Acta Eruditorum*, 292–300, or in: *Mathematische Schriften*, **5**, 226–233.

Leibniz, G.W. 1687. Lettre de M.L. sur un principe general utile à l'explication des loix de la nature par la consideration de la sagesse divine, pour servir de replique à la reponse du R.P.D. Malebranche. *Die philosophischen Schriften*, **3**, 51–55.

Leibniz, G.W. 1701. Mémoire de Mr. G.G. Leibniz touchant son sentiment sur le calcul différentiel. *Mathematische Schriften*, **5**, 350.

Leibniz, G.W. 1713. Epistola ad V. Cl. Christianum Wolfium, Professorem Matheseos Halensem, circa scientiam infiniti. *Acta Eruditorm* (Supplement) or in: *Mathematische Schriften*, **5**, 382–387.

Leibniz, G.W. 1714. Historia et origo calculi differentialis. Manuscript. *Mathematische Schriften*, **5**, 392–410.

Leibniz, G.W. 1993. *De quadratura arithmetica circuli ellipseos et hyperbolae cujus corollarium est trigonometria sine tabulis*, E.Knobloch (ed.), Göttingen: Vandenhoeck & Ruprecht.

Maclaurin, C. 1742. *A treatise of fluxions*. Edinburgh: Ruddimans.

Newton, I. 1665. A method whereby to square those crooked lines which may be squared. Manuscript. *Mathematical Papers*, **1**, 302–13.

Newton, I. 1669. De analysi per aequationes numero terminorum infinitas. Manuscript. *Mathematical Papers*, **2**, 206–247.

Newton, I. 1670–71. De methodis serierum et fluxionum. Manuscript. *Mathematical Papers*, **3**, 32–329.

Newton, I. 1687/1999. *Philosophiae Naturalis Principia Mathematica*. London. English translation: *The Principia, Mathematical Principles of Natural Philosophy, A New Translation by I. Bernard Cohen and Anne Whitman with the Assistance of Julia Budenz, preceded by a Guide to Newton's Principia by I. Bernard Cohen.* Berkeley, Los Angeles, London: University of California Press, 1999.

Newton, I. 1691–1692. De quadratura curvarum. Manuscript. *Mathematical Papers*, **7**, 48–129.

Newton, I. 1704. Tractatus de quadratura curvarum [revision of (Newton 1691–1692)]. *Mathematical Papers*, **8**, 92–167.

Newton, I. 1967–1981. *Mathematical Papers*, D.T.Whiteside (ed.), 8 vols., Cambridge: Cambridge University Press.

Oughtred, W. 1631. *Clavis mathematicae*. Oxford.

Pascal, B. 1659. *Lettres de "A. Dettonville"*. Paris.

Pasini, E. 1993. *Il reale e l'immaginario: la fondazione del calcolo infinitesimale nel pensiero di Leibniz.* Milano/Torino: Edizioni Sonda.

Schneider, I. 1988. *Isaac Newton*. München: Beck.

Scriba, C. 1963. The inverse method of tangents, a dialogue between Leibniz and Newton (1675–1677), *Archive for history of exact sciences*, **2**, 113–37.

Sigurdsson, S. 1992. Equivalence, pragmatic platonism, and discovery of the calculus, in: M. J. Nye et al. (eds), *The invention of physical science*, Dordrecht/ Boston/ London: Kluwer, 97–116.

Simpson, T. 1750. *The Doctrine and Application of Fluxions*. London: J. Nourse.

Truesdell, C. A. 1968. *Essays in the history of mechanics*. Berlin: Springer.

van Maanen, J. 1991. From quadrature to integration: thirteen years in the life of the cissoid. *The Mathematical Gazette*, **75**, 1–15.

Viete, F. 1646. *Opera mathematica*. F. van Schooten (ed.). Leiden.

Wallis, J. 1656. *Arithmetica infinitorum*. Oxford.

Westfall, R. S. 1980. *Never at Rest: A Biography of Isaac Newton*. Cambridge: Cambridge University Press.

Whiteside, D. T. 1970. The mathematical principles underlying Newton's *Principia Mathematica*. *Journal for the History of Astronomy* **1**, 116–38.

2 The introduction of v (where v is ultimately to be put equal to y) you may regard as something of a red herring! Newton was making the assumption that v exists, where $f(x) < v < f(x + o)$, such that the rectangle ov = curvilinear area $B\beta\delta D$; since this is always possible for a simply convex curve, the equation he formed was, in consequence, *exact*.

3 In modern notation, if $\int_0^x y\,dx = z$, where $z = f(x)$, then $y = dz/dx = f'(x)$: in particular, if $z = [n/(m + n)]ax^{(m+n)/n}$, then $y = ax^{m/n}$.

4 Although, in earlier researches, Newton did sketch in the outline of a geometrical proof of the fundamental theorem of the calculus (on the lines of the proofs subsequently published by Barrow and Gregory) he seems to have later preferred to rely on the reversibility of the operations, so that differentiation and integration are regarded essentially as inverses, the one of the other $\left(\text{i.e. if } z = f(x) = \int y\,dx, \text{ then } \frac{dz}{dx} = f'(x) = y, \text{ and conversely, if } y = f'(x) = \frac{dz}{dx}, \text{ then } z = \int y\,dx = f(x)\right)$.

Exercise 7

Use Newton's method to show that, if $z = \sqrt{(a^2 + x^2)}$, $y = x/\sqrt{(a^2 + x^2)}$.

SA 7

$$z = \sqrt{(a^2 + x^2)}, \quad z^2 = a^2 + x^2, \quad (z + ov)^2 = a^2 + (x + o)^2$$
$$z^2 + 2ovz + o^2v^2 = a^2 + x^2 + 2ox + o^2$$
$$2ovz + o^2v^2 = 2ox + o^2$$
$$zv = x = zy, \ (v = y)$$
$$y = x/z = x/\sqrt{(a^2 + x^2)}$$

C3.6 FLUXIONS AND FLUENTS

Even before writing the *De Analysi* Newton had experimented with other types of notation and other forms of demonstration (see flow diagram, p. 12). In the small tract written in 1666, he developed a fairly comprehensive treatment of a whole range of calculus problems based on the generation of curves by motion. These ideas, which constituted the foundation of what he called his 'method of fluxions' were developed more fully in the 1671 tract and it is from that that we will quote. The passage which we have chosen conveys well the 'flavour' of Newton's *fluxions* and *fluents* and suggests clear links with mediaeval ideas on motion, developed by Galileo, Torricelli and Barrow. Because of this, you may not find it easy to follow.

> It now remains, in illustration of this analytical art, to deliver some typical problems and such especially as the nature of curves will present. But first of all I would observe that difficulties of this sort may all be reduced to these two problems alone, which I may be permitted to propose with regard to the space traversed by any local motion however accelerated or retarded:
>
> 1 — 1 Given the length of the space continuously (that is, at every [instant of] time), to find the speed of motion at any time proposed.
>
> 2 — 2 Given the speed of motion continuously, to find the length of the space described at any time proposed.
>
> So in the equation $x^2 = y$, if y designates the length of the space described in any time which is measured and represented by a second space x as it increases with uniform

Text 17: Newton on fluxions and fluents. From M. E. Baron and H. J. M. Bos, eds. (1974). *Newton and Leibniz*. History of Mathematics: Origins and Development of the Calculus 3. The Open University Press, pp. 22–25.

D E

Quadratura Curvarum.

Quantitates indeterminatas ut motu perpetuo creſcentes vel decreſcentes, id eſt ut fluentes vel defluentes in ſequentibus conſidero, deſignoque literis z, y, x, v, & earum fluxiones ſeu celeritates creſcendi noto iiſdem literis punctatis $\dot{z}, \dot{y}, \dot{x}, \dot{v}$, Sunt & harum fluxionum fluxiones ſeu mutationes magis aut minus celeres quas ipſarum z, y, x, v, fluxiones ſecundas nominare licet & ſic deſignare $\ddot{z}, \ddot{y}, \ddot{x}, \ddot{v}$, & harum fluxiones primas ſeu ipſarum z, y, x, v fluxiones tertias ſic $\dddot{z}, \dddot{y}, \dddot{x}, \dddot{v}$, & quartas ſic $\ddddot{z}, \ddddot{y}, \ddddot{x}, \ddddot{v}$. Et quemadmodum $\dddot{z}, \dddot{y}, \dddot{x}, \dddot{v}$ ſunt fluxiones quantitatum $\ddot{z}, \ddot{y}, \ddot{x}, \ddot{v}$, & hæ ſunt fluxiones quantitatum $\dot{z}, \dot{y}, \dot{x}, \dot{v}$, & hæ ſunt fluxiones quantitatum primarum z, y, x, v : ſic hæ quantitates conſiderari poſſunt ut fluxiones aliarum quas ſic deſignabo, z', y', x', v', & hæ ut fluxiones aliarum z'', y'', x'', v'', & hæ ut fluxiones aliarum z''', y''', x''', v'''. Deſignant igitur $z'', z', z, \dot{z}, \ddot{z}, \dddot{z}, \ddddot{z}$, &c. ſeriem quantitatum quarum quælibet poſterior eſt fluxio præcedentis & quælibet prior eſt fluens quantitas fluxionem habens ſubſequentem. Similis eſt ſeries $\sqrt{az-zz}''$, $\sqrt{az-zz}'$, $\sqrt{az-zz}$, $\dot{\sqrt{az-zz}}$, $\ddot{\sqrt{az-zz}}$, $\dddot{\sqrt{az-zz}}$, ut & ſeries $\frac{az+zz}{a-z}''$, $\frac{az+zz}{a-z}'$, $\frac{az+zz}{a-z}$, $\dot{\frac{az+zz}{a-z}}$, $\ddot{\frac{az+zz}{a-z}}$, $\dddot{\frac{az+zz}{a-z}}$.

Et notandum eſt quod quantitas quælibet prior in his ſeriebus eſt ut area figuræ curvilineæ cujus ordinatim applicata rectangula eſt quantitas poſterior & abſciſſa eſt z : uti $\sqrt{az-zz}'$ area curvæ cujus ordinata eſt $\sqrt{az-zz}$ & abſciſſa z. Quo autem ſpectant hæc omnia patebit in Propoſitionibus quæ ſequuntur.

M PROP.

Extract from William Jones' edition of Newton's Fluxions, *1711 (Turner Collection, University of Keele).*

3 speed: then $2\dot{x}x$ will designate the speed with which the space at the same moment of time proceeds to be described. And hence it is that in the sequel I consider quantities as though they were generated by continuous increase in the manner of a space which a moving object describes in its course.

We can, however, have no estimate of time except in so far as it is expounded and measured by an equable local motion, and furthermore quantities of the same kind alone, and so also their speeds of increase and decrease, may be compared one with another. For these reasons I shall, in what follows, have no regard to time, formally so considered, but from quantities propounded which are of the same kind shall suppose
4 some one to increase with an equable flow: to this all the others may be referred as though it were time, and so by analogy the name of 'time' may not improperly be conferred upon it. And so whenever in the following you meet with the word 'time' (as I have, for clarity's and distinction's sake, on occasion woven it into my text), by that name should be understood not time formally considered but that other quantity through whose equable increase or flow time is expounded and measured.

But to distinguish the quantities which I consider as just perceptibly but indefinitely growing from others which in any questions are to be looked on as known and determined and are designated by the initial letters a, b, c and so on, I will hereafter call them fluents and designate them by the final letters v, x, y and z. And the speeds with which they each flow and are increased by their generating motion (which I might more readily call fluxions or simple speeds) I will designate by the letters $\dot{v}$, $\dot{x}$, $\dot{y}$ and $\dot{z}$: namely, for the speed of the quantity v I shall put $\dot{v}$, and so for the speeds of the other quantities I shall put $\dot{x}$, $\dot{y}$ and $\dot{z}$ respectively.

Notes

Although it may appear to help if we express some of Newton's statements in the notation of the calculus it should be borne in mind that, by doing so, we risk distortion in that we may give the work a degree of clarity and rigour which was absent.

1 If $s = f(t)$, where t is the time and s is the distance, to find the speed, i.e. $v = \dfrac{ds}{dt} = f'(t)$.

2 If $v = [ds/dt] = \phi(t)$, to find s, i.e. $s = \int_0^t \phi(t)\,dt$. These are the two inverse problems from which Newton developed his calculus.

3 If $y = x^2$, $dy/dt = 2x\,dx/dt$. Since x *increases uniformly*, $dx/dt = \dot{x}$ is taken to be constant.

4 Since time can only be measured by considering uniform motion, we can write, $dx/dt = \dot{x} = 1$, $x = t$. The independent variable, x, increasing uniformly, can be used as a 'measure' of time.

5 If v, x, y, z, are *fluents*, i.e., variables increasing or decreasing with time, then $\dot{v}, \dot{x}, \dot{y}, \dot{z}$, represent the *fluxions*, or speeds, (dv/dt, dx/dt, dy/dt, dz/dt) of these quantities.

This may be an appropriate point to say something about Newton's 'dot'-notation, particularly as you may ultimately want to compare it with the notation developed by Leibniz. Newton experimented with dot-notation of one kind or another from 1665 onwards (see flow diagram, p. 12). He did not settle on the 'standard' Newtonian form of dot-notation until late 1691 and, in the original version of the 1671 tract, he used literal symbols l, m, n, r for the *fluxions* of v, x, y, z. In 1710, William Jones made a transcript of the 1671 treatise on fluxions and inserted the dot-notation and this transcript was subsequently copied in all published editions. In the translation we are using, Whiteside has preferred to adhere to the 'standard' dot-notation because it is a great aid to understanding. In England, at any rate, this notation, used to denote differentiation with respect to t (where t is the time), has become familiar and useful. In comparing the Newtonian dot-notation with the notation developed by Leibniz (dx, dy) we should bear in mind that Newton's decision to adhere to a standard form of dot-notation and to use it consistently was certainly made with knowledge of the existence of the Leibnizian notation in Europe.

Exercise 8

If $y = x^3$, what is the fluxion of x? What is the fluxion of y? How is the fluxion of y related to the fluxion of x? What are x and y called? Which variable is taken by Newton to move uniformly?

SA 8

$\dot{x}$; $\dot{y}$; $\dot{y} = 3x^2\dot{x}$; x and y are called *fluents*, x is taken to move uniformly so that $\dot{x} = k$ (k normally is taken to be 1).

Let us now consider how, given a relation between the *fluent* quantities Newton set about finding a relation between the *fluxions* of these quantities.

Text 17: Newton on fluxions and fluents. From M. E. Baron and H. J. M. Bos, eds. (1974). *Newton and Leibniz*. History of Mathematics: Origins and Development of the Calculus 3. The Open University Press, pp. 22–25.

DEMONSTRATION

The moments of the fluent quantities (that is, their indefinitely small parts, by addition of which they increase during each infinitely small period of time) are as their speeds of flow. Wherefore if the moment of any particular one, say x, be expressed by the product of its speed $\dot{x}$ and an infinitely small quantity o (that is, by $\dot{x}o$), then the moments of the others, v, y, z, [...], will be expressed by $\dot{v}o, \dot{y}o, \dot{z}o$, [...] seeing that
1 $\dot{v}o$, $\dot{x}o$, $\dot{y}o$ and $\dot{z}o$ are to one another as $\dot{v}, \dot{x}, \dot{y}$ and $\dot{z}$.

Now, since the moments (say, $\dot{x}o$ and $\dot{y}o$) of fluent quantities (x and y, say) are the infinitely small additions by which those quantities increase during each infinitely
2 small interval of time, it follows that those quantities x and y after any infinitely small interval of time will become $x + \dot{x}o$ and $y + \dot{y}o$. Consequently, an equation which expresses a relationship of fluent quantities without variance at all times will express that relationship equally between $x + \dot{x}o$ and $y + \dot{y}o$ as between x and y; and so
3 $x + \dot{x}o$ and $y + \dot{y}o$ may be substituted in place of the latter quantities, x and y, in the said equation.

Let there be given, accordingly, any equation $x^3 - ax^2 + axy - y^3 = 0$ and substitute $x + \dot{x}o$ in place of x and $y + \dot{y}o$ in place of y: there will emerge

4
$$(x^3 + 3\dot{x}ox^2 + 3\dot{x}^2o^2x + \dot{x}^3o^3) - (ax^2 + 2a\dot{x}ox + a\dot{x}^2o^2)$$
$$+(axy + a\dot{x}oy + a\dot{y}ox + a\dot{x}\dot{y}o^2) - (y^3 + 3\dot{y}oy^2 + 3\dot{y}^2o^2y + \dot{y}^3o^3) = 0.$$

Now by hypothesis $x^3 - ax^2 + axy - y^3 = 0$, and when these terms are erased and the rest divided by o there will remain

$$3\dot{x}x^2 + 3\dot{x}^2ox + \dot{x}^3o^2 - 2a\dot{x}x - a\dot{x}^2o + a\dot{x}y + a\dot{y}x + a\dot{x}\dot{y}o - 3\dot{y}y^2$$
$$-3\dot{y}^2oy - \dot{y}^3o^2 = 0.$$

But further, since o is supposed to be infinitely small so that it be able to express the moments of quantities, terms which have it as a factor will be equivalent to nothing
5 in respect of the others. I therefore cast them out and there remains $3\dot{x}x^2 - 2a\dot{x}x + a\dot{x}y + a\dot{y}x - 3\dot{y}y^2 = 0$, as in Example 1 above.

It is accordingly to be observed that terms not multiplied by o will always vanish, as also those multiplied by o of more than one dimension; and that the remaining terms after division by o will always take on the form they should have according to the rule. This is what I wanted to show.[1]

Notes

1 The little 'o' which we saw as a general increment in the *De Analysi* has now become an 'infinitely small period of time', say δt.

2 All variables are *fluent* quantities and their *moments* are correspondingly expressed by the products of their respective velocities and the time 'o'. We can think of $\dot{x}o, \dot{y}o$, as $(\mathrm{d}x/\mathrm{d}t)\delta t$, $(\mathrm{d}y/\mathrm{d}t)\delta t, \ldots$

3 If $f(x, y) = 0$ expresses a relationship between x and y which is valid at all *times*, then

$$f(x, y) = f(x + \dot{x}o, y + \dot{y}o)$$
$$= f(x + (\mathrm{d}x/\mathrm{d}t)\delta t, y + (\mathrm{d}y/\mathrm{d}t)\delta t)$$

4
$$x^3 - ax^2 + axy - y^3 = 0$$
$$= (x + \dot{x}o)^3 - a(x + \dot{x}o)^2 + a(x + \dot{x}o)(y + \dot{y}o) - (y + \dot{y}o)^3$$

The steps followed are, successively: (i) expand, (ii) remove common terms from both sides, (iii) divide by o, (iv) delete terms containing o, 'since o is supposed to be infinitely small'.

5 The relation, $3\dot{x}x^2 - 2a\dot{x}x + a\dot{x}y + a\dot{y}x - 3\dot{y}y^2 = 0$, can be rewritten in the form, $\dot{y}/\dot{x} = \dfrac{3x^2 - 2ax + ay}{3y^2 - ax} = \dfrac{\mathrm{d}y}{\mathrm{d}x} = -f_x/f_y$, where f_x and f_y are the partial derivatives of $f(x, y)$ with respect to x and y respectively. (See M100[2], Unit 15.)

1 *NMP*, III, pp. 79–81.

2 The Open University (1971) M100 *Mathematics: A Foundation Course*, The Open University Press.

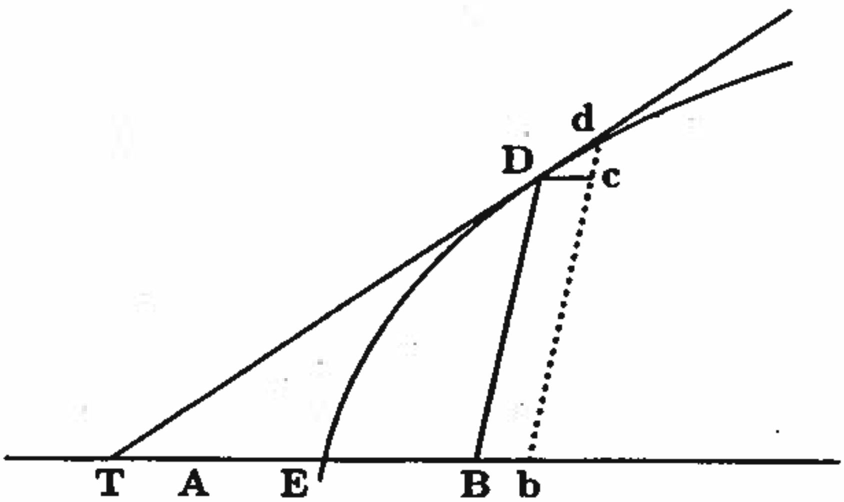
Text 18: Newton on the method of drawing tangents. From D. T. Whiteside, ed. (1964). *The Mathematical Works of Isaac Newton*. Vol. 1. Johnson Reprint Corp.

PROBLEM IV.

To draw Tangents to Curves.

The First manner.

Tangents may be variously drawn according to the various relation of curves to right lines; and first, let BD be a right line or ordinate in a given angle to another right line AB, as a base or absciss, and terminated at the curve ED; let this ordinate move thro' an indefinite small space to the place *bd*, so that it may be increased by the moment *cd*, while AB is increased by the moment B*b* to which D*c* is equal and parallel, let D*d* be produced till it meet with AB in T, and this line will touch the curve in D or *d*, and the triangles *dc*D, DBT will be similar; so that TB : BD :: D*c*, or B*b* : *cd*. Since therefore the relation of BD to AB is exhibited by the equation by which the nature of the curve is determined, seek for the relation of the Fluxion by PROB. I. Then take TB to BD in the ratio of the Fluxion of AB to the Fluxion of BD, and TD will touch the curve in the point D.



EXAMPLE I. Calling AB $= x$ and BD $= y$, let their relations be $x^3 - ax^2 + axy - y^3 = 0$, and the relation of the Fluxion will be $3\dot{x}x^2 - 2a\dot{x}x + a\dot{x}y - 3\dot{y}y^2 + a\dot{y}x = 0$, so that $\dot{y} : \dot{x} :: 3xx - 2ax + axy + ay : 3y^2 - ax ::$ BD or (y) : BT. Therefore

$$\mathrm{BT} = \frac{3y^3 - axy}{3x^2 - 2ax + ay};$$

therefore the point D being given, and thence DB and AB, or y and x, the length will be given by which the tangent TD is determined.

Text 19: Leibniz' process of discovery. From M. E. Baron and H. J. M. Bos, eds. (1974). *Newton and Leibniz*. History of Mathematics: Origins and Development of the Calculus 3. The Open University Press, pp. 42–43.

C3.13 GLIMPSES OF A PROCESS OF DISCOVERY

We have now mentioned the three important ideas which underlie Leibniz's invention of the calculus:

1 Leibniz's interest in symbolism and notation in connection with his idea of a general symbolic language;

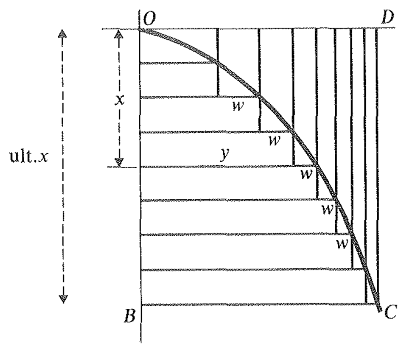
2 The insight that summing of sequences and taking their differences are inverse operations, and that similarly determining quadratures and tangents are inverse operations;

3 The characteristic triangle and its use in deriving general transformations of quadratures (like the transmutation).

In the period 25 October–11 November 1675, Leibniz combined these ideas in a series of studies on the analytic treatment of infinitesimal problems, which contain the invention of the calculus. They are known to us because the manuscripts in which Leibniz jotted down his thoughts, more or less as they came to him, are still extant. These manuscripts, dated 25, 26, 29 October and 1 and 11 November 1675, form a most precious record of a process of invention. It is not often that we are able to follow the successive steps in a major mathematical discovery, and in this section we will indicate these steps and illustrate them by fragments of the original texts.

Leibniz's starting point was the study of relations between quadratures, expressed analytically (in formulae) by means of the symbolism introduced by Cavalieri (see Unit C2 pp. 13–8). That is, he wrote 'omn.*l*' (abbreviation for *omnes l*, 'all *l*'), for the quadrature of a curve whose ordinates are *l*.

To give you the flavour of this starting point of Leibniz's study, here is an argument from the manuscript of 26 October. The text is very brief, it consists only of the sentences we quote[1] and a series of formulae, so we have added some explanation.



Consider a sequence of equidistant ordinates y of a curve as in the figure (which is an amplification of Leibniz's figure). The differences of the ys are called w. The area OCD is the sum of all rectangles xw. Now $x \times w$ is the statical moment of w with respect to the horizontal axis. (Statical moment = weight × distance to axis; in this case the weight of w is taken equal to its length.) Therefore area OCD is the sum of the moments of the differences w. Now area OCD is the complement of area OCB in the rectangle $OBCD$, and the area OCB is, in Cavalieri's terminology, the sum of all 'terms' y. Hence:

> The moments of the differences about a straight line perpendicular to the axis are equal to the complement of the sum of the terms.

Now the ws are the differences of the ys, so that conversely the 'terms' y are the sums of the w. So if we take any sequence with terms w and replace in the preceding sentence 'differences' by 'terms' and 'terms' by 'sum of the terms' we have:

> and the moments of the terms are equal to the complement of the sum of the sums.

Leibniz expresses this result in Cavalierian symbolism:

$$\underbrace{\text{omn.}\overline{xw}}_{\text{moments of the terms } w} \;\sqcap\; \underbrace{\underbrace{\text{ult.}x,\ \overline{\text{omn.}w.},}_{\text{total}} - \underbrace{\overline{\text{omn.omn.}w}}_{\text{sum of the sums of the terms}}}_{\text{complement of the sum of the sums of the terms}}$$

1 Child, J. M. (1920) *The Early Mathematical Manuscripts of Leibniz*, London.

Text 19: Leibniz' process of discovery. From M. E. Baron and H. J. M. Bos, eds. (1974). *Newton and Leibniz*. History of Mathematics: Origins and Development of the Calculus 3. The Open University Press, pp. 42–43.

$\sqcap$ is Leibniz's symbol for equality; he uses overlining where we would use brackets; the commas are separating symbols; ult. stands for *ultimus* (last), meaning the last terms of the sequence. You should note the central role of the theory of difference sequences in this argument, see Section C3. 12 p. 36.

Now Leibniz plays with this formula, and derives other formulae from it purely analytically, without making use of a figure. He does this by substituting special variables in the place of w, and he interprets the results as relations between quadratures. In this way he finds for instance:

$$\text{omn.}az \sqcap \text{ult.}x, \text{omn.}\frac{az}{x} - \text{omn.omn}\frac{az}{x}$$

$\left(\text{by substitution } xw = az, w = \dfrac{az}{x}\right)$ and

$$\text{“omn.}a \sqcap \text{ult.}x, \text{omn.}\frac{a}{x} - \text{omn.omn.}\frac{a}{x}$$

$\left(\text{by substitution } xw = a, w = \dfrac{a}{x}\right)$

Leibniz interprets the last equation:

> the last theorem expresses the sum of the logarithms in terms of the known quadrature of the hyperbola.

$y = \dfrac{a}{x}$ is the equation of the rectangular hyperbola, hence omn.$\dfrac{a}{x}$ is the quadrature of the hyperbola. Now this quadrature is a logarithm (we would say $\int \dfrac{a\,\mathrm{d}x}{x} = \log x$ for some base for the logarithm), so omn.omn.$\dfrac{a}{x}$ is the sum of the logarithms. So the equation indeed expresses the sum of the logarithms in terms of the quadrature of the hyperbola.

You should compare this way of deriving transformations of quadratures with Leibniz's study on the transmutation, and note the advantage of a symbolism through which these transformations can be performed by means of formulae instead of by inspection of complicated figures.

Exercise 19

Leibniz also derived from his basic formula the relation

$$\text{omn.}\frac{a}{x} \sqcap x, \overline{\text{omn.}\frac{a}{x^2}} - \text{omn.}\overline{\text{omn.}\frac{a}{x^2}}$$

Could you imagine how?

SA 19

By using the substitution $w = \dfrac{a}{x^2}$.

Three days later (29 October) we find Leibniz exploring the operational rules for the symbol omn., noting for instance that omn.yz is not equal to omn.y × omn.z. In this investigation Leibniz suddenly chooses a new symbol instead of omn.:

> It will be useful to write $\int$ for omn., so that $\int l$ = omn.l, or the sum of the ls.

$\int$ is the long script s, it stands for *summa*, sum, so that the symbol is shorter and applies better to Leibniz's conception of the quadrature: the *sum* of the terms, rather than the Cavalierian 'all terms'. Leibniz writes $\int\int$ for omn.omn., he stresses that the differences between the terms are infinitely small and he writes simple quadrature relations in the new symbolism:

Text 20: Bishop Berkeley's The Analyst. From D. E. Smith (1959). *A source book in mathematics*. 2nd ed. 2 vols. New York: Dover Publications, Inc., pp. 627–634. Adopted from J. Lützen and K. Ramskov, eds. (1999). *Kilder til matematikkens historie*. 2nd ed. København: Matematisk Afdeling, Københavns Universitet, pp. 95–99.

Berkeley

A Discourse Addressed to an Infidel Mathematician

Though I am a stranger to your person, yet I am not, Sir, a stranger to the reputation you have acquired in that branch of learning which hath been your peculiar study; nor to the authority that you therefore assume in things foreign to your profession; nor to the abuse that you, and too many more of the like character, are known to make of such undue authority, to the misleading of unwary persons in matters of the highest concernment, and whereof your mathematical knowledge can by no means qualify you to be a competent judge. [...]

Whereas then it is supposed that you apprehend more distinctly, consider more closely, infer more justly, and conclude more accurately than other men, and that you are therefore less religious because more judicious, I shall claim the privilege of a Freethinker; and take the liberty to inquire into the object, principles, and method of demonstration admitted by the mathematicians of the present age, with the same freedom that you presume to treat the principles and mysteries of Religion; to the end that all men may see what right you have to lead, or what encouragement others have to follow you. [...]

The Method of Fluxions is the general key by help whereof the modern mathematicians unlock the secrets of Geometry, and consequently of Nature. And, as it is that which hath enabled them so remarkably to outgo the ancients in discovering theorems and solving problems, the exercise and application thereof is become the main if not the sole employment of all those who in this age pass for profound geometers. But whether this method be clear or obscure, consistent or repugnant, demonstrative or precarious, as I shall inquire with the utmost impartiality, so I submit my inquiry to your own judgment, and that of every candid reader. — Lines are supposed to be generated[1] by the motion of points, planes by the motion of lines, and solids by the motion of planes. And whereas quantities generated in equal times are greater or lesser according to the greater or lesser velocity wherewith they increase and are generated, a method hath been found to determine quantities from the velocities of their generating motions. And such velocities are called fluxions: and the quantities generated are called flowing quantities. These fluxions are said to be nearly as the increments of the flowing quantities, generated in the least equal particles of time; and to be accurately in the first proportion of the nascent, or in the last of the evanescent increments. Sometimes, instead of velocities, the momentaneous increments or

[1] *Introd. ad Quadraturam Curvarum.*

decrements of undetermined flowing quantities are considered, under the appellation of moments.

By moments we are not to understand finite particles. These are said not to be moments, but quantities generated from moments, which last are only the nascent principles of finite quantities. It is said that the minutest errors are not to be neglected in mathematics: that the fluxions are celerities, not proportional to the finite increments, though ever so small; but only to the moments or nascent increments, whereof the proportion alone, and not the magnitude, is considered. And of the aforesaid fluxions there be other fluxions, which fluxions of fluxions are called second fluxions. And the fluxions of these second fluxions are called third fluxions: and so on, fourth, fifth, sixth, etc., *ad infinitum*. [...] But the velocities of the velocities — the second, third, fourth, and fifth velocities, etc. — exceed, if I mistake not, all human understanding. [...]

Berkeley then discusses specific examples and different ways of finding the fluxions.

[...] But whether this method be more legitimate and conclusive that the former, I proceed now to examine; and in order thereto shall premise the following lemma: — "If, with a view to demonstrate any proposition, a certain point is supposed, by virtue of which certain other points are attained; and such supposed point be itself afterwards destroyed or rejected by a contrary supposition; in that case, all the other points attained thereby, and consequently thereupon, must also be destroyed and rejected, so as from thenceforward to be no more supposed or applied in the demonstration."[2] This is so plain as to need no proof.

Now, the other method of obtaining a rule to find the fluxion of any power is as follows. Let the quantity x flow uniformly, and be it proposed to find the fluxion of x^n. In the same time that x by flowing becomes $x + o$, the power x^n becomes $\overline{x+o}^n$, i.e., by the method of infinite series

$$x^n + nox^{n-1} + \frac{nn-n}{2}oox^{n-2} + \&c.,$$

and the increments

$$o \text{ and } nox^{n-1} + \frac{nn-n}{2}oox^{n-2} + \&c.$$

are one to another as

$$1 \text{ to } nx^{n-1} + \frac{nn-n}{2}ox^{n-2} + \&c.$$

Let now the increments vanish, and their last proportion will be 1 to nx^{n-1}. But it should seem that this reasoning is not fair or conclusive. For when it is said,

[2]Berkeley's *lemma* was rejected as invalid by James Jurin and some other mathematical writers. The first mathematician to acknowledge openly the validity of Berkeley's *lemma* was Robert Woodhouse in 1803.

let the increments vanish, i. e., let the increments be nothing, or let there be no increments, the former supposition that the increments were something, or that there were increments, is destroyed, and yet a consequence of that supposition, i. e., an expression got by virtue thereof, is retained. Which by the foregoing lemma, is a false way of reasoning. Certainly when we suppose the increments to vanish, we must suppose their proportions, their expressions, and everything else derived from the supposition of their existence, to vanish with them. [...]

I have no controversy about your conclusions, but only about your logic and method: how you demonstrate? what objects you are conversant with, and whether you conceive them clearly? what principles you proceed upon; how sound they may be; and how you apply them? [...]

The great author of the metod of fluxions felt this difficulty, and therefore he gave in to those nice abstractions and geometrical metaphysics without which he saw nothing could be done on the received principles: and what in the way of demonstration he hath done with them the reader will judge. It must, indeed, be acknowledged that he used fluxions, like the scaffold of a building, as things to be laid aside or got rid of as soon as finite lines were found proportional to them. But then these finite exponents are found by the help of fluxions. Whatever therefore is got by such exponents and proportions is to be ascribed to fluxions: which must therefore be previously understood. And what are these fluxions? The velocities of evanescent increments. And what are these same evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities? [...]

And, to the end that you may more clearly comprehend the force and design of the foregoing remarks, and pursue them still farther in your own meditations, I shall subjoin the following Queries: —

[...]

Qu. 4. Whether men may properly be said to proceed in a scientific method, without clearly conceiving the object they are conversant about, the end proposed, and the method by which it is pursued? [...]

Qu. 8. Whether the notions of absolute time, absolute place, and absolute motion be not most abstractely metaphysical? Whether it be possible for us to measure, compute, or know them?

[...]

Qu. 16. Whether certain maxims do not pass current among analysts which are shocking to good sense? And whether the common assumption, that a finite quantity divided by nothing is infinite, be not of this number?[3] [...]

Qu. 31. Where there are no increments, whether there can be any *ratio* of in-

[3]The earliest exclusion of division by zero in ordinary elementary algebra, on the ground of its being a procedure that is inadmissible according to reasoning based on the fundamental assumptions of this algebra, was made in 1828, by Martin Ohm, in his *Versuch eines vollkommen consequenten Systems der Mathematik*, Vol. I, p. 112. In 1872, Robert Grassmann took the same position. But not until about 1881 was the necessity of excluding division by zero explained in elementary school books on algebra.

crements? Whether nothings can be considered as proportional to real quantities? Or whether to talk of their proportions be not to talk nonsense? [...]

Qu. 63. Whether such mathematician as cry out against mysteries have ever examined their own principles?

Qu. 64. Whether mathematicians, who are so delicate in religious points, are strictly scrupulous in their own science? Whether they do not submit to authority, take things upon trust, and believe points inconceivable? Whether they have not *their* mysteries, and what is more, their repugnances and contradictions? [...]

Text 21: J. Lützen (2003). "The Foundation of Analysis in the 19th Century". In: *A History of Analysis*. Ed. by H. N. Jahnke. History of Mathematics 24. Providence (Rhode Island): American Mathematical Society. Chap. 6, pp. 155–195.

CHAPTER 6

The Foundation of Analysis in the 19th Century

JESPER LÜTZEN

6.1. Introduction

The 19th century has often been called the age of rigour. This is a correct characterization in the sense that analysis was given a foundation that we still recognize as satisfactory. The rigourisation was not just a question of clarifying a few basic concepts and changing the proofs of a few basic theorems; rather it invaded almost every part of analysis and changed it into the discipline we now learn in high schools and at universities. The movement towards rigour can even be seen as a process of creation. It produced whole new areas of mathematics, in particular the important point set topological underpinnings of analysis dealing with entirely new concepts such as pointwise and uniform continuity (and convergence), compactness, completeness, etc.

However, it would be wrong to assume that in the 19th century the problem of rigour was considered to be the most pressing question in analysis. The great majority of mathematicians, and even those we will encounter in this chapter, worked mainly on technical questions extending and applying the analytical theories they inherited from their predecessors. In fact, the developments of new technical theorems of mathematics provided one of the main backgrounds for the growing interest in foundational questions. Fourier series were particularly important in this respect since they challenged the old ideas about the concepts of function, integral, convergence, continuity, etc., but differential equations, potential theory, elliptic functions and other areas also contributed to the rigourisation process.

Teaching was another major main motivation behind the rigourisation of analysis. Several mathematicians found themselves in an awkward situation when they had to teach the introduction to analysis, and therefore they decided to reform it. This was the direct background for Cauchy's and Weierstrass's reforms and of Dedekind's and Méray's construction of the real numbers. A process of emancipation of mathematics from science added to the feeling that the foundations of analysis had to be revised. In the 18th century, and even in the beginning of the 19th century in France, analysis was closely connected to theoretical physics. This meant that the correctness of the rules of analysis could be corroborated by their success in applications; more specifically the existence of, e.g., solutions of differential equations or sums of series was inferred from the physical situation. However, during the first half of the 19th century, in particular in Germany, high schools and universities rather than technical high schools became the centres of mathematical training and research. Combined with the neo-humanist movement, this led to the development of pure mathematics as an independent field (cf. (Jahnke 1990)). It became important to give mathematics, including analysis, a solid foundation of its own, independent of applications.

At the same time analysis separated itself from geometry. Since Euclid, geometry had been considered the best-founded part of mathematics and although the concept of number had been enlarged to include irrational and transcendental numbers, most mathematicians sought the justification of the enlarged concept of number in Euclid's theory of magnitudes. This general picture changed in the course of the 19th century. Many gaps were discovered in Euclid's arguments and alternative geometries were constructed so that Euclid's authority was questioned. It was pointed out that basic theorems of analysis that had hitherto been based on geometric intuition needed a firmer basis. In particular several mathematicians sought to prove the intermediate value theorem which states that a continuous function that attains both positive and negative values on an interval will attain the value zero.

In Chapter 4 we saw that already in the 18th century several mathematicians tried to base analysis on algebra instead of geometry. This approach was largely rejected in the 19th century. Instead the natural numbers and arithmetic offered the firm ground, and around 1870 "arithmetization" became a slogan. The real (and complex) numbers were constructed from the rationals which were in turn constructed from the natural numbers (see Chapter 10), and analysis was based directly on this new construct bypassing geometry entirely. Even though Pasch, Peano, Pieri and Hilbert gave a firm axiomatic foundation of geometry at about the same time, it never regained its role as the basis of analysis. On the contrary Hilbert proved that geometry is consistent if arithmetic is consistent.

One can divide the rigourisation of analysis into two periods: a French period dominated by Cauchy and a German period dominated by Weierstrass. This reflects the generally accepted picture of 19th-century mathematics according to which France was the leading mathematical nation till around the middle of the century after which Germany took the lead.

6.2. The concept of function

Since Euler, calculus had been a theory of functions. But what is a function? The meaning of this concept changed over time. As we saw in Chapter 4, Euler presented two definitions: in the *Introductio* a function was defined as an analytic expression (i.e., a formula) containing constants and variables, but in the *Institutiones Calculi Differentialis* it was defined as a variable depending on another variable. We find the same double vision in Lagrange's textbooks. However, in Cauchy's *Cours d'analyse*, which was the first textbook to herald the new era of rigour, functions were defined exclusively as variables depending on other variables:

> If variable quantities are so joined between themselves that, the value of one of these being given, one can conclude the values of all the others, one ordinarily conceives these diverse quantities expressed by means of the one among them, which then takes the name independent variable; and the other quantities expressed by means of the independent variable are those which one calls functions of this variable. (Cauchy 1821), (Cauchy 1831), Transl. (Rüthing 1984, 74)

One year later Fourier distanced himself even more explicitly from those who saw functions as analytic expressions. In his main work *Théorie analytique de la Chaleur* he wrote:

In general, the function $f(x)$ represents a succession of values or ordinates each of which is arbitrary. An infinity of values being given to the abscissa x, there are an equal number of ordinates $f(x)$. All have actual numerical values, either positive or negative or nul. We do not suppose these ordinates to be subject to a common law; they succeed each other in any manner whatever, and each of them is given as it were a single quantity. (Fourier 1822, 500), Transl. (Rüthing 1984, 74)

Dirichlet accepted this definition in his paper on Fourier series and defined a continuous function as follows:

Let us suppose that a and b are two definite values and x is a variable quantity which is to assume, gradually, all values located between a and b. Now, if to each x there corresponds a unique, finite y in such a way that, as x continuously passes through the interval from a to b, $y = f(x)$ varies likewise gradually, then y is called a continuous function of x for this interval. (Dirichlet 1837, 135), Transl. (Rüthing 1984, 74)

This definition, which emphasizes the one valuedness of $f(x)$ was taken over almost verbatim by Riemann (Riemann 1851). If we only look at the definitions, it looks as if the concept of function defined as a general dependence between variables goes back to Euler and was used rather systematically after 1820. For this reason Youschkevich (Youschkevich 1976) has named this concept Euler's concept of function.

However, in the history of mathematics and science it is often insufficient to consider how concepts are defined; one needs also to consider how they are used. This will often lead to a different and more complex story.

Immediately following his definition of a function in the *Cours d'analyse* Cauchy divided functions into various classes, the first being the explicit functions of which he mentioned $\log x$, $\sin x$, $x + y$, x^y, xyz, etc., as examples.

However, when only the relations between the functions and the variables are given, i.e. the equations that these quantities must satisfy, such that these equations are not solved algebraically, the functions, that are not immediately expressed in terms of the variables are called *implicit functions*. (Cauchy 1821, 32).

Here Cauchy clearly implied that functions are either explicit or implicit, i.e., that they are always given through some equation or expression. His division into simple and composit functions leaves the same impression. Moreover, when in Chapter 8 Cauchy passed from real to complex functions, he simply remarked:

When the constants or variables contained in a function, after having first been considered real are then assumed to be imaginary, the notation that has been used to express the said function cannot be maintained in the calculations except through new conventions that can fix the sense of the notation under this latter hypothesis.

I shall return to this quotation below. Here it suffices to note that in order to be able to talk about the "constants" and "variables" in an (arbitrary) function, Cauchy must clearly have had analytic expressions in mind. Still, contrary to Euler's and Lagrange's arguments, Cauchy's proofs and other concepts (in particular continuity) did not rely on the concept of an analytic expression.

Fourier, on the other hand, consciously avoided implying that functions were analytical expressions. Yet in his "proof" of the convergence of Fourier series of an "arbitrary" function f, he explicitly used the fact that when "α differs infinitely little from x, the value of $f\alpha$ coincides with fx" (Fourier 1822, §423); i.e., he assumed that any function is continuous in the modern sense. Cauchy and Fourier were not exceptions. It was quite usual for early 19th-century mathematicians to define functions in a general way and then implicitly or explicitly ascribe various additional properties to them in the course of the arguments. Much of the movement of rigour consisted precisely in a growing awareness that one can only use properties of the functions which have been stated explicitly.

The first to live up to this ideal was Dirichlet, who in his papers on Fourier series formulated his convergence result for the piecewise continuous, piecewise monotonous functions, and he used only these assumptions in the proof. For this reason one often follows (Hankel 1870) and names the modern concept of function after Dirichlet.

Dirichlet's concept of function only gradually replaced Euler's analytical expressions in the textbooks. As late as 1870 Hankel remarked:

> One person defines functions essentially in Euler's sense, the other requires that y must change with x according to a law, without giving an explanation of this obscure concept, the third defines it in Dirichlet's manner, the fourth does not define it at all. However, everybody deduces from his concept conclusions that are not contained in it. (Hankel 1870, 67)

Cauchy's definitions of the basic concepts of analysis

Variables and constants

> A variable quantity is a quantity that one considers as being able to receive successively several different values .... On the contrary a constant quantity is a quantity that receives one fixed and determined value .... (Cauchy 1821, 19)

Limit

> When the values successively attributed to the same variable approach a fixed value indefinitely, in such a way as to end up by differing from it as little as one could wish, this last value is called the *limit* of all the others. (Cauchy 1821, 19), Transl. (Fauvel; Gray 1987, 566)

Infinitely small quantity

> When the successive numerial values of the same variable decrease indefinitely in such a way as to fall below any given number, this variable becomes what one calls an *infinitesimal* or an *infinitely small* quantity. A variable of this kind has zero for its limit. (Cauchy 1821, 19), Transl. (Fauvel; Gray 1987, 566)

Continuity

> Let $f(x)$ be a function of the variable x and suppose that for each value of x between two given limits this function always takes a unique and finite value. If, having a value of x between these limits, one attributes to the variable x an infinitely small increase α, the function itself increases by the difference
>
> $$f(x+\alpha) - f(x)\,, \tag{6.1}$$
>
> which depends simultaneously on the new variable α and the value of x. This done, the function $f(x)$ will be, between the two limits assigned to the variable x, a *continuous* function of this variable if, for each value of x intermediate between

these limits the numerical value of the difference

$$f(x+\alpha)-f(x) \tag{6.2}$$

decreases indefinitely with α. In other words, *the function $f(x)$ will remain continuous with respect to x between the given limits if, between these limits an infinitely small increase in the variable always produces an infinitely small increase in the function itself.* One says furthermore that the function $f(x)$ is, in the neighbourhood of a particular value attributed to x, a continuous function of this variable, whenever it is continuous between two limits of x, however close, which contain that value of x. (Cauchy 1821, 43), Transl. (Fauvel; Gray 1987, 566-567)

Convergence

A series is an indefinite sequence of quantities

$$u_0,\ u_1,\ u_2,\ u_3,\ \text{etc......} \tag{6.3}$$

which succeed each other according to a fixed law. These quantities themselves are the different *terms* of the series considered. Let

$$s_n = u_0 + u_1 + u_2 + \cdots + u_{n-1} \tag{6.4}$$

be the sum of the first n terms, where n is an arbitrary integer. If the sum s_n approaches a certain limit S indefinitely for increasing values of n, then the series is said to be *convergent*, and the limit in question is called the sum of the series. On the contrary, if the sum s_n approaches no fixed limit when n increases indefinitely, the series is divergent and has no sum. (Cauchy 1821, 114)

Derivative

When the function $y = f(x)$ is continuous between two given limits of the variable x, and one assigns a value between these limits to the variable, an infinitely small increment of the variable produces an infinitely small increment in the function itself. Consequently, if we then set $\Delta x = i$, the two terms of the *difference quotient*

$$\frac{\Delta y}{\Delta x} = \frac{f(x+i)-f(x)}{i} \tag{6.5}$$

will be infinitesimals. But while these terms tend to zero simultaneously, the ratio itself may converge to another limit, either positive or negative. This limit, when it exists, has a definite value for each particular value of x; but it varies with x. $\cdots$ The form of the new function which serves as the limit of the ratio $(f(x+i)-f(x))/i$ will depend upon the form of the given function $y = f(x)$. In order to indicate this dependence, we give to the new function the name derivative and we designate it, using a prime, by the notation y' or $f'(x)$. (Cauchy 1823, 22).

Integral

Suppose that the function $y = f(x)$ is continuous with respect to the variable x between the two finite limits $x = x_0$, $x = X$. We designate by $x_1, x_2, \cdots, x_{n-1}$ new values of x placed between these limits and suppose that they either always increase or always decrease between the first limit and the second. We can use these values to divide the difference $X - x_0$ into elements

$$x_1 - x_0,\ x_2 - x_1,\ x_3 - x_2, \ldots, X - x_{n-1}\,, \tag{6.6}$$

which all have the same sign. Once this has been done, let us multiply each element by the value of $f(x)$ corresponding to the left-hand end point of that element: that is, the element $x_1 - x_0$ will be multiplied by $f(x_0)$, the element $x_2 - x_1$ by $f(x_1)$, $\cdots$, and finally the element $X - x_{n-1}$ by $f(x_{n-1})$; and let

$$S = (x_1 - x_0)f(x_0) + (x_2 - x_1)f(x_1) + \cdots + (X - x_{n-1})f(x_{n-1}) \tag{6.7}$$

be the sum of the products so obtained. The quantity S clearly will depend upon 1st: the number n of elements into which we have divided the difference $X - x_0$; 2nd: the values of these elements and therefore the mode of division adopted.

It is important to observe that if the numerical values of these elements become very small and the number n every large, the mode of division will have only an insignificant effect on the value of S. This in fact can be proved as follows. ⋯ Thus, when the elements of the difference $X - x_0$ become infinitely small, the mode of division has only an imperceptible effect on the value of S; and, if we let the numerical values of these elements decrease while their number increases, the value of S ultimately becomes, for all practical purposes [*sensiblement*], constant. Or, in other words, it ultimately reaches a certain limit that depends uniquely on the form of the function $f(x)$ and on the bounding values x_0, X of the variable x. This limit is what is called a *definite integral*. (Cauchy 1823, 122-124), Transl. (Grabiner 1981, 171 and 174).

6.3. Cauchy and the *Cours d'analyse*

Augustin-Louis Cauchy was educated as an engineer at the École Polytechnique and the École des Ponts et Chaussées in Paris but only worked in this profession for a few years. After the restoration of the monarchy in 1815 he began to teach analysis at the École Polytechnique, and the following year he became a member of the Académie des Sciences. He was a staunch supporter of the Bourbon monarchy and a conservative catholic, so after the 1830 revolution he went into voluntary exile first in Torino and then in Prague where he taught mathematics to the son of the dethroned Charles X. In 1838 he returned to Paris but he had to wait for the next revolution (1848) to get a new teaching position, now at the Faculté des Sciences. Cauchy wrote five textbooks and more than 800 papers and is, therefore, next to Euler, the most productive mathematician who ever lived. He made important contributions to such diverse areas as complex function theory, algebra (permutations), the theory of errors, celestial mechanics, and mathematical physics, particularly the theory of elasticity and optics (Belhoste 1991).

He made his contributions to the foundations of analysis in connection with his fifteen years of teaching at the École Polytechnique. At the beginning of this period the curriculum demanded that before teaching differential and integral calculus, the teacher should present the so-called "algebraic analysis" corresponding more or less to volume one of Euler's *Introductio*. Cauchy published his version of this part of the course in 1821. It bore the title *Cours d'Analyse de l'ecole royale polytechnique. Première partie. Analyse algébrique.* The year after it was published, the curriculum underwent great changes and this part of the *Cours d'analyse* was greatly reduced, and eventually disappeared altogether. Therefore, instead of publishing a second part of his *Cours d'analyse*, Cauchy wrote *Résumé des leçons données a l'Ecole royale polytechnique sur le calcul infinitésimal, tome premier* (1823) and finally in 1829 *Leçons sur le calcul différentiel* ((Cauchy 1829) and (Gilain 1989)).

As a teacher Cauchy was a limited success. Most of his students did not appreciate his very theoretical style and even his colleagues and superiors often criticised him for going into too great a length when dealing with the nitty-gritty details of the introductory parts of the course at the expense of the more applicable parts. Yet it is precisely this insistance on the foundations that through his textbooks has made Cauchy famous as the initiator of the movement toward rigour.

This may seem strange to a modern reader looking at Cauchy's definitions of the basic concepts of calculus (see "Cauchy's definitions ⋯" above). To the modern eye Cauchy's definitions may seem wordy, vague and not particularly rigorous. We miss our quantifiers, our ε's, δ's and N's, and in most cases our inequalities. However, as has been pointed out in particular by Grabiner (Grabiner 1981) all these ingredients are clearly present when Cauchy starts using his concepts in proofs. In particular the complicated proofs are strikingly modern in appearance. We get a better understanding of the meaning of the basic concepts when we study how they are used. As Grabiner (Grabiner 1981), Dhombres (Dhombres 1992), Bottazzini (Bottazzini 1990) and others have pointed out, it is the entire architecture of Cauchy's calculus rather than its separate elements that makes it so different from its predecessors. For example, Euler had defined Euler continuous functions as the functions that were given by one analytic expression and had implied that calculus was applicable to such functions but perhaps not so clearly applicable to Euler discontinuous functions. However, it is hard to point to a specific place in Euler's textbooks where this precise distinction is central in a proof. Cauchy, on the other hand, defined a new concept of continuity that was highly operational in the sense that it intervened in a precise way in several proofs, e.g., the existence of the integral and the solution of functional equations.

Yet, even when read in context it is difficult to rescue Cauchy completely. There remains a certain indeterminateness in several of his definitions and certain problems in some of his central proofs that were only resolved by the subsequent development of his ideas. We shall now discuss Cauchy's central concepts one after the other, commenting on their novelty and investigating their origin. We will discuss how they fit into the overall structure of Cauchy's calculus, comparing them with modern concepts and pinpointing unclear points and problematic arguments.

Grattan-Guinness (Grattan-Guinness 1970a) has argued that Cauchy "stole" several ideas from Bolzano's paper of 1817 (about which see below). This claim has been rejected by several historians (see, e.g., (Freudenthal 1971)) and as we shall see, there is no reason to believe that Cauchy owed any debt to Bolzano. There are enough similarities between certain points in the works of Euler, Lagrange, Lacroix, Poisson and the young Cauchy to suggest natural roots of Cauchy's concepts, theorems and proofs (see in particular (Grabiner 1981) and (Bottazzini 1986)).

Before we start the detailed analysis of Cauchy's concepts, it is worth pointing out that the title "Analyse Algébrique" of Cauchy's first book may give the wrong impression of Cauchy's approach. In the introduction Cauchy himself characterized his new methods by distancing himself from the algebraic metaphysics underlying Euler's and Lagrange's introduction to calculus:

> As for methods, I have sought to give them all the rigor that one demands in geometry, in such a way as never to revert to reasoning drawn from the generality of algebra. Reasoning of this kind, although commonly admitted, particularly in the passage from convergent to divergent series and from real quantities to imaginary expressions, can, it seems to me, only occasionally be considered as inductions suitable for presenting the truth, since they accord so little with the precision so esteemed in the mathematical sciences. We must at the same time observe that they tend to attribute an indefinite extension to algebraic formulas, whereas in reality the larger part of these formulas exist only under certain conditions and for certain values of the quantities that they contain. (Cauchy 1821, ii-iii), Transl. (Bottazzini 1986, 102).

Euler seems to have been of the opinion that every algebraic expression had a natural meaning for all complex values of the variables; it was the job of the mathematician to find these values. On the other hand, Cauchy insisted that analytic expressions only have values where we have defined them. If we want to extend analytic expressions beyond the domain where they have initially been defined, a new *definition* is needed. We saw in the previous section Cauchy emphasized this particularly in the case where we want to extend a real function to the complex domain. In the introduction to the *Cours d'analyse* he mentioned this as one of the claims that his readers might find hard to accept. Although Cauchy confused the concept of function with the concept of an analytic expression to a certain degree, he did not accept the generality of these expressions which was central to Euler's and Lagrange's metaphysics of analysis.

6.3.1. Variables and limit. Cauchy already distanced himself from Euler in his definition of a variable quantity (cf. "Cauchy's definitions···"). Euler defined it as "an undetermined or a general numerical quantity which includes all determined values without exception" (Euler 1748, 1. Kapitel §2). Cauchy's variables attain different values but not necessarily all values; i.e., they can be limited to a given interval. Another difference is that Cauchy's concept is dynamical whereas Euler's is closer to a modern concept of an arbitrary or generic element of a set. In particular Cauchy's variables can have limits. This seems odd to the modern reader who gives meaning to the compound statement

$$f(x) \to a \quad \text{for } x \to b \tag{6.8}$$

but not to the statements $f(x) \to a$ or $x \to b$ separately. However, the difference between Cauchy and modern-day conception almost vanishes when we consider how Cauchy used the limit concept. When applied to sequences s_n, it is always understood that n tends to infinity, and in other cases also (see, e.g., the definition of continuity) there are in fact always two variables in play of which one is a function of the other so that (6.8) easily captures what Cauchy had in mind. For example consider the following theorem:

2ND THEOREM. If the function $f(x)$ is positive for very large values of x and the ratio

$$\frac{f(x+1)}{f(x)}$$

converges to the limit k when x increases indefinitely, then the expression

$$[f(x)]^{\frac{1}{x}}$$

will at the same time converge to the same limit. (Cauchy 1821, 58)

Here there is no doubt about the meaning of the terminology, in particular if we read the proof which begins:

PROOF. Let us first assume that the quantity k which is necessarily positive has a finite value and let ε denote a number, as small as one wants. Since the increasing values of x make the ratio

$$\frac{f(x+1)}{f(x)}$$

converge to the limit k, one may give the number h a value so large that for x equal to or larger than h, the said ratio is constantly enclosed between the limits

$$k - \varepsilon \, , \; k + \varepsilon \, .$$

(Cauchy 1821, 59) □

As emphasized by Grabiner (Grabiner 1981), Cauchy has here substantiated with all the quantifiers, ε's and N's, and inequalities, what his definition was meant to cover, and we see that here at least it corresponds exactly to our modern concept of limit. D'Alembert (and even Newton) had built the calculus on a concept of limits of (geometric) magnitudes or variables, but Cauchy clarified the earlier ideas and even changed them. For example, it had been customary to insist that a variable could not surpass or become equal to its limit. Cauchy discarded such unnecessary restrictions.

It should be mentioned that Cauchy's concept of limit differed from our modern concept in at least one respect; Cauchy sometimes allowed a variable (or a sequence) to have more than one limit. For example, he formulated the root-test for a series with positive terms as follows (in the notation of (6.3) and (6.4)):

1ST THEOREM. Search the limit or the limits towards which the expression $(u_n)^{\frac{1}{n}}$ converges when n increases indefinitely and denote by k the greatest of these limits, or, in other words the limit of the greatest values of the said expression. The series (6.3) will be convergent if $k < 1$, and divergent if $k > 1$. (Cauchy 1821, 121)

In the subsequent proof it becomes clear first that for every $U > k$ there exist an n_0 such that for $n \geq n_0$ the quantity $(u_n)^{1/n}$ will be smaller than U and second that for every $U < k$ there exist arbitrarily large numbers n such that $(u_n)^{1/n} > U$. Thus "limit" in this case means nothing but "condensation point" and the greatest limit is precisely what we call lim sup. In most other cases, e.g., Cauchy's definition of the sum of a series, it is understood that there can only be one limit.

6.3.2. Infinitely small quantities. Cauchy writes that a variable having 0 as its limit *becomes* infinitely small. This of course leaves open what an infinitely small quantity (an infinitesimal hereafter) *is*. However, in the sequel Cauchy simply assumes that a variable tending to 0 *is* an infinitesimal; for example,

> Let α be an infinitely small quantity, that is a variable whose numerical value decreases indefinitely. (Cauchy 1821, 38)

Infinitesimals were used by Cauchy in several places in his *Cours d'Analyse* and other textbooks (e.g., in the definition of continuity) as well as in his research papers; their role in Cauchy's rigorous calculus has been discussed recently. The standard interpretation, also adopted by Grabiner, is that the limit concept is the central one and infinitesimals only enter as useful abbreviations for variables having the limit zero. Belhoste (Belhoste 1991, 65, 70) has argued that Cauchy was forced to introduce infinitesimals in his lectures because the program of the school and some of the other teachers insisted that he should use this approach to calculus instead of limits (see also (Gilain 1989)).

Laugwitz (Laugwitz 1989) has characterized this interpretation as "ridiculus". He points out that Cauchy did not otherwise give in to the pressures of his colleagues and emphasizes that Cauchy in his *Avertissement* to (1823) gives the use of infinitesimals as a principal aim:

> It has been my principal aim to try and reconcile the rigour that I have followed as a rule in my *Cours d'analyse*, with the simplicity that follows from the direct consideration of infinitely small quantities.

Laugwitz as well as Robinson (Robinson 1967) and to a certain degree Lakatos (Lakatos 1978) offer a nonstandard reading of Cauchy. Nonstandard analysis is a recent theory of infinitesimals created by Robinson and in a different form by Laugwitz and Schmieden. In this reading, infinitesimals (rather than limits) are fundamental concepts. In particular Laugwitz and Robinson claim that Cauchy's variables not only run through all values that correspond to our modern real numbers, but also through infinitesimals as well as sums of real numbers and infinitesimals. We can save many of Cauchy's problematic theorems and proofs in this way. I find such a revaluation of Cauchy interesting because it highlights how historians of mathematics unconsciously read modern post-Weierstrassian ideas into Cauchy's work. However, I am not convinced that it is better to read nonstandard analytic ideas into Cauchy. In fact, when Cauchy defined variable quantities as ranging over several "values", he had not yet defined the infinitesimals, only numbers that measure "magnitudes" and "quantities", that is, just numbers preceded by a + or −. If the nonstandard reading of Cauchy is correct, "magnitudes" should be ordered in a non-Archimedian way and this clashes with the Euclidean theory. Moreover, it is hard to explain why infinitesimals are later defined *afterwards* as variable quantities tending to zero.

Cauchy's acceptance of infinitesimals was not an obvious choice for a rigourist. From about 1780 these quantities had been generally rejected as unrigorous. For example, Lacroix in his *Traité élementaire* (Lacroix 1820) had used limits but not infinitesimals. Cauchy acknowledged the simplicity of the infinitesimals but redefined them. Euler's and Leibniz's infinitesimals were constant, but Cauchy chose to define them as variables of a specific kind. He might have obtained the idea to this step from Carnot (Carnot 1797) (see (Laugwitz 1989, 205)). Moreover, Cauchy downplayed the concept of infinitesimals in one fundamental way, namely, in his definition of differentials. Leibniz and Euler considered them as infinitely small quantities whose ratio corresponded to what Lagrange called the derived function. Cauchy insisted that differentials were finite quantities (see (Cauchy 1831) quoted in (Laugwitz 1989, 205)). We shall return to his definition below.

6.3.3. Continuity. The most novel and probably most central concept in Cauchy's *Cours d'analyse* is the notion of continuity (cf. "Cauchy's definitions···"), which differed strikingly from the widely accepted Eulerian notion of continuity. Euler's notion was algebraic and global in nature; Cauchy's was what we could anachronistically call topological and local in nature. This step from the global to the local was in full harmony with Cauchy's rejection of the "generality of algebra". Euler's concept must have looked suspect to anyone who accepted Fourier's ideas (see 7.1.2). Indeed, the Fourier series of an Euler discontinuous function $f(x)$ such as $|x|$ (suitably continued outside $[-\pi, \pi]$) gave an analytic expression (in Euler's terminology) of the form

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\alpha)d\alpha \\ &+ \frac{1}{\pi}\sum_{n=1}^{\infty}\left(\cos nx\int_{-\pi}^{\pi} f(\alpha)\cos n\alpha d\alpha + \sin nx\int_{-\pi}^{\pi} f(\alpha)\sin n\alpha d\alpha\right) \end{aligned} \tag{6.9}$$

so that $|x|$ would be Euler continuous. Later Cauchy himself showed with other examples such as

$$f(x) = \begin{cases} x & \text{for} \quad x \geq 0 \\ -x & \text{for} \quad x \leq 0 \end{cases} = \sqrt{x^2} = \frac{2}{\pi}\int_0^{\infty} \frac{x^2 dt}{t^2 + x^2} \tag{6.10}$$

that "a simple change of notation is often enough to transform an [Euler-]continuous function into an [Euler-]discontinuous function and vice versa" (Cauchy 1844, 145). Thus, unless we accept that the continuity of a function depends on the way it is written down, we must reject Euler's concept as ambiguous. Fourier did not go that far, but Cauchy did. Where did Cauchy get his alternative definition from? Functions with jumps had been studied and given the name discontiguous by Arbogast in his rather geometric investigation of the nature of solutions of partial differential equations (Arbogast 1791). Moreover, Euler's definition of continuity, although widespread, had not been used consistently. For example, Lagrange had ascribed a property to continuous functions which resembles Cauchy's later definition. In (Lagrange 1797, §14) he tried to prove that provided h is sufficiently small the power series for $f(x+h) - f(x)$ is such that any term of the series is greater than the sum of all the later terms. To this end he argued that a certain function $h \cdot P$, which is zero for $h = 0$, is "continuous" from this point:

> ... thus it approaches the axis little by little before intersecting it and consequently approaches it by a quantity smaller than any given quantity such that one can always find an abcissa i corresponding to an ordinate less than a given quantity, and then every value smaller than i also corresponds to an ordinate smaller than the given quantity. (Lagrange 1797, 28)

Lagrange here starts with an increase in the dependent variable and asks for a corresponding increase of the independent variable. He is therefore even closer to our formulation than Cauchy was. As we saw in 6.2, Fourier also used a property similar to that of Cauchy.

In his own early investigations of definite integrals (see Chapter 8) Cauchy also discovered the importance of what he later called continuity for the validity of the fundamental theorem of calculus:

$$\int_{b'}^{b''} \varphi'(z)dz = \varphi(b'') - \varphi(b') \,. \tag{6.11}$$

After stating this equality, he remarked:

> Nevertheless this theorem is only true in the case where the function [φ] increases or decreases in a continuous way between the two given limits. But if the function suddenly jumps from one value to another when the variable increases insensibly between the limits of integration, then the difference between these two values must be subtracted from the definite integral as it is usually taken, and each of the sudden jumps that the function can make necessitates a correction of the same nature. (Cauchy 1814, Introduction)

Later in the paper he formalized this: If Z is a point where φ jumps, then

> ... denoting by ζ a very small quantity, one has
>
> $$\varphi(Z+\zeta)-\varphi(Z-\zeta)=\Delta\,, \tag{6.12}$$
>
> then the ordinary value of the integral, that is,
>
> $$\varphi(b'')-\varphi(b')\,,$$
>
> must be reduced by the quantity Δ .... (Cauchy 1814, 2. Part §3)

Thus early on Cauchy saw that rather than Euler continuity the property of having or not having jumps was of direct importance when proving theorems about functions, and he formulated an expression for the jump that anticipated his later definition of continuity.

The property described by Lagrange corresponds to continuity at a point, but Cauchy did not define this concept in his *Cours d'analyse.* Cauchy saw discontinuity as happening in a point; continuity, however, happened in an interval (possibly in a neighbourhood of a point). In this way Cauchy retained some of the intuitive and philosophical idea of continuity (indeed, it is unclear which property is continued by a function that is continuous in one point) while giving it a characterization that was crucial in several of his later proofs.

In recent years people have argued about what exactly Cauchy meant by continuity: Did he mean pointwise continuity or uniform continuity or something else? Cauchy actually gives two definitions, first one without infinitesimals and then one using infinitesimals (cf. "Cauchy's definitions ⋯"). The first definition very clearly specifies a value of the variable x and states that $f(x+\alpha)-f(x)$ tends to zero with α. This sounds suspiciously like pointwise continuity. The second formulation does not speak of a specific value of x but of the increase of the "function". This can be interpreted as uniform continuity. The definition seems ambigious. If we follow the same procedure as we used when analyzing Cauchy's concept of limit and look at how he used the concept of continuity, uniformity seems to be understood. For example, Cauchy did not write that a function such as $\frac{a}{x}$ is continuous in the interval $(0,\infty)$ (which would be false if continuity means uniform continuity). Instead he wrote that it is continuous in a neighbourhood of every point in this interval, which is indeed true even when we think of uniform continuity. Moreover, and more compellingly, he used uniform continuity in two of his proofs: (1) in the proof of the existence of the integral of a continuous function (see below) and (2) in a strong form in the proof of the following theorem:

> 1ST THEOREM. If the variables $x, y, z\cdots$ have the fixed and definite quantities $X, Y, Z\cdots$ as their respective limits and the function $f(x,y,z\ldots)$ is continuous with respect to each of its variables $x,y,z\cdots$ in the neighbourhood of the particular system of values
>
> $$x=X,\quad y=Y,\quad z=Z\cdots,$$
>
> then $f(x,y,z\ldots)$ has $f(X,Y,Z,\ldots)$ as its limit. (Cauchy 1821, 47)

For Cauchy the proof is simple. He observes that the numerical value of

$$\begin{aligned}&f(X+\alpha,Y,Z,\ldots)-f(X,Y,Z,\ldots)\\ \text{and}\quad&f(X+\alpha,Y+\beta,Z,\ldots)-f(X+\alpha,Y,Z,\ldots)\\ \text{and}\quad&f(X+\alpha,Y+\beta,Z+\gamma,\ldots)-f(X+\alpha,Y+\beta,Z,\ldots),\quad\text{etc.}\end{aligned}$$

"decrease indefinitely with the value of the variables α, β, γ" and therefore so does the numerical value of

$$f(X+\alpha, Y+\beta, Z+\gamma, \dots) - f(X, Y, Z, \dots)\,.$$

In order for this proof to work, one must assume a certain uniformity in the smallness of, e.g.,

$$f(X, Y, Z+\gamma, \dots) - f(X, Y, Z, \dots)$$

with respect to the variables $X, Y, Z, \dots$. For these reasons Giusti (Giusti 1984) argued that Cauchy defined uniform continuity. Some infinitesimal reading of Cauchy lead to the same conclusion. It is interesting to notice that Ampère, who took over many of Cauchy's concepts and methods in his teaching at the École Polytechnique, formulated Cauchy's definition in a way that unquestionably corresponded to uniform continuity (Ampere 1824, 11-12).

6.3.4. Sum of a series. During the 17th century the concept of "convergence" of series had been used in several senses, one being that the terms go to zero, another one being that the partial sums s_n (6.4) tend to a fixed limit. Euler sometimes used the second definition in his *Institutiones Calculi Differentialis* (Euler 1755, §110). Cauchy was therefore not particularly revolutionary when he picked the latter definition in his *Cours d'analyse.* What was new was his rather strict use of the $\varepsilon - N$ characterization of convergence in several of his proofs and in particular his insistence all through his textbooks that divergent series have no sums. During the preceding century mathematicians had freely operated with divergent series and Euler had even tried to formalize a definition of their sums (cf. Chapter 4). Cauchy was well aware that he would shock the mathematical community when he claimed that divergent series have no sums (see Introduction to the *Cours d'analyse* (Cauchy 1821, iv)).

This fundamental claim made it necessary for Cauchy to establish the convergence of series before trying to find their sums. To this end he proved several convergence tests. The first and fundamental one is the famous Cauchy criterion. He first made it clear that if a series converges, its n-th term must converge towards zero,

> ... but this condition is not sufficient, and it must also be true for increasing values of n that the different sums
>
> $$u_n + u_{n+1}\,,$$
> $$u_n + u_{n+1} + u_{n+2}\,,$$
> $$\text{etc.}\dots$$
>
> that is, the sums of the quantities
>
> $$u_n\,, \quad u_{n+1}\,, \quad u_{n+2}\,, \text{etc.}\dots$$
>
> taken from the first, in whatever number we wish, will always end up having numerical values that are constantly smaller than any assignable limit. Conversely, when these various conditions are satisfied, the convergence of the series is assured. (Cauchy 1821, 116)

Thus Cauchy established that a convergent series is a Cauchy series (its partial sums s_n form a Cauchy (or fundamental) sequence). However, when he came to the converse (which we would consider the deep part), he simply waved his hands. In modern treatment this converse is derived from the completeness of the real numbers (or is even taken as the definition of completeness) which must be either postulated as an axiom or obtained from a construction of the real numbers. Cauchy's work does not contain either way out and he could not have appealed to the underlying concept of magnitude because Euclid does not have axioms which ensure the completeness of his magnitudes.

This missing account of completeness is a fundamental lacuna which appears in several other places in Cauchy's analysis, in particular in his proof of the intermediate value theorem and in the proof of the existence of the integral of a continuous function.

Cauchy's treatment of convergence tests is otherwise exemplary. He does not explicitly state the comparison test but he uses it for a particular series and proves it in this case by appealing to the Cauchy test. He then proves the root test by a similar comparison with a geometric series (I quoted the beginning of the proof above). Finally he uses the theorem quoted in 6.3.1 to derive the quotient test from the root test. He also establishes other tests.

These convergence tests were not all new. For example d'Alembert used the quotient test, and the comparison test had been considered obvious. What are new are the rigorous proofs of the tests and the fundamental importance attached to them.

The most famous problem in Cauchy's version of calculus is the following theorem that connects his two concepts, convergence and continuity:

> 1ST THEOREM. When the different terms of the series (6.3) are functions of the same variable x, and continuous with respect to this variable in the neighbourhood of a particular value for which the series is convergent, then the sum s of the series is also a continuous function of x in the neighbourhood of this particular value. (Cauchy 1821, 120)

Cauchy's proof runs as follows: Represent the sum of the series s as the sum

$$s = s_n + r_n \tag{6.13}$$

where s_n is the n-th partial sum.

> This being given, consider the variation of these three functions when α increases by an infinitely small quantity α. The variation of s_n is for all possible values of n an infinitely small quantity, and the variation of r_n will become insensible together with r_n if n is given a very large value. Therefore the variation of the function s cannot be but an infinitely small quantity. (Cauchy 1821, 120)

Traditionally historians of mathematics have characterized this theorem and its proof as false because it is false if we give the terms involved their modern meaning. On the other hand Cauchy's conclusion holds true if we assume that the series converges uniformly in the neighbourhood of x. For this reason several historians have tried to rescue Cauchy by claiming that Cauchy had uniform convergence in mind (e.g., (Giusti 1984); the nonstandard reading of Cauchy has led to a similar conclusion or yet another interpretation (cf. Spalt's (Spalt 1992) complete rereading of Cauchy). In this case it is impossible to refer to Cauchy's definition, since he does not define the sum of a series of functions separately. However, it is difficult to rescue Cauchy entirely because he himself later acknowledged that his theorem "cannot be accepted without restriction" (Cauchy 1853). It is hard to escape the conclusion that Cauchy's concepts were somewhat vague at this point. Soon after its publication the problematic theorem caught the attention of several mathematicians and the resulting discussion shaped the idea of uniformity. I shall return to this development below.

6.3.5. Derivative. Cauchy used Lagrange's term "the derivative" and his notation f', but he rejected Lagrange's definition of this function based on power series. Arguing against Lagrange's definition, he first pointed out that the Taylor series of a function need not converge, and secondly, that even if it converges, it does not necessarily represent the expanded function. As an example he mentioned the function $e^{-\frac{1}{x^2}}$ to which he attributed the value 0 for $x = 0$. All its derivatives at zero are equal to zero so its Taylor series is everywhere zero. It is therefore convergent everywhere but only equal to $e^{-\frac{1}{x^2}}$ for $x = 0$ (Cauchy 1823, 229-230). For this reason Cauchy postponed the discussion of Taylor series until he could give an expression for the remainder, that is, until after he had introduced the integral.

Instead, Cauchy followed Lacroix (Lacroix 1820) and defined the derivative as the limit of the difference quotient (the reformulation in terms of infinitesimals was his own twist of the definition). He also borrowed the meaning of the differential $df(x)$ from Lacroix, although he gave it a slightly different definition as

$$\lim_{\alpha \to 0} \frac{f(x+\alpha h) - f(x)}{\alpha} = df(x) \,. \tag{6.14}$$

He easily proved that

$$df(x) = hf'(x) \,. \tag{6.15}$$

Here he considered h as a finite constant. (In modern theory we think of h as a variable so that $df(x)$ is a linear function of h.) Since dx is by definition equal to $x' \cdot h = h$, Cauchy could write

$$df(x) = f'(x)dx \tag{6.16}$$

or

$$f' = \frac{df(x)}{dx} \,. \tag{6.17}$$

Thus Cauchy could use the Leibnizian terminology but still assume that the differentials were finite quantities rather than infinitesimals.

In the definition of the derivative Cauchy started out by assuming that f is continuous, but having formed the difference quotient and its limit, he was cautious to state "if it exists". This seems to set the stage for the introduction of the concept of differentiablity. However, Cauchy did not introduce this concept and in the subsequent chapters of his book he simply assumed that f was continuous (or he assumed nothing), even though he differentiated it any number of times. It is as though Cauchy still clung to the 18th-century idea of a "safe domain" in which analysis was more or less universally valid. With Euler and d'Alembert this domain consisted of all functions; with Cauchy it consisted of the continuous functions. Similar confusion characterized Cauchy's early work on complex functions. Only around 1850 did he formulate a concept of complex differentiability (cf. Chapter 8).

6.3.6. Integral. Cauchy broke radically from his predecessors with his definition of the integral (see "Cauchy's definitions ⋯"). Leibniz had considered integrals as sums of infinitesimals but from the Bernoullis onwards it had been customary to define integration as the inverse process of differentiation. This made the indefinite integral the primary concept and had made integral calculus an appendix to differential calculus. Fourier was the first to change this picture. He realized that in order to calculate Fourier coefficients (7.12c) for arbitrary functions f he could no longer rely on differential calculus, since differentiation of nonanalytically given functions did not necessarily make sense. Therefore he focused on the definite integral $\int_a^b f(x)dx$ (putting the limits of integration at the top and bottom of the integral sign is in fact Fourier's idea) and stressed that it meant the area between the curve and the axis (Fourier 1822, §229).

Cauchy followed Fourier when he focused on the definite integral, but instead of relying on a vague notion of area, Cauchy defined the definite integral as the limit of a "left sum". This was much more precise and it allowed him to *prove* that the integral exists for a continuous function.

This proof is one of the masterpieces of Cauchy's *Calculus*: Corresponding to a division

$$a < x_1 < x_2 < \cdots < x_{n-1} < b$$

of the interval $[a, b]$, Cauchy formed the "left sum"

$$S_1 = (x_1 - a)f(a) + (x_2 - x_1)f(x_1) + \cdots + (b - x_{n-1})f(x_{n-1})\,. \tag{6.18}$$

According to a theorem in the *Cours d'analyse* (Corollary to Theorem 3 in the Préliminaires) this sum is equal to $(b-a)\cdot M$, where M is a "mean value" between the values $f(a), f(x_1), \ldots, f(x_{n-1})$. Since f is continuous, M must by the intermediate value theorem (see 6.4.2) be of the form

$$M = f(a + \theta(b - a)) \tag{6.19}$$

where $0 \leq \theta \leq 1$. Thus

$$S_1 = (b - a)f(a + \theta(b - a))\,. \tag{6.20}$$

Cauchy now subdivided the interval $[a, b]$ and went on to compare the resulting new value S_2 of the left sum with the old value S_1. According to the argument above, the contribution to S_2 stemming from intervals being inside $[a, x_1]$ can be written

$$(x_1 - a)f(a + \theta_0(x_1 - a)) \tag{6.21}$$

where $0 \leq \theta_0 \leq 1$. A similar argument applies to the contributions from the subintervals lying inside the other intervals of the first division. Thus

$$\begin{aligned} S_2 = {} & (x_1 - a)f(a + \theta_0(x_1 - a)) \\ & + (x_2 - x_1)f(x_1 + \theta_1(x_2 - x_1)) \\ & + \cdots + (b - x_{n-1})f(x_{n-1} + \theta_{n-1}(b - x_{n-1}))\,. \end{aligned} \tag{6.22}$$

Writing

$$\begin{aligned} f(a + \theta_0(x_1 - a)) &= f(a) \pm \varepsilon_0\,, \\ f(x_1 + \theta_1(x_2 - x_1)) &= f(x_1) \pm \varepsilon_1\,, \\ &\vdots \\ f(x_{n-1} + \theta_{n-1}(b - x_{n-1})) &= f(x_{n-1}) \pm \varepsilon_{n-1}\,, \end{aligned} \tag{6.23}$$

Cauchy rewrote S_2 in the form:

$$\begin{aligned} S_2 &= (x_1 - a)f(a) + (x_2 - x_1)f(x_1) + \cdots + (b - x_{n-1})f(x_{n-1}) \\ &\quad \pm \varepsilon_0(x_1 - a) \pm \varepsilon_1(x_2 - x_1) \pm \ldots \pm \varepsilon_{n-1}(b - x_{n-1}) \\ &= S_1 + (b-a)M(\pm\varepsilon_0, \pm\varepsilon_1, \ldots, \pm\varepsilon_{n-1}) , \end{aligned} \tag{6.24}$$

where $M(\pm\varepsilon_0, \pm\varepsilon_1, \ldots, \pm\varepsilon_{n-1})$ is a "mean value" between the $\pm\varepsilon$'s. Cauchy continued:

> We may add that the elements $x_1 - a, x_2 - x_1, \cdots, b - x_{n-1}$ (i.e. the lengths of the intervals of the original division) have very small values, each of the quantities $\pm\,\varepsilon_0, \pm\,\varepsilon_1, \cdots, \pm\,\varepsilon_{n-1}$ will be very close to zero, and therefore the same will be true for the sum $[(b-a)M(\pm\,\varepsilon_0, \pm\,\varepsilon_1, \cdots, \pm\,\varepsilon_{n-1})]$. Granting this, $\cdots$ we see that we would not change perceptibly the value of S that was calculated by a mode of division in which the elements of the difference $b - a$ have very small numerical values if we went to a second mode of division in which each of those elements was further subdivided into others. (Cauchy 1823, 124-125)

Cauchy then went on to compare the left sums corresponding to two arbitrary divisions of $[a, b]$ by constructing a common subdivision consisting of all the points in each of the two original divisions. From this he concluded, as quoted in "Cauchy's definitions $\cdots$", that when the divisions become finer and finer, the corresponding left sums will approach each other and therefore they reach a certain limit that is called the integral. The last step obviously requires a completeness property. It should also be emphasized that when Cauchy concluded that $M(\pm\varepsilon_0, \pm\varepsilon_1, \ldots, \pm\varepsilon_n)$ is small when the division is fine, he drew on the definition of continuity interpreted as uniform continuity.

Cauchy also proved the fundamental theorem of calculus and so tied his new concept of integral to the old one.

Why did Cauchy change the definition of the integral and where might he have obtained the idea of the new definition? As we saw in 6.3.3, Cauchy had early on showed that the value of the definite integral may differ from the difference between the values of primitive function at the end points. Moreover, his own and Poisson's work showed that complex integrals may depend on the "path of integration" (see Chapter 8). This may well have suggested to Cauchy that the antiderivative was not a sound basis for the definition of the integral. Euler and his contemporaries had already used left sums to approximate integrals, and Lacroix and Poisson had tried to prove that they converge to the integral in a suitable sense. One can find many elements of Cauchy's arguments in these papers as well as in Lagrange's proof of the fundamental theorem of calculus (see (Grabiner 1981, Chapter 6)), and it is very possible that Cauchy built on these sources. Yet Cauchy's treatment is much clearer and only assumed the explicitly stated concept of continuity while the earlier arguments had relied on (but had not explicitly stated) the existence of f' and f'' and on monotoneity of the functions. And most importantly, Cauchy changed the technique from being a numerical approximation procedure to being a definition.

This shows, as Grabiner has emphasized, that we should not seek the origins of Cauchy's rigorous calculus in the formal algebraic metaphysics of the 17th century but in the numerical procedures from this period that brought forth an "algebra of inequalities".

In the case of the integral the earlier approximation techniques led Cauchy to a definition which allowed him to prove the existence of the integral for a specific type of functions. No one seems to have asked this existence question before, nor could it have been answered with the earlier definition. Cauchy also proved general existence theorems in the theory of differential equations. Instead of asking how to integrate a special function or a special differential equation (that is, finding an analytic expression for the solution), Cauchy began the process of establishing the existence of the integral for a wide class of functions (or differential equations). He thus started an important process towards a qualitative mathematics which was carried further by Sturm-Liouville theory and by Poincaré (see Chapter 11).

Cauchy (Cauchy 1823, 140-144) defined the integral of a function f which is discontinuous at $x_1, x_2, \ldots, x_m$ but continuous in the intervals (x_0, x_1), $(x_1, x_2), \ldots, (x_m, x)$ as

$$\int_{x_0}^{x} f(x)dx = \lim_{\varepsilon\to 0}\left(\int_{x_0}^{x_1-\varepsilon\mu_1} f(x)dx + \int_{x_1+\varepsilon\nu_1}^{x_2-\varepsilon\mu_2} f(x)dx + \ldots + \int_{x_n-\varepsilon\nu_n}^{x} f(x)dx\right) \quad (6.25)$$

if this limit exists and is independent of the positive quantities μ_i, ν_i. If the limit depends on the choice of μ_i, ν_i, he defined the "principal value" as the limit (6.25) when $\mu_i = \nu_i = 1$ for all i. He dealt with integrals over unbounded intervals in a similar way by letting the upper and (or) lower limits tend to $+\infty$ and (or) $-\infty$.

This way of dealing with piecewise continuous functions should be compared with the more elegant (and stronger) way later devised by Riemann, who used Cauchy's original definition for all functions where the process converges, instead of restricting the definition to continuous functions as Cauchy had done (see Chapter 9).

6.3.7. Functional equations and the binomial theorem. In order to give a feeling for the closely knit structure of Cauchy's *Cours d'analyse*, we shall look at his proof of the binomial theorem (see also (Dhombres 1992) and (Bottazzini 1990, LXXIV))

$$(1+x)^\mu = 1 + \frac{\mu}{1}x + \frac{\mu(\mu-1)}{2!}x^2 + \ldots \quad (6.26)$$

which was one of the corner stones of the "analyse algébrique", the other one being the fundamental theorem of algebra. By the ratio test the series is convergent in the interval $(-1, 1)$ (Cauchy 1821/1885, 105). For a fixed x Cauchy denoted the sum of this series by $\varphi(\mu)$. The problem is to show that $\varphi(\mu) = (1+x)^\mu$. To this end Cauchy drew on his study of the simple functional equations such as

$$\psi(\mu)\cdot\psi(\mu') = \psi(\mu+\mu') . \quad (6.27)$$

In Chapter V of his *Cours d'analyse* he showed that the only *continuous* solution to this equation is $\psi(\mu) = A^\mu$ (with A a constant). His proof has become standard. From equation (6.27) Cauchy deduced in a straightforward way that

$$\psi(m) = \psi(1)^m \quad (6.28)$$

first for $m \in \mathbb{N}$ and then for $m \in \mathbb{Q}_+$. Since every real number is a limit of rationals and since $\psi(1)^\mu$ is continuous as a function of μ, Cauchy concluded that the equation

$$\psi(\mu) = \psi(1)^\mu \quad (6.29)$$

also holds for $\mu \in \mathbb{R}^+$, and by a simple argument it also holds for $\mu \in \mathbb{R}$. The step from $\mathbb{Q}$ to $\mathbb{R}$ is the decisive one. Here Cauchy's definition of continuity comes in as a crucial operational concept. Though functional equations of this kind had been studied by many mathematicians from the time of Euler, Cauchy was the first to treat the step from $\mathbb{Q}$ to $\mathbb{R}$ in a satisfactory manner.

Cauchy then showed (Cauchy 1821/1885, 108-109) that the right-hand side $\varphi(\mu)$ of (6.26) satisfies equation (6.27). This is a simple consequence of his famous theorem on Cauchy multiplication of series:

6TH THEOREM. Let the same things be given as in the previous theorem, [that is, let

$$\begin{cases} u_0, u_1, u_2, \ldots, u_n\,, & \text{etc.,} \\ v_0, v_1, v_2, \ldots, v_n\,, & \text{etc.,} \end{cases} \tag{6.30}$$

be two convergent series with the sums s und s', respectively.] If each of the series (6.30) remains convergent if one reduces its different terms to their numerical values, then

$$\begin{cases} u_0v_0\,, u_0v_1 + u_1v_0\,, u_0v_2 + u_1v_1 + u_2v_0\,, \ldots \\ \ldots u_0v_n + u_1v_{n-1} + \ldots + u_{n-1}v_1 + u_nv_0\,, \text{ etc.} \end{cases} \tag{6.31}$$

is a new convergent series which has $s \cdot s'$ as its sum. (Cauchy 1821, 132-133)

Since $\varphi(\mu)$ is the sum of a convergent series of continuous functions (of μ), Cauchy's theorem on the sum of a series of continuous functions states that it is continuous. In addition to an application to power series, this is the only crucial place where Cauchy used this problematic theorem. By the uniqueness theorem mentioned above, $\varphi(\mu)$ must be equal to $\varphi(1)^\mu$ or $(1+x)^\mu$. This completes Cauchy's highly original proof of the binomial theorem.

6.4. Gauss, Bolzano and Abel

6.4.1. Gauss. Two other mathematicians, Gauss and Bolzano, had arrived at ideas concerning the foundation of analysis very similar to those of Cauchy independently and even earlier. In 1850 Gauss wrote to Schumacher:

> It is characteristic of mathematics of our modern times (contrary to antiquity) that our sign language gives us a lever that reduces the most complicated arguments to a certain mechanism. In this way science has gained infinitely in richness, but as the business is usually run, it has lost equally much in beauty and solidity. How often is this lever only used mechanically, although the authorization to do so in most cases implies certain tacit assumptions. I insist that by all application of the calculus, by all applications of concepts one should remain conscious about the original conditions, and never without authorization consider the results of the mechanism as one's property. However, the usual trend is that one claims that analysis has a general character, and when someone else does not admit that the results generated in this way have been proved, one demands that he must prove that they are false. However, one can only demand this from a person who claims that the result is false, not from a person who does not accept that a result is proved, when it relies on a mechanism whose original essential conditions are not at all satisfied in the existing situation. It is often like that in the case of divergent series. Series have a clear meaning when they converge; this clear meaning vanishes with this condition of convergence, and it changes nothing essential whether one uses the word sum or value. However a letter is too short to explain everything further. – For example, consider paper money as a metaphor for the above machine. It can profitably be used to produce great works, but its use is only solid, as long as I am sure that it can at any moment be exchanged to hard cash. (Gauss 1850/1865, 434-435)

The similarity to Cauchy's introduction to the *Cours d'analyse* is striking. There is an attack on the belief in the generality of the mechanism of analysis and a particular ban on divergent series. In other letters and manuscripts Gauss also discussed the problem of extending functions outside of the domain where they are initially defined especially from $\mathbb{N}$ to $\mathbb{R}$ (in the case of the Γ function) and from $\mathbb{R}$ to $\mathbb{C}$. Though the above letter dates from 1850, the ideas expressed in it go back to Gauss's youth, i.e., long before the publication of Cauchy's text books. In a series of manuscripts from around 1800 Gauss began to analyze the "foundation of the theory of infinite series" in connection with a discussion of trigonometric series. In one of these notes (Werke X^1, 390-394) Gauss defined lim sup and lim inf for a series very accurately. This discussion of the concepts was much more rigorous than the one found later in Cauchy's *Cours d'analyse*, where, as we saw above, the "greatest limit" appeared suddenly in the root test, but where only the subsequent proof revealed its clear meaning. However, Gauss never wrote systematically on the foundations of analysis in print. He raised the question of the handling of infinite series in his dissertation (Gauss 1799) on the fundamental theorem of algebra, and in his paper (Gauss 1813) on the hypergeometric series

$$F(\alpha, \beta, \gamma, x) = 1 + \frac{\alpha\beta}{1 \cdot \gamma}x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{2!\gamma(\gamma+1)}x^2 + \dots , \tag{6.32}$$

he used the quotient test to discuss the convergence properties of this series and pointed out that it makes no sense to ask for a value of the series when it does not converge. However, as in so many other areas of mathematics, Gauss had little impact on the development of the foundation of analysis because of his reluctance to publish anything but the ripe fruits.

6.4.2. Bolzano. Bolzano also had only a limited influence on the development of analysis, but for different reasons. He was a philosopher-theologian who lived in Prague far from the main mathematical centers, and he published no new technical results in mathematics. His works and even his name remained virtually unknown for about half a century despite the fact that he dug deeper into the foundations of analysis than any of his contemporaries. His most important paper *Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwei Werthen, die ein entgegensetztes Resultat gewähren wenigstens eine reelle Wurzel der Gleichung liege* (Bolzano 1817) is devoted to the intermediate value theorem.

According to Bolzano "a function $f(x)$ varies according to the law of continuity for all values x inside or outside of certain limits just means that: if x is some such value, the difference $f(x+\omega) - f(x)$ can be made smaller than any given quantity provided that ω can be taken as small as we please" (Bolzano 1817, 162). In the applications of the concept and in the more precise definition of Bolzano's posthumously published *Funktionenlehre* (Bolzano 1930), it becomes clear that $f(x+\omega) - f(x)$ must remain smaller than the given "grösse" when ω is numerically smaller than some ω_0. This seems to be our definition of pointwise continuity.

Bolzano then introduced fundamental (or Cauchy) sequences and "proves" that they converge towards a "constant quantity". This proof is not satisfactory. What he tries to prove is first that it is not impossible to assume that the limit is "beständig" (i.e., constant) and, second, that this limit is unique and can be determined as accurately as one wishes. From this theorem he then proves the following *Lehrsatz*, which is "of the highest importance".

THEOREM. If a property M does not belong to *all* values of a variable x, but does belong to *all* values which are *less* than a certain u, then there is always a quantity U, which is the greatest of those of which it can be asserted that all smaller x have property M. (Bolzano 1817, 174)

In modern terms this states the existence of the supremum of a nonempty set which is bounded above, namely, the set of values having "Property M" (or rather the infimum of the set of values that do not have "Property M"). The proof is entirely correct; Bolzano constructed a sequence of numbers a_i for which property M holds but such that it does not hold for $a_i + \frac{D}{2^k}$ where $k(i) \to \infty$ for $i \to \infty$. He then observed that this sequence is a Cauchy sequence and thus, according to the preceding theorem, has a limit U. This is the desired "supremum". He pointed out that U need not have property M, emphasising the difference between sup and max.

Finally Bolzano proved that when f and φ are continuous in $[\alpha, \beta]$ and $f(\alpha) < \varphi(\alpha)$ whereas $f(\beta) > \varphi(\beta)$, then there exists a value $x \in (\alpha, \beta)$ for which $f(x) = \varphi(x)$. He did this by looking at property $M : f(x') < \varphi(x')$. This is a property of the kind described in the previous theorem, so there is a supremum x of the set of x''s having the property. From the definition of continuity Bolzano then proved in detail that $f(x) = \varphi(x)$. He then applied the theorem to polynomials having first shown that they must be continuous.

Cauchy also considered the intermediate value theorem in his *Cours d'analyse*. In the main text he just appealed to geometric intuition, but in a note entitled *On the numerical solution of equations* he used a numerical procedure of Lagrange to supply a "proof"; cf. (Grabiner 1981):

Let $f(x)$ be a continuous function on $[x_0, X]$ for which $f(x_0) < 0 < f(X)$, and let m be a given natural number greater than one. Consider the sequence

$$f(x_0)\,,\ f\left(x_0 + \tfrac{h}{m}\right), \ldots, f\left(X - \tfrac{h}{m}\right)\,,\ f(X)\,, \tag{6.33}$$

where $h = X - x_0$. There are necessarily two consecutive terms, say $f(x_1), f(X')$ such that $f(x_1) < 0 < f(X')$ (Cauchy but not Bolzano ignored the trivial case where one of the terms is zero). Divide the interval $[x_1, X']$ in the same way in m subintervals and pick in this sequence two consecutive terms such that $f(x_2) < 0 < f(X'')$, etc. The continuation of the process will lead to an increasing series $x_0, x_1 < x_2, < \ldots$ and an decreasing series $X > X' > X'' > \ldots$ such that any quantity of the first series is smaller than any quantity of the second and such that the difference $(X^{(n)} - x_n)$ will end up differing "as little as desired" from zero. Cauchy then observed:

> From this one can conclude that the general terms of the series $[x_0, x_1, \cdots]$ and $[X, X', X'', \cdots]$ converge to a common limit. (Cauchy 1821, 379)

Cauchy proved (in much less detail than Bolzano) that the common limit, x say, satisfies $f(x) = 0$.

It is interesting to compare Cauchy's and Bolzano's procedures.

1) Bolzano did not use infinitesimals in his definition or proofs. Cauchy did.
2) Bolzano's definition of continuity is clearer than Cauchy's and sounds more pointwise. In his *Funktionenlehre* he even remarked that continuity does not imply uniform continuity, but he never fully appreciated the importance of uniformity.

3) Both Cauchy and Bolzano relied on the completeness of the real numbers. However, Bolzano's understanding of this concept seems to have been deeper than Cauchy's. Cauchy relied on the completeness also in the formulation of the Cauchy criterion and in the definition of the integral, but he did not connect these instances. Bolzano, on the other hand, used the "Cauchy criterion" to deduce the supremum property and the intermediate value theorem. Moreover, where Cauchy's universe of quantities arose from measurement of magnitudes, Bolzano in his manuscript, later called "Theorie der rellen Zahlen" (Rychlik 1962), tried (unsuccessfully by modern standards) to base the real numbers on the concept of rational (or natural) numbers alone.
4) Both Lagrange and Ampère had tried to prove that all functions (continuous in a certain sense in Ampère's case) have derivatives except for isolated values of the variable, Bolzano in his posthumously published *Funktionenlehre* (Bolzano 1930) constructed a continuous function that he could prove not to be differentiable on a dense set (in fact his function is nowhere differentiable). Though Cauchy did not try to prove the erroneous theorem that any continuous function could be differentiated, he gave, as we have seen above, his readers the impression that it was true.

Though Bolzano exceeded his contemporaries as far as rigour in analysis is concerned, he also jumped to erroneous conclusions. For example, he tried to prove a wrong theorem about termwise differentiation of series and he argued that his non-differentiable function was continuous because it was a sum of a series of continuous functions (just as Cauchy had done) (Jarnik 1981, 55). Still, if Bolzano's works had been known when they were written, they would clearly have been influential in the development of the foundation of analysis. As it were, they were rediscovered by Hankel and H. A. Schwarz only around 1870, by which time they had nothing but historical interest.

6.4.3. Abel. A third mathematician who began the reform of the foundations of mathematics was Abel. In 1826 he wrote to his professor, Hansteen:

> I will apply all my strength to bringing more light into the vast darkness that unquestionably exists in analysis. It totally lacks any plan and system, so it is really very strange that it is studied by so many and worst of all, that it is not treated rigorously at all. There are very few theorems in the higher analysis which have been proved with convincing rigour. Everywhere one finds the unfortunate method of concluding from the special to the general, and it is very strange that after such a procedure there exist only few of the so-called paradoxes. (Abel 1902, 22)

Earlier the same year he had been more specific in a letter to his friend Holmboe:

> On the whole divergent series is a devilry, and it is a shame that one dares to base any demonstration on them. One can deduce whatever one wants when one uses them, and they have done much harm and caused many paradoxes. Can you think of anything more terrible than saying that
>
> $$0 = 1 - 2^n + 3^n - 4^n + \text{ etc.}$$
>
> where n is a positive integer. Risum teneatis amici. My eyes have been opened in the most amazing way; indeed when you except the simplest cases, e. g. the geometric series, there hardly exists in all of mathematics a single infinite series

the sum of which has been determined in a rigorous way. In other words the most important part of mathematics is without a foundation. Most of it is correct, that is true, and that is very strange. I will try to find the reasons for that. A highly interesting problem. – I do not think you could state many theorems concerning infinite series against whose proof I could not make well founded objections. Do so and I will answer. – Even the binominal [sic] formula is not rigorously proved yet ⋯ . Taylor's Theorem, the basis of all higher mathematics is equally badly founded. I have only found one rigorous proof; it is by Cauchy in his *Resumé des leçons sur le calcul infinitesimal*. He shows that one has

$$\varphi(x+\alpha) = \varphi x + \alpha\varphi' x + \frac{\alpha^2}{2}\varphi'' x + \dots$$

as long as the series is convergent (but one uses it without more ado in all cases) [Abel had clearly not read the *Resumé* carefully. Whether he knew Cauchy's proof of the binomial theorem in the *Cours d'analyse* is unclear] ...

On the whole until now the theory of infinite series is very badly founded. – One applies all operations to infinite series as though they were finite, but is that permissible? Hardly. – Where is it proved that one gets the differential of an infinite series by differentiating each term? – The same holds true of multiplication, division etc. of infinite series .... (Abel 1902, 16-18)

Here Abel spotted many of the weaknesses in the arguments of his contemporaries, including some (like termwise differentaition) that had been overlooked by Gauss and Cauchy. In his letter to Hansteen he announced that he would publish several small papers on these questions in Crelle's Journal, but probably due to his early death he only published one such paper (Abel 1826) dealing with the binomial theorem. The most interesting part of this paper is the introduction where he repeats several of the critical remarks from the letters quoted above and the subsequent general theorems about series. About the latter theorems Abel remarked:

The excellent work by Cauchy: "*Cours d'analyse de l'école polytechnique*" which ought to be read by every analyst who loves rigour in mathematical investigations will serve as my guide. (Abel 1826, 313)

In passing, we should remark that Abel gave a more critical picture of Cauchy's personality and style in a letter he wrote to Holmboe after he had arrived in Paris later in the year 1826:

Cauchy is fou and you can't get anywhere with him although for the moment he is the mathematician who knows how mathematics must be done. His things are excellent but he writes very vaguely. At first I understood almost nothing of his works, now I am doing better. Now he has a series of papers printed under the title *Exercises des Mathématiques*. I buy and read them diligently ⋯ Cauchy is tremendously catholic and bigoted. A very strange thing for a mathematician. Otherwise he is the only one working in pure mathematics. Poisson, Fourier, Ampère etc. etc. are only occupied with magnetism and other physical things. (Abel 1902, 43)

Abel took over all Cauchy's definitions, including his concept of an infinitesimal, without feeling compelled to make them more precise. He read Cauchy's definition of convergence in the case of a series of functions to mean pointwise convergence, and therefore found Cauchy's theorem concerning the continuity of the sum of a convergent series of continuous functions to be incorrect, or as he phrased it:

It seems to me that this theorem has exceptions. For instance the series

$$\sin\varphi - \tfrac{1}{2}\sin 2\varphi + \tfrac{1}{3}\sin 3\varphi - \dots \text{ etc.} \tag{6.34}$$

is discontinuous for each value $(2m+1)\pi$ of φ, where m is an integer. It is well known that there is a host of series with similar properties. (Abel 1826, 316)

The series in question is the Fourier series of $\frac{1}{2}x$ which is indeed discontinuous in the points $2(m+1)\pi$, $m \in \mathbb{Z}$.

To replace Cauchy's faulty theorem, Abel stated what is now called Abel's theorem:

THEOREM IV. When the series

$$f(\alpha) = v_0 + v_1\alpha + v_2\alpha^2 + \cdots + v_m\alpha^m + \dots \tag{6.35}$$

converges for a definite value δ of α, then it will converge for any smaller value of α and such that $f(\alpha - \beta)$ will approach the limit $f(\alpha)$ for constantly decreasing values of β, provided that α is smaller than or equal to β. (Abel 1826, 314)

This is followed by the crucial

THEOREM V. Let

$$v_0 + v_1\delta + v_2\delta^2 + \dots \text{ etc.} \tag{6.36}$$

be a [convergent] series in which $v_0, v_1, v_2, \dots$ are continuous functions of one and the same variable x between the limits $x = a$ and $x = b$, then for $\alpha < [\delta]$ the series

$$f(x) = v_0 + v_1\alpha + v_2\alpha^2 + \dots \tag{6.37}$$

is convergent and is a continuous function of x between the same limits. (Abel 1826, 315)

Abel's proof of Theorem IV is correct in principle but not particularly clearly phrased. For example, it is not made clear how the crucial uniformity property comes into play. It is in fact highly questionable if Abel saw this problem, for in the proof of Theorem V he ignored this property. As Kronecker later pointed out, Abel's proof of Theorem V is basically flawed at this point.

It is clear that Theorem V can replace Cauchy's faulty theorem in Cauchy's own proof of the binomial theorem. Abel proceeded along similar lines, but since he was interested in complex values of x, his proof and the associated convergence considerations were more complicated.

We now turn to two particular problems that forced mathematicians to sharpen ideas of rigour in analysis, namely, (1) the problem of convergence of Fourier series and (2) the analysis of Cauchy's theorem on the continuity of the sum of a convergent series of continuous functions.

6.5. Convergence of Fourier series

When Abel in 1826 referred to the trigonometric series

$$\sin\varphi - \tfrac{1}{2}\sin 2\varphi + \tfrac{1}{3}\sin 3\varphi - \dots \tag{6.38}$$

as a counterexample to Cauchy's theorem, he could do so in a rigorous fashion, for in the same paper Abel proved its convergence and determined its sum from his own complex version of the binomial theorem. He could not refer to a general proof of convergence of Fourier series because such a rigorous proof did not yet exist.

To be sure, several arguments had been given. Fourier himself in his *Théorie analytique de la chaleur* (Fourier 1822, §423) gave the following argument: He interchanged integration and summation once more in (6.9) and transformed it into

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\alpha)\lim_{j\to\infty} K_j(\alpha - x) \tag{6.39}$$

where the kernel K_j is defined by

$$K_j(r) = \cos jr + \sin jr\frac{\sin r}{1-\cos r}\,. \tag{6.40}$$

He then argued that for $j = \infty$ the infinitely fast oscillations of $\cos jr$ will make the $\cos jr$ term of the integral zero. The same holds true for the $\sin jr\frac{\sin r}{1-\cos r}$ term, except in an infinitely small neighbourhood of $r = 0$ (or $\alpha = x$) where $\frac{1}{1-\cos r}$ is infinite. In this neighbourhood $f(\alpha)$ can be replaced by $f(x)$ (where, as we pointed out above, Fourier used Cauchy continuity without saying so) so that (6.39) becomes

$$\frac{1}{2\pi}f(x)\int_{-\pi}^{\pi}\sin jr\frac{r}{\frac{1}{2}r^2} \tag{6.41}$$

which Fourier could show to be equal to $f(x)$. In modern terms, Fourier's proof is based on the observation that $\lim_{j\to\infty} K_j(r)$ is a δ distribution (near zero).

When Fourier published his argument, his rival Poisson had already published his own argument (Poisson 1820) concerning cosine series. His idea was that while it is hard to handle the Fourier series $\sum a_n \cos nx$ (where the a_n's are given by the Fourier integrals), it is easy to see what happens if the series is multiplied by the terms of the geometric series $\sum p^n$ for $p \in (0, 1)$. The resulting series

$$\sum_{n=1}^{\infty} p^n a_n \cos nx \tag{6.42}$$

is convergent and Poisson found its sum in terms of the so-called Poisson integral. He then put $p = 1$ in this expression and used questionable arguments to show that the result was $f(x)$. Of course the problem here is that this does not prove that the original series is convergent.

To Cauchy this was the main question, for in his opinion if the series was not convergent, it had no sum. In a paper (Cauchy 1827) he first gave an argument along the lines used by Poisson in order to prove that the sum "est équivalente" to $f(x)$ (whatever he might have meant by that). "But," he continued, "it is important to prove its convergence." He then went on to apply his newly discovered residue theorem in order to transform the Fourier series into a series $\sum_{n=1}^{\infty} \nu_n$ where

$$\begin{aligned}\nu_n = {} & \frac{1}{2n\pi\sqrt{-1}}e^{-\frac{2n\pi}{a}x\sqrt{-1}}\int_0^{\infty} e^{-z}\left[f\left(a+\frac{az}{2n\pi}\sqrt{-1}\right) - f\left(\frac{az}{2n\pi}\sqrt{-1}\right)\right]dz \\ & -\frac{1}{2n\pi\sqrt{-1}}e^{\frac{2n\pi}{a}x\sqrt{-1}}\int_0^{\infty} e^{-z}\left[f\left(a-\frac{az}{2n\pi}\sqrt{-1}\right) - f\left(-\frac{az}{2n\pi}\sqrt{-1}\right)\right]dz\,.\end{aligned} \tag{6.43}$$

However, for very large values of n each of the integrals contained in expression (6.43) will sensibly be reduced to

$$[\beta_n =] - \frac{1}{2n\pi}[f(a) - f(0)]\frac{\sin 2n\pi}{a} \,. \tag{6.44}$$

But, it is clear that the series having the expression (6.44) as its general term will be a convergent series. (Cauchy 1827, 16)

With this remark Cauchy ended his proof implying implicitly that when $\sum \beta_n$ is convergent, $\sum \nu_n$ must also be convergent. As Dirichlet (Dirichlet 1829) pointed out, this conclusion is unwarranted as can be seen from the example

$$\sum_{n=1}^{\infty} \beta_n = \sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}} \quad \text{and} \quad \sum_{n=1}^{\infty} \nu_n = \sum \frac{(-1)^n}{\sqrt{n}}\left(1 + \frac{(-1)^n}{\sqrt{n}}\right) \,. \tag{6.45}$$

The first series is convergent and the second one is divergent, but the ratio between the n-th terms of the two series tends to one as n tends to infinity. Dirichlet correctly rejected Cauchy's proof for this reason. Moreover, Dirichlet argued that the use of complex function theory was inapplicable when the function f is not given as an analytic expression, because it is unclear which values one should assign to it outside of $\mathbb{R}$ where it is originally defined.

Dirichlet's criticism was published in his paper *Sur la convergence des séries trigonométriques qui servent a représenter une fonction arbitraire entre des limites données* (Dirichlet 1829) (a revised German version was published in 1837). In this paper he published his own convergence proof: He considered the $(n+1)$-st partial sum of the Fourier series (6.9) and transformed it in a way different from Fourier into:

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(\alpha)\frac{\sin(n+\frac{1}{2})(\alpha - x)}{2\sin\frac{1}{2}(\alpha - x)}d\alpha \,. \tag{6.46}$$

The kernel involved is now called the Dirichlet kernel. He then proved the main theorem:

Let h denote a positive quantity less than or equal to $\frac{\pi}{2}$, and g a positive quantity smaller than h, then the integral

$$\int_g^h f(\beta)\frac{\sin i\beta}{\sin\beta}d\beta \tag{6.47}$$

in which the function $f(\beta)$ is continuous between the limits of integration and always increasing or decreasing from $\beta = g$ to $\beta = h$, will converge to a certain limit when the number i becomes larger and larger. This limit is equal to zero, except when g has the value zero; in this case it has the value $\frac{\pi}{2}f(0)$. (Dirichlet 1829, 128)

With this at hand he could prove that (6.46) converges to $f(x)$ when f is continuous and monotone, and from this he generalized the convergence result to piecewise continuous and piecewise monotone functions with the refinement that in the points of discontinuity the Fourier series converges to

$$\lim_{\varepsilon\to 0}\tfrac{1}{2}(f(x+\varepsilon) + f(x-\varepsilon)) \,.$$

Dirichlet characterized his proof as rigorous and except for one detail we would agree today. At the end of the paper he loosely indicated that it should be possible to generalize the result to functions f that do not fulfil the above requirements as long as any interval $[a, b] < [-\pi, \pi]$ contains a subinterval $[\alpha, s]$ on which f is continuous (in modern terminology, f's points of discontinuity are nowhere dense). If f does not fulfil this assumption, Dirichlet felt that the Cauchy integral of $f(x)\cos mx$ and $f(x)\sin mx$ would lose their meaning. As an example he mentioned

$$f(x) = \begin{cases} c & \text{for} \quad x \in \mathbb{Q}, \\ d & \text{for} \quad x \in \mathbb{R}\backslash\mathbb{Q} \end{cases} \tag{6.48}$$

when $c \neq d$ (the first truly nonanalytically given function). However, he admitted that

> ... in order that this [the generalisation of the result] can be done with all the clarity that one may desire, it requires some details related to the fundamental principles of the infinitesimal calculus that will be explained in another note. (Dirichlet 1829, 132)

Dirichlet never published the promised paper, but from a letter in 1853 to Gauss it appears that he remained optimistic about the project (as did Gauss). The hope of such a generalization was shattered in 1873/1876 when du Bois-Reymond gave an example of a continuous function whose Fourier series diverges at a point (indeed, at a dense set of points).

6.6. Cauchy's theorem and uniform convergence

There have been differing interpretations of Cauchy's problematic theorem on the sum of a series of continuous functions. The classical interpretation is that Cauchy made a subtle mistake (or was not clear enough) and overlooked the problem with the Fourier series presented by Abel. Grattan-Guinness ((Grattan-Guinness 1970a) and (Grattan-Guinness 1970b)), on the other hand, interprets Cauchy's theorem as a frontal attack on Fourier and his series. This interpretation is rather odd since Cauchy himself in 1827 published a convergence proof for Fourier series. Trying to discredit a theory by disproving a theorem one believes to be true seems an odd strategy. As we saw above, other authors interpret Cauchy's basic notions in such a way that the theorem becomes true. However, this clashes with Cauchy's own statement in 1853 that Fourier series (in particular a series similar to the one mentioned by Abel) are "always convergent for real values of x", and that they provide true counterexamples against his theorem.

As we saw above, Abel used what Lakatos (Lakatos 1976) has termed the method of exception barring, separating out a safe domain (or at least a domain he believed to be safe) where a special case of the theorem holds true. Seidel (Seidel 1847) and Stokes (Stokes 1849) carried out a deeper analysis twenty years later. According to Lakatos, Seidel discovered the method of proofs and refutations in this connection. The central idea of this method is that if one finds a counterexample to a theorem, one should examine the proof of the theorem in order to uncover a (Lakatos writes "the") hidden assumption unconsciously used in the proof but which does not hold in the counterexample. In this way one can rephrase a new version of the theorem and mathematics has progressed. Seidel describes this methodology accurately in the introduction of his paper:

> When one takes for granted the just obtained assurance that the theorem does not hold in general, i.e. that its proof is based on a hidden assumption, and one subjects the assumption to a closer examination, then it is not difficult to discover

the hidden hypothesis. One can then conversely conclude that this hypothesis cannot be fulfilled by series that represent discontinous functions. (Seidel 1847, 36-37)

Yet, instead of formulating a positively stated replacement of Cauchy's theorem, his main result focussed on a precise description of the exceptional case:

THEOREM. If a convergent series represents a discontinuous function of a quantity x and its terms are continuous functions, then in the immediate neighbourhood of the point where the function jumps there exist values of x where the series converges arbitrarily slowly. (Seidel 1847, 37).

Seidel did not give an explicit definition of "arbitrarily slow convergence" but the proof shows that it is a way to describe the lack of uniformity of the convergence near the point of discontinuity. Indeed, if we follow Cauchy's notation and let

$$s(x) = \sum_{x=1}^{\infty} f_i(x)\ , \quad s_n(x) = \sum_{i=1}^{n} f_i(x) \tag{6.49}$$

and

$$r_n(x) = \sum_{i=n+1}^{\infty} f_i(x), \tag{6.50}$$

then in order to show continuity of $s(x)$, we must consider the quantity

$$\begin{aligned} |s(x+h) - s(x)| &= |s_n(x+h) - s_n(x) + r_n(x+h) - r_n(x)| \\ &\leq |s_n(x+h) - s_n(x)| + |r_n(x+h)| + |r_n(x)|\ . \end{aligned} \tag{6.51}$$

Seidel now argues that the main problem is whether for every $\rho > 0$ we can find a $n_0 \in \mathbb{N}$ such that

$$\begin{aligned} r_{n_0}(x+h) &< \rho, \\ r_{n_0+1}(x+h) &< \rho, \\ r_{n_0+2}(x+h) &< \rho\ , \end{aligned} \tag{6.52}$$

for *all* h in a fixed neighbourhood of x. If this is the case, Cauchy's theorem holds because then we can first pick n so large that this is satisfied which will make both of the last terms of (6.50) less than ρ. Afterwards one can use the continuity of s_{n_0} to make the first term less than ρ for h sufficiently small. If the above requirement does not hold, Seidel argues that the convergence must be arbitrarily slow in x. Thus, taking Seidel's proof into consideration, his "arbitrarily slow convergence" seems to be the negation of what is often called uniform convergence in a neighbourhood of x:

$$\exists \delta > 0 \forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \in \mathbb{N} \forall y \in [x-\delta, x+\delta] : n > n_0 => |r_n(y)| < \varepsilon\ . \tag{6.53}$$

At least it is clear that the n_0 is chosen such that it is independent of y. It is less obvious that the length δ of the neighbourhood does not depend on ε. If it does, i.e., if we interchange the first two quantifiers in (6.53), we get a concept that is called uniform convergence in the point x.

Stokes (Stokes 1849) defined infinitely slow convergence similarly, but his treatment of the problem was not an attack of Cauchy, whom he did not mention, and he gave the problem a different twist. He considered a sequence of functions v_n

which he implicitly assumed to be continuous in an interval $[0, a]$ and he let u_n denote $v_n(0)$. He assumed that $\sum_{n=1}^{\infty} v_n(h)$ is convergent for all $h \in [0, a]$ and put

$$(6.54) \qquad V(h) = \sum_{n=1}^{\infty} v_n(h) \text{ for } h \neq 0 \text{ and } U = \sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} v_n(0)\ .$$

> The limit of V [i.e., $\lim_{h\to 0} V(h)$] can never differ from U unless the convergency of the series (6.54) $\left[\sum_{n=1}^{\infty} v_n(h)\right]$ becomes infinitely slow when h vanishes.
>
> The convergency of the series is here said to become infinitely slow when, if n be the number of terms which must be taken in order to render the sum of the neglected terms numerically less than a given quantity e which may be as small as we please, n increases beyond all limit as h decreases beyond all limit. (Stokes 1849, 281)

Here the problem has explicitly been stated as a question about interchangeability of two limit procedures since the problem is whether the two quantities

$$(6.55) \qquad \lim_{h\to 0} \sum_{n=1}^{\infty} v_n(h) \quad \text{and} \quad \sum_{n=1}^{\infty} \lim_{h\to 0} v_n(h)$$

are equal. Stokes's verbal formulations are difficult to formalize since it is not always clear in which order the quantifiers come. Yet, taken at face value, Stokes's description of the infinitely slow convergence seems to be the negation of the following:

$$(6.56) \qquad \exists \delta > 0 \forall \varepsilon > 0 \exists\, n_0 \in \mathbb{N} \forall y \in [x, x+\delta] : |r_{n_0}(y)| < \varepsilon\ .$$

If $r_n(y)$ does not become identically equal to 0 for any n in any neighbourhood of x, this is in fact equivalent to requiring that one can take n larger than a given $n_0 \in \mathbb{N}$. Thus under this assumption Stokes actually used the so-called quasiuniform convergence in the neighbourhood of x ($= 0$) which can in modern terms be described as follows:

$$(6.57) \qquad \exists \delta > 0 \forall \varepsilon > 0 \forall n_0 \in \mathbb{N} \exists n > n_0 \forall y \in [x-\delta, x+\delta] : |r_n(y)| < \varepsilon$$

apart from the facts that Stokes only seems to consider a neighbourhood to the right of $x = 0$.

Stokes (Stokes 1849, 282) even proved the converse theorem that if the convergence does not become infinitely slow near X ($= 0$), $V = U_0$ and in the concluding remark of this proof Stokes wrote in words a sentence that we might translate as follows: The negation of the infinitely slow convergence is characterized by

$$(6.58) \qquad \forall \varepsilon > 0 \forall N \in \mathbb{N} \exists \delta > 0 \exists n > N \forall y \in [x-\delta, x+\delta] : |r_n(y)| < \varepsilon\ .$$

This is called quasiuniform convergence in the point x.

Grattan-Guinness (Grattan-Guinness 1970b, 117) translates Stokes's concept into (6.56), Hardy (Hardy 1918) interprets it as (6.57), and both authors therefore declare the converse theorem to be false. If, however, we assume that (6.58) is what Stokes had in mind, he would be the first to have found the correct theorem stating that the sum is continuous in x if and only if the series converges quasiuniformly in the point x (Dini 1878, §95, 43-44).

I do not think it makes sense to declare any of the above concepts (6.56)–(6.58) to be Stokes' concept of uniform convergence. The truth is that he did not have a completely precise idea of the meaning of infinitely slow convergence. Seidel's concept is more precise, but it is worth remarking that neither of the two were capable of transferring their new understanding to other problematic theorems. Thus Stokes used Cauchy's theorem about termwise integration of an infinite series in his work without remarking that there is also a flaw in the argument here which can be repaired using uniformity. This application of the concept is of course not so obvious for Stokes (or Seidel) when we keep in mind that neither of them described uniform convergence in an interval, but rather concentrated on what goes on at or near a point.

When, as a result of a remark in a paper by his students Briot and Bouquet, Cauchy returned to the problematic theorem after many years of silence, he came close to the concept of uniform convergence in an interval. This, however, is not obvious from his new statement of the theorem:

THEOREM I. If the different terms of the series

$$u_0,\ u_1,\ u_2, \ldots, u_n,\ u_{n+1}, \ldots \tag{6.59}$$

are functions of the real variable x and continuous with respect to this variable between given limits; if moreover the sum

$$u_n + u_{n+1} + \ldots + u_{n'-1} \tag{6.60}$$

always becomes infinitely small for infinitely large values of the integers n and $n' > n$, then the series (6.57) will be convergent and the sum s of the series (6.57) will be a continuous function of the variable x between the given limits. (Cauchy 1853, 33)

The key word that separates this statement from his previous statement is "always" but only in the proof does it become clear what it covers:

> Imagine now, that by attributing to n a suffiently large value one can, for all values of x contained between the given limits, render the module of the expression (6.60) (for arbitrary n'), and subsequently the module of r_n, smaller than an arbitrarily small number ε. (Cauchy 1853, 32)

Thus, "always" covers the concept "uniform Cauchy sequence in an interval" from which Cauchy immediately concluded "uniform convergence in an interval". Cauchy carefully showed that a Fourier series similar to Abels' (Cauchy did not mention Abel) does not "always" converge in this sense, which explains why its sum is discontinuous.

6.7. Weierstrass

Before Seidel, Stokes and Cauchy had analyzed the topic of uniform convergence, the property had been used extensively by the young Weierstrass. Already in 1838 in a paper on elliptic functions Weierstrass's teacher Gudermann used the phrase "convergence in a uniform way" when the "mode of convergence" of a series $\sum\limits_{n=1}^{\infty} f_n(x, \varphi, \psi)$ is independent of the variables φ and ψ. He thought that it was a "remarkable fact" when an infinite series (or product) converges "in a uniform way", but he did not give an exact definition, nor did he use the property in the proof of any theorems. Weierstrass, who probably learned about the concept in

Gudermann's course on elliptic functions in 1839–1840 used it in a critical way in a paper of 1841, where he showed that if a series of analytic functions converges uniformly in a connected domain, its sum is analytic, and one can differentiate it term by term. However, since this paper remained unpublished until the publication of Weierstrass's *Werke* in 1894, the mathematical world did not learn about his essential use of uniform convergence until he began to lecture at the University of Berlin in 1856. The concept of uniform convergence was only a small piece in his complete refoundation of analysis. His lectures were given as a four semester cycle consisting of the following courses:

Theory of analytic functions,
Theory of elliptic functions,
Applications of elliptic functions to geometry and mechanics,
Theory of Abelian functions.

With some variations, he ran through this cycle sixteen times from 1857 to 1887 (Dugac 1973, 62). Weierstrass explained his approach to the foundation of analysis in the beginning of the first course. Unlike the last three courses, the content of the first one was not published during his lifetime. However, the main ideas soon became known through the testimony, notes and works of the many German and foreign students who gathered in Berlin, not in the least because they wanted to follow Weierstrass's lectures. Some of these notes have recently been published in full or in parts (Dugac 1973), (Weierstrass 1988a), and (Weierstrass 1988b).

When Weierstrass came to Berlin, he was also elected a member of the Academy, and he gave a talk (reprinted in (Weierstrass 1988a), (Weierstrass 1988b)) on this occasion, where he discussed the works he had written while still a Gymnasium teacher. They dealt with elliptic and abelian functions, subjects that remained at the heart of his cycle of lectures. It is remarkable that he emphasized the importance of applications to physics but did not even mention the foundation of analysis as a worthwhile area of research. His interest in this field seems to have been aroused by his teaching and by discussions with Kronecker (as can be seen from the notes of Casorati). Later in their careers, Weierstrass and Kronecker came to disagree sharply on the proper foundation of mathematics, e.g., the status of the real numbers, and more generally on actual infinite sets, but (as pointed out by Bottazzini (Bottazzini 1986, 260-264)), at the beginning they shared a belief in the insufficiency of many crucial ideas and proofs of analysis. They also agreed that the true basis should be found in the arithmetic of the natural numbers.

This tendency had surfaced in Weierstrass's lectures by 1864. From this time on he began his lectures on the theory of analytic functions with a construction of the real numbers (see Chapter 10). He continued with a general study of functions and series and applied the results in a special discussion of power series which then formed the basis of the theory of analytic functions.

Weierstrass's approach to the foundation of analysis found in his general discussion of functions and series is very similar to the modern aproach, so there is no point in going through it in detail. Yet let me emphasize a few points. Through his construction of the real numbers, Weierstrass solved questions concerning completeness that had eluded Cauchy and Bolzano. He became famous for his epsilonic style. Cauchy had used quantifiers ε's, δ's, n_0's, inequalities, etc., in his more complicated proofs, but Weierstrass used the technique in all proofs and also in his definitions. For example, in 1861 he defined a continuous function in the following way:

> If $f(x)$ is a function of x and x is a definite value, then the function will change into $f(x+h)$ when x is replaced by $x+h$. The difference $f(x+h) - f(x)$ is called the change that the function undergoes while x is changed into $x+h$. Now, if it is possible to determine a limit δ such that for all values of h, with absolute value smaller than δ, $f(x+h)-f(x)$ will become smaller than any arbitrarily small quantity, then one says that infinitely small changes of the argument correspond to infinitely small changes of the function. Indeed, if the absolute value of a quantity can become smaller than an arbitrarily small quantity, then one says that it can become infinitely small. When a function is of such a nature that infinitely small changes of the argument correspond to infinitely small changes of the function, then one says that it is a continuous function of the argument or that it varies continuously with this argument. (Dugac 1973, 119-120)

We see that Weierstrass, at least in his early lectures, still used the concept of an infinitely small quantity. However, he only used it as a handy abbreviation that was mostly removed by his successors. His definition is entirely unambiguous, corresponding to pointwise continuity. In a similar way Weierstrass defined the limit of a function and of a series distinguishing clearly between pointwise and uniform convergence in an interval. He used the latter to save Cauchy's theorem and also showed that it answered Abel's question on the termwise differentiation of series. He showed that the termwise integration of series was not universally valid as had been assumed earlier but that it was valid for uniformly convergent series. From being an ad hoc idea, uniform convergence had now turned into a central property.

The distinction between pointwise and uniform convergence had been raised in connection with the study of trigonometric series; no such problem gave rise to a discussion of the distinction between pointwise and uniform continuity. However, with Weierstrass's $\varepsilon - \delta$ formalism at hand, the distinction became almost self-evident and in 1872 Heine separated the two concepts and proved that a continuous function on a closed bounded interval is uniformly continuous. Dirichlet had already formulated this theorem in his 1854 lectures on integration (Dugac 1989, 91). Toward the end of the century Borel isolated the property of "compactness" and used a method similar to Heine's to prove that a closed bounded interval is compact (the so-called Heine-Borel theorem) (see (Dugac 1989)).

Heine's definition appeared in his *Funktionenlehre* of 1872. Though Heine was not a student of Weierstrass, he knew of his approach to analysis through the Weierstrass students Cantor and H. A. Schwartz and followed it closely in his paper. This, together with two talks by Weierstrass in 1870 and 1872, was the first glimpse that the public got of Weierstrass's methods.

Weierstrass's talks challenged two widespread beliefs. The first (Weierstrass 1870/1895) concerned the distinction between a maximum and a supremum (or minimum and infinimum). Although Bolzano had called attention to the essential difference, several existence proofs had relied on a confusion of these concepts, the most famous being the so-called Dirichlet principle (see Chapters 12 and 13).

Weierstrass's second paper (1872/1895), which he presented to the Academy in Berlin in 1872, exhibited the function

$$f(x) = \sum_{n=1}^{\infty} b^n \cos(a^n x\pi) \tag{6.61}$$

where a is odd, $b \in [0, 1)$ and $ab > 1 + \frac{3}{2}\pi$, as an example of a continuous function which is nowhere differentiable. We have seen that Bolzano had already found (but not published) a similar function and Weierstrass reported that Riemann had given another example in his lectures, namely,

$$f_1(x) = \sum_{n=1}^{\infty} \frac{\sin(n^2 x)}{n^2} . \tag{6.62}$$

However, it was (and still is) unclear if Riemann had claimed that f_1 is nowhere differentiable or only that it is nondifferentiable in a dense set. Weierstrass could not prove that f_1 is nowhere differentiable and therefore found his own example. In fact Gerver in 1970 showed that f_1 is differentiable in the values $a\pi$, where a is of the form $\frac{2m+1}{2k+1}$ (Neuenschwander 1978). Weierstrass's function contradicted an intuitive feeling held by most of his contemporaries to the effect that continuous functions were differentiable except in "special points". It created a sensation and, according to Hankel, disbelief when du Bois-Reymond published it in 1875.

6.8. Pathological functions and the new style in analysis

Weierstrass's continuous nowhere differentiable function became the most well known of a large number of pathological functions constructed around 1870. Cauchy's example (see 6.3.5) of a C^∞-function that is not represented by its convergent Taylor series and Dirichlet's nonintegrable function (see 6.5) can be seen as early examples of pathological functions. Riemann (Riemann 1854) gave a series of other examples in connection with his study of trigonometric series and the integral (see Chapter 9). Trigonometric and Fourier series as well as the theory of integration gave rise to many bizarre functions (Chapter 9) of which I have mentioned du Bois-Reymond's continuous function that cannot be developed in a Fourier series. Hankel (Hankel 1870) and Darboux (Darboux 1875) constructed a series of pathological functions, and the former even invented a method such that if he had a function with a particular singularity in one point, he could often construct a new function that possessed this property on a dense set of points. He called this method "condensation of singularities".

The pathological functions heralded a new trend in mathematical analysis. Where, earlier, new types of functions had arisen from or been forced upon mathematicians by applications, mathematicians now actively sought unpleasant functions in the framework of pure mathematics in order to delineate the limits of concepts such as function, continuity, differentiability, integrability, etc.

Several mathematicians were highly critical of this new trend. For example, Poincaré expressed his scepticism as follows:

> For half a century we have seen a mass of bizarre functions which appear to be forced to resemble as little as possible honest functions which serve some purpose. ...
>
> In former times when one invented a new function, it was for a practical purpose; today one invents them purposely to show up defects in the reasoning of

our fathers and one will deduce from them only that. (Poincaré 1899, 130-131), Transl. (Kline 1972, 973)

However, the pathological functions showed that Dirichlet's concept of function was too general to serve directly as a basis of analysis. According to Hurwitz, Weierstrass himself went so far as to declare the general Dirichletian function concept to be "totally untenable and unfruitful. In fact it is impossible to deduce any general properties of functions from it" (Dugac 1973, 116). In a sense one can view Weierstrass's work on the foundations of analysis as a search for the most useful concept of function (see (Dugac 1973, 71)). Laugwitz (Laugwitz 1992) has emphasized that for Weierstrass "the final goal is always the representation of a function". By "representation" Weierstrass meant analytic or arithmetic "representation". In this sense Weierstrass's view of the concept of function was a continuation of Euler's and Lagrange's algebraic point of view, and it stood in stark contrast to the more conceptual point of view put forward by the Göttingen mathematicians such as Riemann. Most of Weierstrass's contemporaries and followers seems to have given up the idea of considering analytic expressions as a class of particularly well-behaved functions. After all, Weierstrass's approximation theorem and Fourier series showed that the class was vast (later work by Baire and Lebesgue (Lebesgue 1905) brought this out even more clearly) and Hankel showed that one could construct analytic expressions having all the different singularities he considered in his 1870 paper. Since the integral theory proved that one could derive essential theorems even for rather singular functions, mathematicians stuck with the general Dirichlet concept of function. Of course they had to give up the hope that analysis could be universally valid in one specific subregion. Instead each theorem required its own explicitly stated assumptions. A theorem, like the mean value theorem, that had been written earlier in the form of just one equation now took the form:

PROCLAIM. Let A be a ... subset of $\mathbb{R}$ which is ... (closed, open ... everywhere dense, measurable simply connected, etc.) and let f be a function defined on A (or $\overline{A}$ or ...) which is C^n or ... in A and ... on $\overline{A}$ and let x_0 be a point of A such that ..., then ....

In this way, by proceeding according to the method of proofs and refutations, analysis gained both rigour and generality, but it lost in elegance and simplicity and was estranged from intuition and physical applications. Many mathematicians regretted this tendency, but it was hard to escape once the pathological functions had entered the Eden of "naive" analysis.

The outcome of this process of rigorization was much more than the old theorems suitably rephrased and provided with rigorous proofs. It was the creation of a whole new point set topological and measure theoretical basis for analysis with its own new concepts and its own results. This basis was eventually almost disconnected from its source, and its theorems, such as the Heine-Borel theorem, the Bolzano-Weierstrass theorem, etc., became ends in themselves. In this way the insistence on rigour that started out (with Berkeley) as a destructive movement ended up as a strong creative force.

6.9. Diffusion and acceptance of rigourist analysis

In this chapter we have dealt with a number of mathematicians who actively pushed for the rigorous foundation of analysis, and we have only considered those of their works that deal with foundational questions. However, their ideas did not prevail immediately either in teaching or in research.

When Cauchy introduced his new standard of rigour at the École Polytechnique, he was criticized by his colleagues and superiors for emphasizing foundations at the expense of applications ((Belhoste 1991, 61-86) and (Gilain 1989)). His first co-teacher on the course, Ampère, followed Cauchy some of the way but Navier, who began teaching the course in 1819, emphasized applications and did not conform to Cauchy's standards of rigour. During the 1840s Cauchy's methods were taught again at the school but the professors Sturm and Liouville were not themselves interested in carrying the process of rigorization further (Lützen 1990, 72-76). However, Sturm became responsible indirectly for spreading Cauchy's ideas because his lecture notes, published posthumously in 1857–1859 were widely read (the 14th edition was published as late as 1909).

In other schools, and in particular outside of France, it took much longer before Cauchy's ideas replaced the older style. In England where Lagrange's approach had replaced Newton's fluxional style during the 1810s, partly through the agitation of the so-called Analytical Society, the algebraic formal style flourished well into the second part of the 19th century.

The second step in the rigorization process was felt at the École Polytechnique in 1893–1896 when Jordan introduced Weirstrassian $\varepsilon - \delta$-techniques in the second edition of his *Cours d'analyse*. This textbook was very much read and greatly responsible for spreading the new standards. However, in many universities Cauchyan style analysis was taught well into the 20th century. For example, Sturm's book was used in Copenhagen until 1915. Thus both the Cauchyan and the Weierstrassian reform took about forty years to reach the general classroom.

In research work there was a similar delay. Cauchy himself repeatedly sinned against his own rigorous standards in his research (e.g., he used divergent series) and other mathematicians did the same. In the case of applied mathematics this is not so surprising; these branches have often been developed with a certain laxness toward mathematical rigour. But in the pure mathematical research of Cauchy and his followers, we also find a great deal of freedom as far as the methods are concerned. A particularly striking example is the loosely founded operational techniques by which Cauchy and, in particular, British mathematicians tried to manipulate differential operators. A special branch of this field is the theory of operators of the form $\left(\frac{d}{dx}\right)^{\mu}$, $\mu \in \mathbb{C}$, developed by Liouville and Riemann in the 1830s and 1847, respectively. In (Gispert 1983) Gispert pointed out how Darboux's plea for Weierstrassian rigour in analysis was generally rejected by his French colleagues.

One should not underestimate how difficult it was to understand the new standards of rigour even for mathematicians who were open to the new trends. Liouville admitted to his friend Dirichlet, that he "found it rather difficult to explain (and even to understand) the proof which Abel has given of (his) important theorem" (Dirichlet 1862). Dirichlet was a master of the new standards, and so he was able to clear up the matter "offhand and before [Liouville's] very eyes". Yet, when formulating Dirichlet's principle in potential theory (cf. Chapters 12 and 13), Dirichlet himself sinned against the ideal of rigour. This is all the more surprising because he reproached Steiner for having made a similar error in his proof of the isoperimetric property of the circle. Du Bois-Reymond, who published the Weierstrass continuous, nowhere differentiable function, declared:

> It appears to me that the metaphysics of Weierstrass's function still hides many riddles and I cannot help thinking that entering deeper into the matter will finally lead us to the limits of our intellect. (Du Bois-Reymond 1875, 29)

When Hankel tried to classify the pathological functions he had constructed according to their singularities, he committed what now seem glaring mistakes by confusing topologically small sets, e.g., nowhere dense sets and measure theoretically small sets that have measure equal to zero in some sense. These differences were cleared up around 1880 by Smith, du Bois-Reymond, Volterra and Harnack (see Chapter 9).

6.10. Breaking the rigorous chains

Towards the end of the 19th century the new standards of rigour began to dominate mathematical research, but there was still some resistance. Some of the opposition was due to outright conservatism, but to a certain degree there was a good reason to oppose the rigourous chains imposed on analysis. For mathematicians like Poincaré (and Riemann before him) who had a well-developed mathematical intuition, it would have been crippling to insist that one should clear up every single ε in the arguments. Poincaré's own ingenious publications did in many cases not live up to the ideal of rigour, but when he most clearly sinned against this ideal (e.g., using the Dirichlet principle), he admitted that the argument was only a heuristic one (cf. (Poincaré 1896, 118-119)). Some mathematicians felt that the rigourists had been too radical. For example, when banning divergent series, they excluded many successful arguments in applied physics and astronomy. Heaviside, himself a master of odd arguments using divergent series in electromagnetic theory, expressed it in his usual colourful way:

> I must say a few words on the subject of generalised differentiation and divergent series ... It is not easy to get up any enthusiasm after it has been artificially cooled by the wet blankets of the rigorists. Nevertheless, I have been informed that I have been the means of stimulating some interest in the subject in certain places. Perhaps not in England ... but certainly in Paris ... There will have to be a theory of divergent series, or, say, a larger theory of functions than the present, including convergent and divergent series in one harmonious whole. (Heaviside 1899, §425 and §432)

In Paris, Poincaré (and independently Stieltjes) (1886) succeeded in creating a theory of so-called asymptotic series which rescued many of the earlier arguments with divergent series. Another approach was begun by Frobenius in 1880 and Hölder in 1882 and developed by Cesàro (1890), who defined the sum of a large class of divergent series. Though the series do not approach their limit when the number of terms increase, the sums defined in this way turned out to make sense both in applications and in theoretical work. For example, Fejér (1904) showed that a bounded Riemann integrable function has a Fourier series that is summable with the "correct" sum (cf. (Kline 1972, 196-1121)).

Another area where it turned out that the rigourists of the 19th century had been too radical was in their insistence that only differentiable functions can be differentiated. Heaviside also refers to this in the above quote. Euler had tried to argue that any arbitrary function φ would give rise to a solution $y(x,t) = \varphi(x-t)$ to the wave equation $\frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}$, but such ideas were brushed aside during the

19th century. In Laurent-Schwartz's theory of distributions (Schwartz 1950/1951), Euler's claim makes perfect sense: The derivatives do not necessarily exist as functions but as generalized functions. Schwartz (and Sobolev before him) also made sense of the "δ-function" that had been used by 19th-century mathematicians such as Fourier, Kirchhoff and Heaviside but that had been discarded by rigorous mathematicians (cf. (Lützen 1982)). Infinitesimals that had left the scene around 1870 could be made into perfectly sound objects. In 1960/1961 Robinson constructed a non-Archimedean field extension of $\mathbb{R}$ that contains infinitesimals. With these at hand he could reestablish many of Leibniz's, Euler's and even Cauchy's arguments on a firm basis (cf. (Robinson 1966) and (Laugwitz; Schmieden 1958)).

The discovery of this so-called nonstandard analysis seems to have had an influence on the historiography of the foundations of calculus. As long as there was only one accepted version of this field, namely, the Weierstrassian, the development had often been seen as a natural striving towards this one natural goal. Nonstandard analysis has made it clear that there is not a unique goal so that instead of viewing the historical process as almost inevitable, we must now view it as one among many possible scenarios that has to be explained in its context.

The 20th-century theories of divergent series, generalized differentiation, generalized functions and infinitesimals may even lead someone to consider the rigorous ideals of the 19th century as an unnecessarily restricted or even perverted stage that we have now outgrown. To this, one must answer that the more general ideas of the 20th-century are all based on the rigorous foundation developed in the 19th century. For example, Schwartz's generalized functions (distributions) are defined as functionals on infinitely differentiable functions with compact support equipped with a suitable topology. Thus, though the 19th-century development of the foundations of calculus cannot be considered as a necessary or even a natural development, our modern analysis is firmly rooted in it.

Bibliography

Abel, N.H. 1826. Untersuchung über die Reihe $1+\frac{m}{1}x\frac{m(m-1)}{1\cdot 2}+\frac{m(m-1)(m-2)}{1\cdot 2\cdot 3}x^3+\ldots$, *Journal für die reine und angewandte Mathematik* **1** (1826), 311–339 in *Oeuvres complètes* **1**, 219–250 in *Ostwald's Klassiker* nr 71, Leipzig 1921; (page references to the original paper).

Abel, N.H. 1902. *Festskrift ved Hundredeaarsjubilæet for Niels Henrik Abels Fødsel*, Kristiania 1902. French transl.: *Mémorial publié à l'occasion du centenaire de sa naissance*, Christiania 1902.

Ampère, A.M. 1824. *Précis des leçons sur le calcul différentiel et le calcul intégral*, Paris.

Arbogast, L.F.A. 1791. *Mémoire sur la nature des fonctions arbitraires qui entrent dans les intégrales des équations aux différences partielles*, Sct. Petersburg 1791.

Belhoste, B. 1991. *Augustin-Louis Cauchy. A biography*, Springer-Verlag, New York 1991.

du Bois-Reymond, P. 1875. Versuch einer Classification der willkürlichen Funktionen reeller Argumente nach ihren Aenderungen in den kleinsten Intervallen, *Journal für die reine und angewandte Mathematik* **79** (1875), 21–37.

Bolzano, B. 1817. *Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewären wenigstens eine reelle Wurzel der Gleichung liege*, Prag 1817, *Ostwald's Klassiker* Nr. 153 Leipzig 1905; (English edition by S.B. Russ. Historica Mathematics **7** (1980), 156–185).

Bolzano, B. 1930. *Funktionenlehre*, ed. K. Rychlik, Prague 1930.

Bottazzini, U. 1986. *The Higher Calculus*, Springer, New York 1986.

Bottazzini, U. 1990. Geometrical rigour and "modern" analysis. An introduction to Cauchy's *Cours d'analyse*, pp. xi-clxvii, in Bottazzini's 1990 edition of Cauchy (1821).

Carnot, L. 1797. *Réflexions sur la métaphysique du calcul infinitésimal*, Paris an V. (1797) 2nd enlarged ed. 1813.

Cauchy, A.L. 1814. Mémoire sur la théorie des intégrales définies, *Mémoire presentées a l'Academie Roy. des Sciences par divers savans* **2** (1827), 601–799, in *Oeuvres* (1) **1**, 319–506.

Cauchy, A.L. 1821. *Cours d'analyse de l'École Royale Polytechnique. 1re partie. Analyse algébrique*, Paris 1821. Oeuvres (2) **3** Later editions: Gabay, Paris 1989. CLUEB, Bologna 1990 (ed. Bottazzini).

Cauchy, A.L. 1821/1885. *Algebraische Analysis*, Deutsche Übersetzung von (Cauchy, 1821) von C. Itzigsohn, Springer, Berlin 1885.

Cauchy, A.L. 1823. *Résumé des leçons données a l'école royale polytechnique sur le calcul infinitesimal. Tome premier*, Paris 1823 in Oeuvres (2) **4**, 5–261.

Cauchy, A.L. 1827. Mémoire sur les développements des fonctions en séries périodiques, *Mémoire de l'Académie des Science* **6** (1827) p. 603 in *Oeuvres* (1) **2**, 12–19.

Cauchy, A.L. 1829. *Leçons sur le calcul différentiel*, Paris 1829 in Oeuvres (2) **4**, 263–609. German transl.: *Vorlesungen über die Differentialrechnung*, Braunschweig 1836.

Cauchy, A.L. 1831. Sul calcolo infinitesimale ed in particolare sul calcolo differenziale, *Biblioteca italiana, o sia Gionale di letteratura scienze ed arti complato da vari letterati* **61** (1831), 321–334, in Oeuvres (2) **15**, 161–171.

Cauchy, A.L. 1844. Mémoire sur les fonctions continuées ou discontinuées, *Comptes Rendus de l'Académie des Sciences*, Paris **18** (1844), 116–130, in *Oeuvres* (1) **8**, 145–160.

Cauchy, A.L. 1853. Note sur les séries convergentes dont les divers termes sont des fonctions continues d'une variable réelle ou imaginaire, entre des limites donnés, *Comptes Rendus de l'Académie des Sciences* **36** (1853), p. 454, in *Oeuvres* (1) **12**, 30–36.

Darboux, J.G. 1875. Mémoire sur les fonctions discontinues, *Annales Sci. École Normale Superieure* (2) **4**, 57–112.

Dhombres, J. 1992. Le rôle des équations fonctionnells danc l'Analyse algébrique de Cauchy, *Revue d'Histoire des Sciences* **45** 1992, 25–49.

Dini, U. 1878. *Fondamenti per la teorica delle funzioni di variabili reale*, Piza 1878. German transl. *Grundlagen für eine Theorie der Funktionen einer veränderlichen reellen Grösse*, Leipzig 1892 (page reference to the German transl.).

Dirichlet, P.G. Lejeune 1829. Sur la convergence des séries trigonométriques qui servent a représenter une fonction arbitraire entre des limites données, *Journal für die reine und angewandte Mathematik* **4** (1829), 157–169, in *Werke* **1**, 117–132.

Dirichlet, P.G. Lejeune 1837. Über die Darstellung ganz willkürlicher Funktionen durch sinus- und cosinusreihen, *Repertorium der Physik* **1** (1837), 152–174 in *Werke* **1**, 133–160 in *Ostwald's Klassiker* Nr 116, Leipzig 1900, pp. 1–34.

Dirichlet, P.G. Lejeune 1856/1857 publ. 1876. *Vorlesungen über die im umgekehrten Verhältniss des Quadrats der Entfernung wirkenden Kräfte*, Herausgegeben von F. Grube, Leipzig 1876.

Dirichlet, P.G. Lejeune 1862. Démonstration d'un théorème d'Abel. Note de M. Lejeune Dirichlet communiquée par M. Liouville, *Journal de Mathématiques pures et appliquées* (2) **7** (1862), 253–255; *Werke* **2**, 305–306.

Dugac, P. 1973. Eléments d'analyse de Karl Weierstrass, *Archive for History of Exact Sciences* **10** (1973), 41–176.

Dugac, P. 1989. Sur la correspondance de Borel et le théorème de Dirichlet-Heine-Weierstrass-Borel-Schoenflies-Lebesgue, *Archives Internationales d'Histoire des Sciences* **39** (1989), 69–110.

Euler, L. 1748. *Introductio in analysin infinitorum (2 vols)*, Lausanne 1748 in *Opera Omnia* (1) **8-9**; German transl. Berlin 1788.

Euler, L. 1755/1790. Institutiones calculi differentialis. 2 vol. St. Petersburg, or in: *Opera Omnia*, $\mathbf{II}_{10}$.

Fauvel, J.; Gray, J. 1987. *The History of Mathematics: A Reader*, Milton Keynes: The Open University, 1987.

Fourier, J. 1822. *Théorie analytique de la chaleur*, Paris 1822; in Oeuvres **I**, reprinted by Gabay, Paris 1988.

Freudenthal, H. 1971. Did Cauchy plagiarize Bolzano?, *Archive for History of Exact Sciences* **7** (1970/71), 375–392.

Gauss, C.F. 1799. *Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*, Helmstädt 1799; in Werke **3**, 1–31.

Gauss, C.F. 1813. Disquisitiones generales circa seriem infinitam

$$1 + \frac{\alpha\beta}{1 \cdot \gamma}x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1 \cdot 2 \cdot \gamma(\gamma+1)} + \text{etc.},$$

Commentationes societatis regiae scientiarum Göttenhensis resentiores **2** (1813); in *Werke* **3**, 123–162.

Gauss, C.F. 1840. Allgemeine Lehrsätze in Beziehung auf die im verkehrten Verhältnisse des Quadrats der Entfernung wirkenden Anziehungs- und Abstossungs-Kräfte, *Resultate aus den Beobachtungen des magnetischen Vereins im Jahre 1839* (4) Leipzig 1840; *Werke* **5**, 197–242; *Ostwald's Klassiker* Nr. 2 Leipzig 1889.

Gauss, C.F. 1850/1865. Brief von Gauss an Schumacher, 1 September 1850, *Briefwechsel zwischen Gauss und Schumcaher*, **6** Altona 1865 p. 107, Gauss Werke $\mathbf{X}^1$, 434–436.

Gilain, C. 1989. Cauchy et le Cours d'Analyse de l'Ecole Polytechnique, pp. 3–46 in *Augustin-Louis Cauchy 1789-1857 Bulletin de la Société des Amis de la Bibliothéque de l'Ecole polytechnique* Nr. 5, Paris 1989.

Gispert, H. 1983. Sur les fondements de l'analyse en France, *Archive for History of Exact Sciences* **28** (1983), 37–106.

Giusti, E. 1984. Gli "errori" di Cauchy e i fondamenti dell'analisi, *Bolettino di Storia delle Scienze matematiche* **4** (1984), 24–54.

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

Grattan-Guinness, I. 1970a. Bolzano, Cauchy and the "New analysis" of the Early nineteenth Century, *Archive for History of Exact Sciences* **6** (1969/70), 372–400.

Grattan-Guinness, I. 1970b. *The development of the foundations of mathematical analysis from Euler to Riemann*, MIT Press, Cambridge, MA 1970.

Green, G. 1833/1835. On the determination of the exterior and interior attractions of ellipsoids of variable densities, *Transactions of the Cambrige Philosophical Society* (5) **3** (1835) in *Mathematical Papers*, 187–222.

Hankel, H. 1870. *Untersuchungen über die unendlich oft oscillioenden und unstetigen Funktionen*, Tübingen 1870, wieder gedrucht in *Mathematische Annalen* **20** (1882), 63–112. *Ostwald's Klassiker* Nr. 153, Leipzig 1905; (page reference to Ostwald's Klassiker).

Hardy, G.H. 1918. *Sir George Stokes and the concept of uniform convergence*, Proceedings of the Cambridge Philosophical Society **19** (1918), 148–156.

Heaviside, O. 1899. *Electromagnetic Theory* vol **2**, London 1899.

Heine, E. 1872. Die Elemente der Funktionenlehre, *Journal für die reine und angewandte Mathematik* **74** (1872), 172–188.

Hilbert, D. 1900. Über das Dirichletsche Prinzip, *Jahresber. Deut. Math. Verein.* **8** (1900), 184–188; *Journ. Reine Angew. Math.* **129** (1905), 63–67; *Ges. Abhandl.* **3**, 10–14.

Jahnke, H.N. 1990. *Mathematik und Bildung in der Humboldtschen Reform.* Göttingen: Vandenhoeck & Ruprecht.

Jarnik, V. 1981. *Bolzano and the Foundations of Mathematical Analysis*, Society of Czechoslovak Mathematicians and Physicists, 1981.

Kline, M. 1972. *Mathematical Thought from Ancient to Modern Times.* New York: Oxford University Press.

Lacroix, S.F. 1820. *Traité élementaire de calcul différentiel et de calcul intégral*, 3rd ed. Paris 1820.

Lagrange, J.L. 1797. *Theorie des fonctions analytiques*, Paris au V (1797). (2nd ed. 1813), (3rd ed. 1842), in Oeuvres **9**.

Lagrange, J.L. 1806. *Leçons sur les calculs des fonctions*, 2nd ed. Paris 1896 in Oeuvres **10**.

Lakatos, I. 1976. *Proofs and refutations. The logic of mathematical discovery*, eds. J. Worrall and E. Zahar. Cambridge University Press, Cambridge 1976.

Lakatos, I. 1978. Cauchy and the continuum, *The Mathematical Intelligencer* **1** (1978), 151–161, also pp. 43–60 in Lakatos: *Mathematics, Science and Epistemology* 1980, Philosophical papers **2**, 43–60.

Laugwitz, D. 1987. Infinitely Small Quantities in Cauchy's Textbooks, *Historia Mathematica* **14** (1987), 258–274.

Laugwitz, D. 1989. Definite values of infinite sums: Aspects of the foundations of infinitesimal analysis around 1820, *Archive for History of Exact Sciences* **39** (1989), 195–245.

Laugwitz, D. 1992. "Das Letzte Ziel is immer die Darstellung einer Funktion". Grundlagen der Analysis bei Weierstrass 1886, *Historia Mathematica* **19** (1992), 341–355.

Laugwitz, D; Schmieden, C. 1958. Eine Erweiterung der infinitesimalrechnung. Mathematische Zeitschrift, 69, 1–39.

Lebesgue, H. 1905. Sur les fonctions représentables analytiquement, *Journal de Mathématiques Pures et Appliques* **1** (1905), 139–216.

Lützen, J. 1982. *The Prehistory of the Theory of Distributions*, Springer-Verlag, New York, Heidelberg, Berlin 1982.

Lützen, J. 1990. *Joseph Liouville 1809-1882. Master of pure and applied mathematics*, Springer-Verlag, New York 1990.

Neuenschwander, E. 1978. Riemann's Example of a Continuous, 'nondifferentiable' function, *The Mathematical Intelligencer* **1** (1978), 40–44.

Poincaré, H. 1896. La méthode de Neumann et le problème de Dirichlet, *Acta Mathematica* **20** (1896–1897), 59–142; *Oeuvres* **9**, 202–272.

Poincaré, H. 1899. La logique et l'intuition dans la science mathématique et dans l'enseignement, *L'enseignement mathématique* **1** (1899), 157–162, in *Oeuvres* **11**, 129–133.

Poisson, S.D. 1820. Mémoire sur la manière d'exprimer les fonctions par des séries de quantités périodiques, *Journal de l'Ecole Polytechnique* **11** (18th cahier), 417–489.

Riemann, B. 1851. *Grundlagen für eine allgemeine Theorie der Funktionen einer veränderlichen complexen Grösse*, Inaugural dissertation Göttingen 1861. 2nd ed. Göttingen 1867. *Gesammelte Math. Werke* pp. 3–45.

Riemann, B.G.F. 1854. *Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe. Habilitationsschrift*, Universität Göttingen, or in: Gesammelte Mathematische Werke, 259–303.

Riemann, B. 1856/1867. *Ueber die Darstellbarkeit einer Function durch eine trigonometrische Reihe*, Habilitationsschrift 1854 in *Abhandlungen der Königschen Gesellschaft der Wissenschaften zu Göttingen* **13** (1867); in *Werke*, 227–271.

Robinson, A. 1966. *Nonstandard analysis*, Amsterdam 1966.

Robinson, A. 1967. *The Metaphysics of the Calculus*, pp. 28-40, in, Problems in the Philosophy of Mathematics (ed. I. Lakatos), Amsterdam 1967.

Rychlik, K. 1962. *Theorie der reellen Zahlen in Bolzano's handschriftlichem Nachlasse*, Wissenschaftlicher Redakteur Ladislaw Rieger, Prag 1962.

Rüthing, D. 1984. *Some Definitions of the Concept of Function from Joh. Bernoulli to N. Bourbaki*, The Mathematical Intelligencer **6** (1984), 72-77.

Schwartz, L. 1950/1951. *Théorie des distributions*, vol **1** Paris 1950, Vol **2** Paris 1951.

Seidel, P.L. 1847. Note über eine Eigenschaft der Reihen, welche discontinuirliche Funktionen darstellen, *Abhandlungen der Math. Phys. Klasse der Kgl. Bayrischen Akademie der Wissenschaften* **5** (1847), 381–394; in *Ostwald's Klassiker* Nr. **152**, 35–45, (page references to Ostwald's Klassiker).

Spalt, D. 1992. Aufklärung des Cauchy-Problems, *Preprint Fachbereich Mathematik der Technischen Hochschule Darmstadt nr. 1457*, 1992.

Stokes, G.G. 1849. On the critical values of the sums of periodic series, *Transactions of the Cambridge Philosophical Society* **8** (1849), 533–583, *Mathematical and Physical Papers* **1**, 236–313.

Sturm, C. 1857-1859. *Cours d'analyse de l'Ecole Polytechnique*, (ed. Prouhet) 2 vols., Paris 1857, 1859.

Thomson, W. (later Lord Kelvin) 1847. Note sur une équation aux différences partielles qui se présente dans plusieurs questions de physique mathématique, *Journal de mathématiques pures et appliques* **12** (1847), 493–496. *Cambridge and Dublin Math. Journal* **3** (1848), 84–87, *Math. and Phys. Papers* **1**, 93-96.

Weierstrass, K. 1841/1894. *Zur Theorie der Potenzreihen*, *Werke* **1**, 67–74.

Weierstrass, K. 1870/1895. Über das sogenannte Dirichletsche Prinzip, in Werke **2**, 49–54, reprinted in Weierstrass 1988a.

Weierstrass, K. 1872/1895. Über continuirliche Funktionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten bezitsen, in *Werke* **2**, 71–74, reprinted in Weierstrass 1988a.

Weierstrass, K. 1988a. *Ausgewählte Kapitel aus der Funktionenlehre. Vorlesung, gehalten in Berlin 1886*, (ed. R. Siegmund-Schultz). Teubner Archiv zur Mathematik Leipzig, 1988.

Weierstrass, K. 1988b. *Einleitung in die Theorie der analytischen Funktionen. Vorlesung Berlin 1878 in einer Mitschrift von Adolf Hurwitz*, (ed. W. Scharlau and P. Ullrich). DMV Vieweg. Braunschweig 1988.

Youschkevich, A.P. 1976. The concept of function up to the middle of the 19th century, *Archive for History of Exact Sciences* **16** (1976), 37–85.

Bolzano, Cauchy and the "New Analysis" of the Early Nineteenth Century

I. GRATTAN-GUINNESS

Communicated by J. E. HOFMANN

Summary

In this paper[1] I discuss the development of mathematical analysis during the second and third decades of the nineteenth century; and in particular I assert that the well-known correspondence of new ideas to be found in the writings of BOLZANO and CAUCHY is *not* a coincidence, but that CAUCHY had read one particular paper of BOLZANO and drew on its results without acknowledgement. The reasons for this conjecture involve not only the texts in question but also the state of development of mathematical analysis itself, CAUCHY both as personality and as mathematician, and the rivalries which were prevalent in Paris at that time.

Contents

1. Introduction 372
2. The Common Ideas in BOLZANO and CAUCHY 373
3. The New Analysis 378
4. The Old Analysis 381
5. CAUCHY's Originality as a Mathematician 384
6. The State of Parisian Mathematics 387
7. CAUCHY's Personality 393
8. The Availability and Familiarity of BOLZANO's Work 395
9. The Personal Relations between BOLZANO and CAUCHY 397
10. Epilogue 398
Index of Names 400

1. Introduction

The central theme of this paper is an historical conjecture concerning the development of mathematical analysis in the early nineteenth century. It is well known that the major event was the publication in 1821 of the *Cours d'Ana-*

[1] This paper is a revised and greatly expanded version of a lecture entitled "Did Cauchy read Bolzano before writing his *Cours d'Analyse?*" given at the *Problemgeschichte der Mathematik* seminar at Oberwolfach, West Germany, on the 26th November, 1969. I wish to thank Professors J. E. HOFMANN and C. J. SCRIBA for their invitation to this seminar.

The text draws frequently on my history of *The Development of the Foundations of Mathematical Analysis from Euler to Riemann* and *Joseph Fourier 1768—1830*, which are both to be published by the M.I.T. Press and are referred to in later footnotes as *Foundations* and *Fourier*, respectively. The latter work was written with the collaboration of Dr. J. R. RAVETZ, and the former with the help of his detailed criticism: I wish to record here my indebtedness to his assistance.

lyse of AUGUSTIN-LOUIS CAUCHY (1789—1857),[2] in which CAUCHY presented a new type of analytical reasoning far superior to previous ideas for the development of analysis — limits, functions, the calculus, and so on. CAUCHY'S achievement was the so-called "arithmeticisation" of analysis, a method whose development and application has been a major interest for mathematicians ever since.

It has been also well-known for some time that CAUCHY had been anticipated in his basic ideas of the new analysis by an obscure pamphlet published in Prague in 1817 by BERNARD BOLZANO (1781—1848). In contrast to the broad programme of CAUCHY'S book, BOLZANO devoted his little work to the proof of a theorem which he described in its title: "Purely analytical proof of the theorem, that between any two values [of a function $f(x)$] which guarantee an opposing result [in sign] lies at least one real root of the equation [$f(x)=0$]."[3] The "pure analysis" which BOLZANO produced in his proof is exactly that which we find greatly developed and extended in CAUCHY'S *Cours d'Analyse* and his later writings on analysis.

I do not believe that we have here an example of a remarkable coincidence of new ideas. Such occurrences are of course well-known in the history of science, but I shall argue for the conjecture that in this case CAUCHY was well acquainted with BOLZANO'S paper and that he drew on its novelties without ever making acknowledgement to him.

The argument for this thesis is not based on new documentary evidence: there is no reference to BOLZANO'S work among the scattered fragments of CAUCHY'S papers and letters, no library record of CAUCHY'S reading or borrowing BOLZANO'S paper, no copy of it in his personal library (which in fact has been dispersed). My reasons for the conjecture are circumstantial and related to intellectual matters, and involve not only the general development of analysis at that time but also that aspect of the growth of science which is ignored all too often by its historians — the social and educational situation of the period, and the personalities of the principal characters.

2. The Common Ideas in Bolzano and Cauchy

We consider first the directly corresponding results in the two works, in each case in its general historical setting.

2.1. Continuity of a Function. Normally the continuity of a function was then identified with its description by a single algebraic expression, and the function was usually thought to be differentiable: in fact, under EULER'S influence the

[2] A.-L. CAUCHY, *Cours d'Analyse de l'Ecole Royale Polytechnique. 1re Partie: Analyse Algébrique* (1821, Paris) = *Oeuvres*, (2) **3**. No further parts of this work were published: it is referred to in later footnotes as *Cours*.

[3] B. BOLZANO, "Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege," (1817, Prague) = *Abh. Königl. Böhm. Gesell. Wiss.*, (3) **5** (1814—17: publ. 1818), 60 pp. = Ostwald's Klassiker, No. 153 (ed. P. JOURDAIN: 1905, Leipzig), 3—43. French trans. in *Rev. d'Hist. Sci. Appl.*, **17** (1964), 136—164: there have also been various other translations and issues. The paper is referred to in later footnotes as *Beweis*.

term "continuous" was usually confined to functions which we now call "differentiable".[4] There were efforts to move away from this view, including by EULER himself; but nobody had come at all close to the formulation of continuity given by BOLZANO and CAUCHY:

BOLZANO: "A function $f(x)$ varies according to the law of continuity for all values of x which lie inside or outside certain limits, is nothing other than this: if x is any such value, the difference $f(x+\omega)-f(x)$ can be made smaller than any given quantity, if one makes ω as small as one ever wants to."[5]

CAUCHY: "The function $f(x)$ will remain continuous with respect to x between the given limits, if between these limits an infinitely small increase of the variable always produces an infinitely small increase of the function itself".[6]

One of the most interesting and important features of this formulation of continuity is that it extends the old formulation beyond that of differentiability, for it also encompasses functions with corners. I think that BOLZANO was aware of the extension in 1817, for in later manuscripts he studied the distinction between the new continuity and differentiability to the extent of constructing a continuous non-differentiable function of the type studied later only by the school of WEIERSTRASS in the 1870's.[7] But CAUCHY seems to have seen the new idea only as a reformulation of the old one when he wrote the *Cours d'Analyse*, for the examples he gave there of continuous functions were all of standard differentiable algebraic expressions, with the functions x^a for negative a, and $\frac{a}{x}$, regarded as "discontinuous" at $x=0$ since they then became infinite.[8] In fact, he explicitly discussed the distinction only in a paper of 1844, and then in a way which tried to give the impression that he had known it all along:

"In the works of Euler and Lagrange, a function is called *continuous* or *discontinuous*, according as the diverse values of that function, corresponding to diverse values of the variable ... are or are not produced by one and the same equation .... Nevertheless the definition that we have just recalled is far from offering mathematical precision; for the analytical laws to which functions can be subjected are generally expressed by algebraic or transcendental formulae [that is, by the EULERIAN range of algebraic expressions], and it can happen that various formulae represent, for certain values of a variable x, the same function: then, for other values of x, different functions."

He then quoted the example

$$\sqrt{x^2}=\frac{2}{\pi}\int_0^\infty \frac{x^2}{t^2+x^2}\,dt=\begin{cases} x \text{ if } x\geqq 0\\ -x \text{ if } x\leqq 0\end{cases}, \qquad (1)$$

[4] EULER's classic presentation of his theory of functions was given in the opening sections of both volumes of his *Introductio ad analysin infinitorum* (2 vols: 1748, Lausanne) = *Opera Omnia*, (1) **8**—9.

[5] B. BOLZANO, *Beweis*, preface, part IIa.

[6] A.-L. CAUCHY, *Cours*, 34—35 = *Oeuvres*, (2) **3**, 43.

[7] See B. BOLZANO, *Functionenlehre* (ed. K. RYCHLIK), in his *Schriften*, **1** (1930, Prague), esp. pp. 66—70, 88—89.

[8] A.-L. CAUCHY, *Cours*, 36—37 = *Oeuvres*, (2) **3**, 44—45.

in which the first two forms are "continuous" in EULER's sense while the third is "discontinuous";

"... but the indeterminacy ceases if for Euler's definition we substitute that which I have given [in the *Cours d'Analyse*]".[9]

2.2. Convergence of a Series. A major innovation of the new analysis was the study of the convergence of a series (or of classes of series) as a general problem separate from and indeed prior to that of its summation; but it *would* be wrong to presume that the problem of convergence had previously been ignored or taken for granted. 17^{th} and 18^{th} century mathematicians were perfectly well aware that a series was to be interpreted as a term-by-term addition of its members, and that individual series (usually series of constant terms or certain power series) could be shown to be convergent, especially if they were associated with some geometrical limiting procedure such as the approximation to a curve by a polygon. But this understanding had been endangered during the 18^{th} century, especially by EULER's great ability to devise complicated new methods of summation of series. Today we understand that some of these methods reduce to orthodox summation for orthodox convergent series and some do not; but EULER and his contemporaries seemed to have regarded *all* methods as legitimate, giving "the" sum of the series rather than its sum relative to the method of summation involved. This more sophisticated understanding began to develop only in the 1890's, under the leadership of BOREL:[10] until then, series considered "divergent" (that is, oscillatory series as well as those with an infinite sum) had been banished from analysis under the influence of CAUCHY's work. But he and BOLZANO were not the first to consider the convergence of a series to be an important property worthy of investigation of its own. GAUSS had even advanced as far as a sophisticated convergence test by 1812[11]: FOURIER had already treated the convergence of particular examples of his series in 1807, in his first paper on the diffusion of heat[12]: LAGRANGE had tried to find expressions for the remainder term of a TAYLOR series, in connection with his long held belief that the series could serve as the foundation of the calculus;[13] and LACROIX was also aware of the need for general formulation of convergence.[14] Both BOLZANO and CAUCHY also stressed that the convergence of a series is to be determined only by the tendency of the n^{th} partial sums to a limiting value s as n tended to infinity;[15]

[9] A.-L. CAUCHY, "Mémoire sur les fonctions continues ou discontinues", *C. R. Acad. Roy. Sci.*, **18** (1844), 116—130 (see pp. 116—117) = *Oeuvres*, (1) **8**, 145—160 (pp. 145—146).

[10] For extended discussion, see my *Foundations*, ch. 4.

[11] K. F. GAUSS, "Disquisites generales ...", *Comm. Soc. Reg. Sci. Göttingen Rec.*, **2** (1811—13: publ. 1813), cl. math., 46 pp. = *Werke*, **3**, 123—162: see art. 16. For a history of convergence tests, see the appendix to my *Foundations*.

[12] J. B. J. FOURIER, "Sur la propagation de la chaleur," MS. 1851, *Ecole Nationale des Ponts et Chaussées*, Paris: see arts. 42—43. The publication of this entire manuscript constitutes the body of my *Fourier*: see there ch. 7 on this point.

[13] See especially his *Théorie des fonctions analytiques* ..., (2nd edition: 1813, Paris) = *Oeuvres*, **9**: part 1, arts. 35—40.

[14] See especially his *Traité du calcul différentiel et du calcul intégral* (1st edition: 1797—1800, Paris), **1**, 4—9.

[15] B. BOLZANO, *Beweis*, art. 5. A.-L. CAUCHY, *Cours*, 123—125 = *Oeuvres*, (2) **3**, 114—115.

thus this correspondence is not so striking, although the idea was still then very much a new one. But in both works we find a new type of result, not to be found in any other contemporary writing. BOLZANO had defined a class of series:

"... which possess the property that the variation (increase or decrease) which their value suffers through a prolongation [of terms] as far as desired remains always smaller than a certain value, which again can be taken as small as one wishes, if one has already prolonged the series sufficiently far",[16] and then he proved that for series with this property,

"... there always exists a certain constant value, and certainly only one, which the terms of this series always approach the more, and towards which they can come as close as desired, if one prolongs the series sufficiently far."[17] CAUCHY stated that:

"For the series $\left[\sum_{r=1}^{\infty} u_r\right]$ to be convergent it is yet necessary that for increasing values of n the different sums

$$\begin{aligned}&u_n + u_{n+1}\\&u_n + u_{n+1} + u_{n+2}\\&\&c. \quad . \quad . \quad . \quad .\end{aligned}$$

... finish by constantly achieving numerical values smaller than any assignable limit. Reciprocally, when these various conditions are fulfilled, the convergence of the series is assured."[18]

In other words, they both found a general condition for convergence in terms of the behaviour of $(s_{n+r} - s_n)$ as n tended to infinity: a result of quite profound originality. Contrary to general belief, BOLZANO in fact only asserted the sufficiency of the condition in his paper; his proof is very difficult to follow even with the ideas of his new analysis, and in fact is faulty. The necessity of the condition is far easier to recognise and prove: CAUCHY did prove it, but then avoided difficulties by hinting that sufficiency followed as a consequence (which it does not!):

"the sums s_n, s_{n+1}, ... differ from the limit s, *and consequently among themselves*, by infinitely small quantities."[19]

2.3. Bolzano's Main Theorem. The theorem which BOLZANO actually proved in his paper was the following generalisation of the theorem of his title:

Let $f_1(x)$ and $f_2(x)$ be continuous functions for which $f_1(\alpha) < f_2(\alpha)$ and $f_1(\beta) > f_2(\beta)$: then $f_1(a) = f_2(a)$ for at least one value a of x between α and β. (The basic theorem is the case where $f_2(x) \equiv 0$.)

As a theorem it is most untypical of its time: that is, a general theorem concerning the properties of functions was *not* the kind of result then being sought in analysis. BOLZANO himself saw it rather as a theorem in the theory of equations, as a companion to GAUSS's recent proofs of the decomposition of a polynomial

---

[16] B. BOLZANO, *Beweis*, art. 5.

[17] B. BOLZANO, *Beweis*, art. 7.

[18] A.-L. CAUCHY, *Cours*, 124—125 = *Oeuvres*, (2) **3**, 115—116.

[19] A.-L. CAUCHY, *Cours*, 125 = *Oeuvres*, (2) **3**, 115. My italics.

into linear and quadratic factors.[20] CAUCHY saw it as a theorem of the new analysis, and put it *twice* into the *Cours d'Analyse* (in its restricted form): firstly with a naive geometrical argument, and later, in the part of his book reserved for those with a special interest in analysis, with a condensation argument which seems very much like an unrigorous version of the intricate proof developed in BOLZANO's paper.[21]

2.4. Bolzano's Lemma. A crucial lemma required by BOLZANO to establish the existence of the real root was the following lemma:

"If a property M does not apply to all values of a variable quantity x, but to all those which are smaller than a certain u: so there is always a quantity U which is the largest of those of which it can be asserted that all smaller x possess the property M."[22]

With this extraordinary theorem came another new idea into analysis, completely untypical of its time: the *upper limit* of a sequence of values. It is not to be found explicitly in CAUCHY's *Cours d'Analyse*, but instead we have there a frequent use of phrases like "... the largest value of the expression ..." when calculating limiting values, especially in connection with the development of tests for convergence of a series.[23] As with continuity of a function, CAUCHY was revealingly only partially aware of the significance of the idea; for he used it only as a tool for developing the proofs of his particular theorems and not as a profound device for investigating more sophisticated properties of analysis. Therefore it would be especially surprising if it were CAUCHY's own invention: not until the 1860's was it introduced again and properly used, by the WEIERSTRASS school of analysts.[24]

2.5. The Real Number System. Lastly, a point which is less striking than the others but worth mentioning: the considerations given in both works to the real numbers. In the course of proving his lemma as well as in other parts of his paper BOLZANO had recourse to extended considerations of real numbers, especially regarding the rational or irrational limiting values of sequences of certain finite series of rationals.[25] In later manuscripts he extended these remarks into a full theory of rational and irrational numbers of the type which, like continuous non-differentiable functions and the theorem on upper limits, was next investigated

[20] K. F. GAUSS, "Demonstratio nova altera ..." and "Theorematis de resolubitate ...", *Comm. Soc. Reg. Sci. Göttingen Rec.*, **3** (1814—15: publ. 1816), cl. math., 107—134, and 135—142 = *Werke*, **3**, 31—56, and 57—64.

[21] A.-L. CAUCHY, *Cours*, 43—44 and 460—462 = *Oeuvres*, (2) **3**, 50—51 and 378—380.

[22] B. BOLZANO, *Beweis*, art. 12.

[23] See especially the sections on convergence tests in chs. 6 and 9 of the *Cours*.

[24] There is a distinction between BOLZANO's introduction of an upper limit and CAUCHY's "largest value of the expression ...", in that CAUCHY actually used the *Limes* of a sequence (whose every neighbourhood contains members of the sequence), while BOLZANO defined the upper limit (which does not *necessarily* have this property); but we cannot interpret this distinction as intentional in BOLZANO and CAUCHY's time and I do not know of any recorded awareness of it then. For a brief discussion of the point, see P. E. B. JOURDAIN, "On the general theory of functions," *Journ. rei. ang. Math.*, **128** (1905), 169—210 (pp. 185—188).

[25] B. BOLZANO, *Beweis*, art. 8: see also art. 12.

only by WEIERSTRASS and his followers.[26] CAUCHY wrote just once on the real number system: it was in the *Cours d'Analyse*, where he gave a superficial formal exposition of the real number system. The initial stimulus for this work was foundational questions concerning the representation of complex numbers; but he took the development of the ideas well into BOLZANO's territory, twice including the remark that "when B is an irrational number, one can obtain it by rational numbers with values which are brought nearer and nearer to it"[27] — merely a remark on a property of the real numbers and not as a *definition* of the irrational number in the sense of the later work, as has sometimes been thought. Once again CAUCHY did not fully appreciate the depth of BOLZANO's thought; and yet it is clear from his partial success that he was aware of BOLZANO's ideas, rather than from his partial failure that he was ignorant of them. The striking feature of this remark, as with his interpretation of continuity and his only incomplete use of the upper limit, is that it is there *at all*, rather than that it appears in a mutilated form.

3. The New Analysis

Thus we find a significant collection of unusual results in the two works: yet there is a much stronger and more profound link between them, which cannot be identified by means of precise quotations or references — namely, a *unity of approach*. We have here a good example of the rule that the whole is greater than the sum of the parts, for it is the *homogeneity* and *general applicability* of these new ideas which is their most significant feature. The term "arithmeticisation of analysis" is given to them, because they operate by means of arithmetical differences and proofs within the analysis are based on the arithmetical manipulation of them; but I do not favour this name, partly because it is identified with the later WEIERSTRASSIAN developments of analysis but principally because the arithmeticisation is only at the service of something more profound: the theory of *limit-avoidance*.

When we speak of "introducing the concept of a limit" into analysis, we are actually introducing limit-avoidance, where the limiting value is *defined* by the property that the values in a sequence avoid that limit by an arbitrarily small amount when the corresponding parameter (the index n for the sequence s_n of n^{th} partial sums, say, or the increment α in the difference $\big(f(x+\alpha)-f(x)\big)$ for continuity) avoids its own limiting value (infinity and zero, in these examples). The new analysis formed in BOLZANO's pamphlet and developed in CAUCHY's text-books was nothing else than a complete reformulation of the whole of analysis in limit-avoidance terms, terms which CAUCHY made quite explicit in the introduction to the *Cours d'analyse*:

"When the values successively attributed to a particular variable approach indefinitely a fixed value, so as to finish by differing from it by as little as one wishes, this latter is called the *limit* of all the others."[28]

---

[26] These manuscripts were published in K. RYCHLIK (ed.), *Theorie der reellen Zahlen im Bolzanos handschriftlichen Nachlasse* (1962, Prague).

[27] A.-L. CAUCHY, *Cours*, 409 and 415 = *Oeuvres*, (2) **3**, 337 and 341.

[28] A.-L. CAUCHY, *Cours*, 4 = *Oeuvres*, (2) **3**, 19.

One important aspect of limit-avoidance is that it is *independent* of the continuum of values over which the analysis is conducted. Limit-avoidance can be developed whether an infinitesimal or non-infinitesimal field is being used: the use of the WEIERSTRASSIAN term "arithmeticisation of analysis", applied to the period when WEIERSTRASS excluded infinitesimals from analysis, has led us to forget that its limit-avoiding character was shown also by the earlier period instigated by BOLZANO, who used both types of continuum in his analysis,[29] and CAUCHY, who practiced only infinitesimals throughout his mathematical career. Since WEIERSTRASS's time, we have held a fairly contemptuous view of the infinitesimalists which I regard as unfair. A remarkable amount of pure and applied analysis was developed from the time of NEWTON onwards with the aid of infinitesimals; but there were important foundational difficulties involved in their use, and in fact CAUCHY is a good example of them. These difficulties seem to me to lie especially in the foundations of the calculus, which if we examine from the point of view of limit-avoidance also reveal the attraction that infinitesimals must have had to the founders of the algebraic calculus.

We make our point in the LEIBNIZIAN notation, which not only became the standard system but also contained a key to the difficulties that the infinitesimalists faced. When we calculate the derivative by means of the definition

$$\frac{dy}{dx} =_{\text{Df.}} \lim_{h\to 0}\left[\frac{f(x+h)-f(x)}{h}\right], \tag{2}$$

we may quite easily obtain the value of the derivative involved; but we are left with the important foundational question of *how* that value is obtained in light of the fact that the ratio on the right hand side of (2) becomes $\frac{0}{0}$ when $h=0$. The virtue of infinitesimals, quantities which obeyed the law

$$a+h=a \tag{3}$$

of addition to the "ordinary" numbers, was that, being non-zero they avoided the limiting value and therefore the difficulty of $\frac{0}{0}$; on the other hand, being smaller than "any assignable quantity" (that is, any non-infinitesimal), they effectively allowed the limit to be taken. This view was of course an inconsistent one, but I think that it lay basically behind infinitesimalist reasoning and was the source of its difficulties. The infinitesimal was either zero or non-zero, according to the needs of the moment: thus it could be added to or withdrawn from any quantity in an equation, with the presumed certainty of leaving the mathematical situation described by that equation undisturbed. We may see this as a double-interpretation for the infinitesimal — a limit-avoiding interpretation as a non-zero quantity, and what we may call by contrast a "limit-achieving" interpretation as an essentially zero quantity allowing the limit to be taken. From this distinction there follows a corresponding double-inter-

[29] In the *Beweis* BOLZANO did not explicitly discuss the possible continua, and seemed to have allowed the use of infinitesimals; but later in the year he published another pamphlet, on *Die drei Probleme der Rectification, der Complanation und die Cubirung, ohne Betrachtung des unendlich Kleinen, ... und ohne irgend eine nicht streng erweisliche Voraussetzung gelöst;* ... (1817, Prague) = *Schriften,* 5 (1948, Prague), 67—138.

pretation of $\frac{dy}{dx}$. Let us take a specific example of a derivative, say for the function

$$y = x^3, \tag{3}$$

whose derivative

$$\frac{dy}{dx} = 3x^2 \tag{4}$$

is calculated from

$$\frac{dy}{dx} =_{\text{Df.}} \lim_{h\to 0}\left[\frac{(x+h)^3 - x^3}{h}\right]. \tag{5}$$

When h achieves its limiting value zero (4) gives us the value of the derivative, and so the denoting symbol $\frac{dy}{dx}$ is in fact just a symbol and is *not* to be taken as an arithmetical ratio "$dy \div dx$". Thus it is not valid to multiply through (4) by dx to obtain

$$dy = 3x^2 dx. \tag{6}$$

(6) follows from (4) by turning from the limit-achieving interpretation of $\frac{dy}{dx}$ as a whole symbol to its limit-avoiding interpretation, where it *is* the ratio "$dy \div dx$". For if we avoid the limiting value by the non-zero infinitesimal quantity dx, then we see from the right hand side of (5) that the situation for the increment dy $\left(= d(x^3)\right)$ is given by

$$dy = 3x^2 dx + q, \tag{7}$$

where q is a *second-order* infinitesimal obeying the law

$$a + q = a \tag{8}$$

of addition to "ordinary" or first-order infinitesimal quantities a. $\frac{dy}{dx}$ in this kind of situation, if we wish to consider it, could arise by dividing throughout (7) to give:

$$\frac{dy}{dx} = 3x^2 + \frac{q}{dx}, \tag{9}$$

a result of a *fundamentally different kind from* (4). There is a difference between the two far greater than the first order infinitesimal $\frac{q}{dx}$: we see a basic *qualitative* difference, for $\frac{dy}{dx}$ appears in (4) as a limit-achieving symbol but in (9) as a limit-avoiding ratio. Further, the deduction of (9) from an infinitesimal equation (7) is not necessary to the derivation of (4). For let us suppose that we change continua so that in WEIERSTRASSIAN style we reject the use of infinitesimals. Then (4) and (5) still stand (with the limit now of course taken over the non-infinitesimal field); but (7) and all its consequences, such as (9), disappear altogether for (7) itself changes into the identity

$$0 = 0, \tag{10}$$

whether or not it was true in the infinitesimal continuum.

The ideas that I have presented here are essentially straightforward, and are susceptible of considerable extension; but they are independent of the modern interest in developing a *consistent* theory of infinitesimals.[30] They do not themselves establish a consistent infinitesimalism but at least show that much can be clarified in terms which could have been understood and developed in the infinitesimalist period. Yet they were far from the considerations of the time: in particular, CAUCHY's treatment of the foundations of the calculus was as incoherent and incompetent as any that were ever offered. In his *Résumé des leçons ... sur le calcul infinitésimal* of 1823, the next instalment of his new analysis after the *Cours d'Analyse*, he explicitly rejected LAGRANGE's faith in TAYLOR's series, but he replaced it with an extraordinary theory of the derivative which made simultaneous use of both LAGRANGE's theory of derived functions $f'(x)$, $f''(x)$, ... and also of CARNOT's theory of differentials dx, ddx, ...: infinitesimals not only achieved the limit in CAUCHY's system but they also avoided it, at times by non-infinitesimal amounts, changing their role with every appearance of new and usually unnecessary notation.[31] However, when CAUCHY came to integration he was wonderfully successful, laying out the whole basic structure of the theory of the "CAUCHY integral" (defined in terms of the area as the limit of a sum) in a masterly display of the power of the new analysis of limit-avoidance.

This is what the new analysis was: only in limit-avoidance terms can its full power and subtlety be appreciated, and theorems such as the necessary and sufficient condition for convergence in the diminishing of $(s_{n+r} - s_n)$ — *where the limit s is avoided altogether* — and BOLZANO's theorem on the existence of upper limits, can be seen to their best advantage. Yet to understand BOLZANO and CAUCHY's work we must look at the old as well as the new. What sort of analysis had they replaced?

4. The Old Analysis

We have referred earlier briefly to certain features of 18^{th} century analysis, and it is appropriate now to make more detailed remarks about its character. In speaking of the "old analysis", we are referring only to the subject immediately prior to BOLZANO and CAUCHY's work; and we find that many of its features were the result of problems in other areas of mathematics, especially in the solution of difference and differential equations. Following the leadership of EULER, his contemporaries (mainly D'ALEMBERT, DANIEL BERNOULLI and LAGRANGE) and successors (mainly LAGRANGE, LAPLACE and MONGE) had developed a wide range of solution methods. It is impossible to describe them all in a sentence, but often they involved the construction of exact differentials prior to integration to give functional solutions, or assumptions of particular kinds of solution which led via the conditions of the problem to auxiliary equa-

[30] See A. ROBINSON, *Non-Standard Analysis* (1966. Amsterdam); and also the work initiated by C. SCHMEIDEN & D. LAUGWITZ, "Eine Erweiterung der Infinitesimalrechnung", *Math. Zeitschr.*, **69** (1958), 1—39.

[31] A.-L. CAUCHY, *Résumé des leçons données à l'Ecole Royale Polytechnique sur le calcul infinitésimal. Tome premier* (1823, Paris) = *Oeuvres*, (2) **4**, 5—261. No other volumes were published: see here lecture 5.

tions. The analytical techniques themselves — which involved not only differentiation and integration, but also summation and rearrangement of series (especially power series), manipulations of algebraic expressions, the taking of limiting cases (in moving from difference to differential equations for example), and so on — were normally used as required without consideration of their validity. This is not intended as a criticism, but merely a general statement of the situation: it led to an enormous range of results in pure and applied mathematics which have remained important ever since. Further, there were cases when questions of rigour and validity *did* arise, of which the most important was the problem of the motion of the vibrating string;[32] but in general the situation at the beginning of the 19th century was that not only were such considerations relatively limited but the techniques themselves were susceptible of, and received, plenty of further development without concern for the rigour involved. This is a matter of great importance when considering the "new analysis" of BOLZANO and CAUCHY. Their new foundations, based on limit avoidance, certainly swept away the old foundations, founded largely on faith in the formal techniques; but it would be a mistake of posterior wisdom to assume that old foundations had been in a serious and comprehensive state of decay and were recognised as such by those who were using them. Historians of science seem to be only too ready to make assumptions of this kind when considering "revolutions" in science: they also tend to identify anticipations of a new system in the old one with that new system instead of what they probably were, something else in the old system which was quite different and also interesting. The historiographical point here is the danger of determinism; that because a body of knowledge developed in a particular way, then it must be viewed historically as having been capable of developing *only* that way, certainly from the intellectual point of view and perhaps even chronologically. Yet in fact any situation is always open to a variety of future developments: we must not allow the intermediate historical processes that actually happened to distort our vision of the situation from which they started.

I have already claimed that the new analysis replaced an old analysis which does not seem to have needed such a radical replacement: from the point of view of the BOLZANO-CAUCHY question, it follows that it is all the more surprising that exactly the same type of replacement began to emerge twice within four years. But we must consider also the anticipations of the new system in the old one. The "new analysis" laid great stress on the rigour of processes: did no "old analyst" try to do the same? Yes, certainly, but not in any way resembling the comprehensive and homogeneous character of the new method: they had other ideas which were quite different and also interesting. EULER tried hard, though with little practical success, to produce a consistent infinitesimalism in his "reckoning with zeros", including consideration of different orders of infinitesimal. D'ALEMBERT tended to distrust infinitesimals altogether, while LAGRANGE tried to avoid all limiting processes by *defining* the derivatives of a function in

[32] For a discussion of foundational questions in the light of this problem, see my *Foundations*, Ch. 1; and for an extended account of the solution of differential equations in this period, see C. TRUESDELL, *The rational mechanics of flexible or elastic bodies 1638—1788*, L. Euleri Opera Omnia, (1) **11**, pt. 2 (1960, Zurich).

terms of the coefficients of its expansion as a TAYLOR series. This was "limit-avoidance" of a completely different and considerably less successful kind, and it won few supporters. One of them, however, was ARBOGAST, who tried towards the end of the century to reduce the number of distinctions between types of function to a group based on analytical rather than algebraic or mechanical considerations. L'HUILIER offered a thoughtful essay on the taking of limits: I am sure that CAUCHY read it, for he always used the notation "lim" for a limiting value which L'HUILIER introduced there. But I doubt if he learnt much more from it, for the results obtained are severely limited, being concentrated on the derivative and often providing no more than a re-writing of known ideas. L'HUILIER also criticised (with justice) EULER's use of infinitesimals, and CARNOT took it further into a profound essay on orders of the infinitely small and the interpretation of the LEIBNIZIAN notations as infinitesimals. But perhaps the best example, especially from the point of view of anticipations of BOLZANO and CAUCHY, is LACROIX, the principal text-book writer of the day. He was not an important creative mathematician, but he was capable of some measure of appreciation of contemporary work and he read exhaustively among the earlier literature. I referred earlier to his understanding of convergence of series as a general problem, which he learnt from D'ALEMBERT's vague warnings against divergent series in the 1760's: he also gave in 1806 a formulation of continuity vaguely similar to that of BOLZANO and CAUCHY.[33] Thus we may say that LACROIX anticipated them if we wish; yet it would be more misleading than illuminating to do so, not least to the understanding of LACROIX's results. For one cannot find in LACROIX's writings the general aim that BOLZANO and CAUCHY achieved, not even in the new editions of his works that continued to appear after CAUCHY's text-books were published.

What would have happened if CAUCHY had *not* read BOLZANO? Without doubt, foundational questions would have received discussion, but it seems to me most unlikely that the radical reform that in fact happened would have taken place: rather only parts of that theory would probably have emerged, especially in the convergence of series and the integral as the limit of a sum, while the rest, apparently sound enough, would have received well-meaning but limited examination. But in order to put the old and the new analyses into better perspective we must describe some of the fundamental problems which were current before BOLZANO's paper; and at the same time we shall pass on to further aspects of the CAUCHY-BOLZANO question, aspects which involve not only analysis itself but also the Paris in which CAUCHY was working and the way in which his mathematical genius was inspired.

---

[33] S. F. LACROIX, *Traité élémentaire du calcul intégral* (2nd edition; 1806, Paris): see art. 60. The other works to which we referred explicitly were L. F. A. ARBOGAST, *Mémoire sur la nature des fonctions arbitraires qui entrent dans les intégrales des équations aux différentielles partielles* (1791, St. Petersburg): S. L'HUILIER, *Exposition élémentaire des principes des calculs supérieures* (1786, Berlin), esp. chs. 1 and 11; and L. N. M. CARNOT, *Reflexions sur la métaphysique du calcul infinitésimale* (1st edition: 1797, Paris. 2nd edition: 1813, Paris). On EULER's and LAGRANGE's views on analysis, see A. P. JUSCHKEWITSCH, "Euler and Lagrange über die Grundlagen der Analysis," *Sammelband der zu Ehren des 250. Geburtstages Leonhard Eulers* (ed. K. SCHRODER: 1959, Berlin), 224—244; and on all these and other developments, my *Foundations*, chs. 1 and 3.

5. Cauchy's Originality as a Mathematician

If CAUCHY came to his new ideas independently of BOLZANO, then he perceived a completely novel approach to analysis and detected its superiority over known techniques which themselves were not lacking in power or generality. This kind of achievement is characteristic of certain mathematicians: it reflects their sensitive "intuition for problems", their ability to see far beyond contemporary work into totally new ways of solving current problems, or even of forming new problems of which others were hardly aware. GAUSS is a prime example of such a thinker, with his notebooks already filled with the seeds of most 19th century mathematics within its first decade: BOLZANO shows this ability, too, and to the extent that he was in fact extremely limited in ability at "orthodox" developments of current and popular methods. Thus in 1816, for example, before the flood of his own new thinking, he published a treatise on the binomial series in the style of the old analysis which is really quite remarkably uninteresting.[34] But CAUCHY is a good example of originality of another kind, lacking such sensitivity and feeling for new problems but, when stimulated by the achievements or especially *lack of success* in some contemporary work, would expand the accomplished fragments into immense generalisations and extensions within the same field of research. His monument in mathematics in his theory of functions of a complex variable and their integration, one of the great achievements of all 19th century mathematics. Its origins are to be found in a large paper of 1814 (his 25th year) on the validity of using complex numbers in the evaluation of definite integrals. The technique had been used for decades from time to time, without much consideration of its validity: in particular, in June 1814, LEGENDRE published an instalment of the second volume of his *Exercises du calcul intégral*, a work containing various methods of evaluating definite integrals whose main aim was towards the development of his theory of elliptic integrals.[35] This instalment concerned itself chiefly with integrals whose integrands were the product of rational and trigonometric functions, and it provided the spark for CAUCHY's fire, for from LEGENDRE's work CAUCHY came to the following generalised problem concerning the evaluation of definite integrals: what are sufficient conditions for the validity of using complex variables in such evaluations? His solution was the equality of two mixed partial differentials:

$$\frac{\partial^2}{\partial x\, \partial y}\int f(z)\, dz = \frac{\partial^2}{\partial y\, \partial x}\int f(z)\, dz, \tag{11}$$

where z is a complex function of x and y;

$$z = h(x, y) + i k(x, y) \tag{12}$$

and thus

$$f(z) = u(x, y) + i v(x, y). \tag{13}$$

From this fruitful equation (11) stemmed a variety of general theorems (including the "Cauchy-Riemann equations") and thence hosts of particular integrals,

[34] B. BOLZANO, *Der binomische Lehrsatz und als Folgerung aus ihm der Polynomische und die Reihen, ...* (1816, Prague). The most interesting section is on pp. 27—40.

[35] A.-M. LEGENDRE, *Exercises du calcul intégral sur divers ordres des nombres transcendantes et sur les quadratures* (3 vols: 1811—17, Paris).

including the evaluation of some of LEGENDRE'S. CAUCHY presented his paper in August (1814) to the *Institut de France*, and LEGENDRE was one of its examiners: he rightly praised its many important new results, but had a most interesting and important dispute with CAUCHY over the evaluation of

$$\int_0^\infty \frac{x\cos ax}{\sin bx}\,\frac{dx}{1+x^2}. \tag{14}$$

Put in modern terms, if we regard the integral as a function of $\frac{a}{b}$ then it has a discontinuity of magnitude π at the odd multiple values of its argument. CAUCHY had by separate equations evaluated the left- and right-hand limiting values of the function for $\frac{a}{b}<1$ and $\frac{a}{b}>1$; but in the 1814 instalment of his book LEGENDRE had used a power series expansion method on a generalisation of (14) to produce in a limiting case the *arithmetic mean* of CAUCHY's two evaluations for $\frac{a}{b}=1$, and he could not understand that this new type of algebraic expression — the integral representation — could in fact give a *discontinuous* function. CAUCHY produced a spurious piece of infinitesimal reasoning to resolve the situation to LEGENDRE's satisfaction;[36] but it must have shown him that there were foundational questions in real variable analysis apart from the use of complex numbers with which he would have to deal.

Let us return, however, to the question of CAUCHY's type of mathematical inspiration. We see in this episode that CAUCHY was directly stimulated by LEGENDRE's attempts at integral evaluation to work in exactly the same field, rather than to intuite from it some more general and abstract kind of problem concerned with the use of functions of a complex variable. In the 1814 paper for example, the theory of singularities and residues which he was to produce in later years was given in a *real variable* integral form, which we may write as:

$$\int_{y_1}^{y_2}\int_{x_1}^{x_2}\frac{\partial S}{\partial x}\,dx\,dy-\int_{x_1}^{x_2}\int_{y_1}^{y_2}\frac{\partial S}{\partial x}\,dy\,dx=\int_0^\varepsilon [S(X+p,Y+q) \\ -S(X+p,Y-q)-S(X-p,Y+q)+S(X-p,Y-q)]\,dp, \tag{15}$$

where $\frac{\partial S}{\partial x}$ has an infinity at the point (X, Y) inside the rectangle bounded by the sides, $x=x_1$, $x=x_2$, $y=y_1$ and $y=y_2$.[37] His later fine achievements in the new analysis with the theory of integration may be traced in large part to the issues involved in the profound result (15).

In the following year of 1815 CAUCHY had another large paper ready, this time on the propagation of water-waves.[38] Complex variables were again present,

[36] For a full account of this episode see my *Foundations*, ch. 2.

[37] CAUCHY's paper was "Mémoire sur les intégrales définies", *Mém. prés. Acad. Roy. Sci. div. sav.*, (2) **1** (1827), 601—799 = *Oeuvres*, (1) **1**, 319—506. LEGENDRE's evaluation of the integral (14) is in his [35], **2**, 124.

[38] A.-L. CAUCHY, "Théorie de la propagation des ondes à la surface d'un fluide pesant d'une profondeur indéfinie," *Mém. prés. Acad. Roy. Sci. div. sav.*, (2) **1** (1827), 3—312 = *Oeuvres*, (1) **1**, 4—318.

as they were to be in all of his mathematical output; and integrals were also to be found, for the prominent new feature here was the use of integral methods to solve linear partial differential equations (and thus to use again the integral representation of a function). The inspiration in this case is not so easy to trace, as it is impossible to say how much of FOURIER's then still unpublished work on heat diffusion he had seen; but he knew of POISSON's (lesser) work in the same field, and doubtless he was aware of some results of LAPLACE which we shall discuss later. At all events, in 1817 his further researches brought him to "Fourier's Integral Theorem":

$$f(x) = \frac{1}{\pi} \int_0^\infty f(p)\, dp \int_0^\infty \cos q(x-p)\, dq, \tag{16}$$

in a short paper whose rushed and excited tone suggests that he had really found the result independently of FOURIER.[39] FOURIER acquainted him with his own prior discovery of the theorem, and then CAUCHY certainly did read his manuscripts: not only did he publish an acknowledgement in 1818,[40] but in all his later work on integral solutions to partial differential equations there was a new confidence and dexterity, and again — *extensions* and *generalisations* (to multiple integral solutions, and so on) of what FOURIER had already done.[41]

And then we come to 1821 and the *Cours d'Analyse*: large numbers of theorems on all aspects of real and complex variable function theory, based on the ideas which we listed in our section 2. From where had the inspiration come this time? From within CAUCHY himself? Perhaps; but it is so utterly untypical of his kind of achievement whereas under the hypothesis of his prior reading of BOLZANO it is so perfect an example of it, that it seems difficult not to accept the latter possibility. Perhaps I can best illustrate the force of this point by describing my own researches into the development of the foundations of analysis during this period. I had started naturally enough with CAUCHY's *Cours d'Analyse* and his other contributions to analysis, and in the course of reading other of his writings his need for an initial external stimulus to his genius had become clear to me. Thus I wanted to find the source of the new ideas of the *Cours d'Analyse*, and so I made a special search of all of CAUCHY's work written prior to 1821. I found many important things, especially the 1814 integrals paper and the disagreement over (14) with LEGENDRE, and the affair of 1817 over FOURIER's Integral Theorem (16): there was clearly plenty of motivation for CAUCHY to try to improve analytical techniques. But of the new ideas that were to achieve that aim — of them, to my great surprise, I could find nothing. Only later did I follow up my knowledge that BOLZANO had done "something" in analysis which no-one had read (or so I thought); and I can remember quite clearly the extraordinary effect of reading BOLZANO's 1817 pamphlet and seeing the *Cours d'Analyse* emerging from its pages. I then re-read the *Cours d'Analyse* and found the fine details of

[39] A.-L. CAUCHY, "Sur une loi de réciprocité qui existe entre certaines fonctions", *Bull. Sci. Soc. Philom. Paris* (1817), 121—124 = *Oeuvres*, (2) 2, 223—227.

[40] A.-L. CAUCHY, "Second note sur les fonctions réciproques", *Bull. Sci. Soc. Philom. Paris* (1818), 178—181 = *Oeuvres*, (2) 2, 228—232.

[41] For discussion of these developments, see my *Fourier*, chs. 21 and 22; and BURKHARDT[33], chs. 8—11 *passim*.

correspondence; but more than that, I could see CAUCHY's mind at work in its own individual way, taking the fragments of BOLZANO's thought as he had taken LEGENDRE's morsels and FOURIER's substantial achievements earlier, and producing from them whole new systems of mathematical thought.

But if CAUCHY owed so much to BOLZANO, why did he not acknowledge him? To answer this question, we move more fully into the social situation of the time: to Paris, the centre of the mathematical world.

6. The State of Parisian Mathematics

Almost every mathematician of note at this time either lived in or at least visited Paris. One consequence of this galaxy of brilliance was that a state of intense rivalry and sometimes bitter enmity existed almost continuously in the Parisian scientific circles. Everybody was affected by it, although some less than others; and the reasons were not always purely scientific. There were deep and passionate political or religious disagreements, too, heightened by the Napoleonic era and its violent end and brief resurrection in the mid-1810's. These rivalries pose an exciting and difficult problem for the historian of the period, for their detection and description calls for the most careful reading of even the finest point in the most obscure paper, as well as reading *between* the lines of all the scientific literature of the time. Very little work has been done on these rivalries: indeed, most historians have failed to notice them altogether.[42] But perhaps I can give some idea of how they affected the situation and bore especially upon CAUCHY and his *Cours d'Analyse* by describing two of the most important controversies of the time — as fully as I have been able to disclose them.

We have mentioned FOURIER's name several times, and the first controversy involved his work on heat diffusion. Like GAUSS and BOLZANO, he also had a strong intuition for new problems, and seemingly from about 1802 he began work on the then novel study of the mathematical description of the diffusion of heat in continuous bodies. His early work on the problem proceeded by means of a discrete n-body model, and though he achieved considerable mathematical success a small but vital error in the model itself brought failure to his efforts to obtain a solution for the corresponding continuous bodies by taking n to infinity. Then he had a slight CAUCHY-like inspiration from a small paper of 1804 by BIOT on the propagation of heat in a bar[43] to start again by forming the partial differential equation directly, and in the brief periods of leisure allowed him in the next three years from his duties as Prefect of Isère at Grenoble and from his Egyptological researches he created a genuine revolution of his own: a revolution in mathematical physics, which he took beyond the realm of NEWTONIAN mechanics into a new physical territory of heat diffusion, with its own equations and physical constants and a fresh range of solution methods based on the use of linear equations, the method of separation of variables (then mainly used in solving

---

[42] An exception is H. BURKHARDT[33]: for scattered remarks, see ch. 8 *passim.* See also my *Foundations*, esp. chs. 2—5; and *Fourier*, esp. chs. 21 and 22.

[43] J. B. BIOT, "Mémoire sur la propagation de la chaleur," *Bibl. Brit.*, **27** (1804), 310—329 = *Journ. Mines*, **17** (1804), 203—224. FOURIER never acknowledged BIOT's paper!

ordinary differential equations) and the superposition of special solutions. FOURIER series were only one consequence of these new methods: another was his creation of the basic theory of the misnamed "Bessel functions", and indeed it was there that he showed his mathematical technique at its greatest. By 1807 he had progressed far; but he was unable to solve the problem of heat diffusion in an *infinite* continuous body, and so he wrote up his theoretical achievements and experimental results in a large monograph submitted to the *Institut de France* in December[44]. LAGRANGE and LAPLACE were the most important of the examiners: for various conceptual reasons LAGRANGE was opposed to the whole approach based on separation of variables, but LAPLACE was very impressed and began to take great interest in FOURIER's work. So a struggle began over the reception of FOURIER's paper, with LAPLACE, FOURIER and MONGE (another examiner, and personally close to FOURIER) in support, and opposition from LAGRANGE and — POISSON.

We must consider POISSON for a moment, for in him more than in any other single person lies the key to the Parisian mathematical rivalries. He graduated brilliantly from the *Ecole Polytechnique* in 1803, and to the aging grand masters of Parisian mathematics — LAGRANGE, LAPLACE, LEGENDRE and MONGE — he must have seemed to be the only heir to their crown: FOURIER was so occupied with administrative work at Grenoble that he could not be expected to be achieving substantial mathematical work, while CAUCHY was still only in his early teens. So POISSON was placed in a position of special favour from the beginning of his career which he exploited to the full, especially by means of influential positions on Parisian scientific journals; but over the next twenty years he gradually but steadily lost favour and reputation to FOURIER and then CAUCHY as they emerged and surpassed him in the quality of their work. The 1807 paper of FOURIER was crucial in this development. By 1805 or 1806 POISSON was already aware of some of FOURIER's results and the type of solution that he was trying to develop: he replied not only by applying to FOURIER's diffusion equation in 1806 the ideas of LAGRANGE and LAPLACE on solutions of partial differential equations using power series of functions,[45] but also by publishing a denigrating five-page review of FOURIER's monograph in 1808 in a journal of which he was mathematical editor.[46] However, LAPLACE, acting in his typical political way, maintained his interest in POISSON (and also in BIOT) while gradually changing his interests towards FOURIER's methods and results. In 1809 he published a miscellany on analysis which — without reference to FOURIER — just happened to contain a treatment of the diffusion equation with initial condi-

---

[44] For the references of this manuscript, see [12]; and for a detailed analysis of its contents, see my "Joseph Fourier and the revolution in mathematical physics", *Journ. Inst. Maths. Applics.*, **5** (1969), 230—253. Much new information on FOURIER's life and Prefectural responsibilities is contained in my *Fourier*, ch. 1.

[45] S. D. POISSON, "Mémoire sur les solutions particulières des équations différentielles et des équations aux différences", *Journ. Ec. Polyt.*, cah. 13, **6** (1806), 60—116 (pp. 109—111).

[46] S. D. POISSON, "Mémoire sur la propagation de la chaleur dans les corps solides", *Nouv. Bull. Soc. Philom. Paris*, **1** (1808), 112—116 = FOURIER's *Oeuvres*, **2**, 213—221.

tions over an infinite interval. His solution

$$v = \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-u^2} f\left(x + 2u\sqrt{Kt}\right) du \tag{17}$$

brought into mathematics a result which later was developed as the "Laplace transform"; it may well have been CAUCHY's inspiration to try integral solutions to partial differential equations.[47] It was certainly FOURIER's inspiration, for it showed FOURIER that an *integral*, rather than a series, solution was applicable in the case of an infinite interval and it led him to "Fourier integrals" and thus to his integral theorem (16). Meanwhile, POISSON had been opposing FOURIER's solution method in favour of functional solutions by means of indirect references in the context of the vibration of elastic surfaces;[48] but FOURIER and his supporters eventually managed to secure a prize problem for heat diffusion in the *Institut de France* for January, 1812. To the revision of the manuscript of 1807 FOURIER added a new section on FOURIER integrals, and also two more new parts on physical aspects of heat which were inspired by discussions with LAPLACE. He won the prize, but the criticisms of LAGRANGE in the examiners' report hurt him for the rest of his life:

"... This work contains the true differential equations of the transmission of heat, both in the interior of the bodies and at their surface, and the novelty of the purpose adjoined to its importance has determined the class [of the *Institut*] to crown this work, observing, however, that the manner of arriving at its equations is not free from difficulties and its analysis of integration still leaves something to be desired, both relative to its generality and on the side of rigour."[49]

LAGRANGE died in 1813; but publication of this second paper was no more likely than its predecessor and so FOURIER wrote his book on heat diffusion as the third version of his work. It did not appear until 1822,[50] having been delayed partly by FOURIER's own difficulties in developing the physical aspects of heat (which he eventually omitted and promised for a sequel which was never written); and the 1812 prize paper did not appear until still later.[51] By this time FOURIER

[47] P. S. LAPLACE, "Mémoire sur divers points d'analyse", *Journ. Ec. Polyt.*, cah. 15, **8** (1809), 229—265 (pp. 235—244) = *Oeuvres*, **14**, 178—214 (pp. 184—193).

[48] See especially the preamble to a prize problem on this topic in *Hist. cl. sci. math. phys. Inst. Fr.* (1808: publ. 1809), 235—241. Obviously written by POISSON, it extols the virtues of functional solutions to the wave equations — in implied contrast to FOURIER series solutions which were then available. In controversial circumstances (described in my *Fourier*, ch. 21), POISSON read his own paper on the subject in 1814, which was published as "Mémoire sur les surfaces elastiques", *Mém. cl. sci. math. phys. Inst. Fr.*, (1812), pt. 2 (publ. 1816), 167—225.

[49] Published in FOURIER's *Oeuvres*, **1**, vii—viii. The manuscript is kept in the Archives of the *Académie des Sciences*, Paris.

FOURIER never allied himself closely to LAPLACE, and gave no acknowledgement to LAPLACE in the prize paper. It may be that LAGRANGE's continued general opposition was supplemented by LAPLACIAN annoyance: the remarkable story of the relations between LAPLACE and FOURIER from 1807 until the 1820's is described in my *Fourier*, chs. 21 and 22.

[50] J. B. J. FOURIER, *Théorie analytique de la chaleur* (1822, Paris) = *Oeuvres*, **1**.

[51] J. B. J. FOURIER, "Théorie du mouvement de la chaleur dans les corps solides", *Mém. Acad. Roy. Sci.*, **4** (1819—20: publ. 1824), 185—555; and **5** (1821—22: publ. 1826), 153—246 = *Oeuvres*, **2**, 3—94.

had risen to a strong political position, having been appointed *sécretaire perpétuel* of the *Académie des Sciences* in 1821; and then there developed the second of our major controversies, which directly involved CAUCHY's *Cours d'Analyse* — the convergence problem of FOURIER series.

FOURIER series contain many of the problems which we tackle by means of the new analysis, but we have not yet described any of FOURIER's work in that field. The reason is that, although he understood all the basic analytical problems — convergence, the possibility of discontinuous functions, the integral as an area — before both BOLZANO and CAUCHY had begun their work, he was not strongly attracted to pure analysis as a study and so did not develop his own understanding to the extent of that which he was capable.[52] Doubtless CAUCHY was aware of this fact, for in the *Cours d'Analyse* he put the following theorem:

"When the different terms of the series $\left[\sum_{r=1}^{\infty} u_r\right]$ are functions of the same variable x, continuous with respect to that variable in the vicinity of a particular value for which the series is convergent, the sum of the series is also a continuous function of x in the vicinity of that particular value."[53]

The theorem is remarkable for its falsehood: it was known in its day to be false, and indeed CAUCHY knew it was refuted when he put it in his book. But to find the reasons why it was included, we must examine the type of counter-examples which were then known. They were in fact FOURIER series:

$$f(x) = \tfrac{1}{2} a_0 + \sum_{r=1}^{\infty} (a_r \cos r x + b_r \sin r x), \tag{18}$$

where

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(u)\, du, \tag{19}$$

$$a_r = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(u) \cos r u\, du, \quad r = 1, 2, \ldots \tag{20}$$

$$b_r = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(u) \sin r u\, du, \quad r = 1, 2, \ldots. \tag{21}$$

The trigonometric functions are continuous, and so the series on the right hand side of (18) is covered by CAUCHY's theorem: thus if $f(x)$ is discontinuous, the series cannot be convergent to it. But FOURIER had produced several series of discontinuous functions, and had shown by direct consideration of their n^{th} partial sums that they were convergent; and since 1815 POISSON had found that he had had to abandon his belief in functional and power series solutions in favour of FOURIER series solutions, and he had found similar examples also. So what was CAUCHY's purpose in stating his theorem? There was of course an intellectual aspect to it, for CAUCHY did have a proof: suffice it to say for now that the

[52] In the 1807 manuscript[12], see arts. 42—43, 64—74: in the 1811 paper[51], see part 1, 269—273 and 304—316: in the book[50] (mostly written by 1815), see arts. 177—179 and 222—229.

[53] A.-L. CAUCHY, *Cours*, 131—132 = *Oeuvres*, (2) 3, 120.

distinctions between modes of uniform and non-uniform convergence which resolve the difficulty were not noticed by anybody until the 1840's, that CAUCHY's theorem had some role to play in their development, and that shortly afterwards, in his last years, he wrote a pathetic paper of his own on the subject presenting the same type of idea without any reference to recent work.[54] But on the personal side, there was a message to FOURIER and POISSON between the lines of his theorem: "your trigonometric series may be very interesting, but do you have a general convergence proof for them? Do your series not affront the results of the new analysis?"

The later developments of this rivalry read almost like a novel.[55] Briefly, POISSON had already published a general proof in 1820 based on rather crude manipulations of the "Poisson integral"

$$\int_{-\pi}^{+\pi} \frac{(1-p^2)\,f(\alpha)}{1-2p\cos(x-\alpha)+p^2}\,d\alpha;^{56} \tag{22}$$

but, while he never abandoned it, it impressed few of his contemporaries. If CAUCHY knew it when he wrote the *Cours d'Analyse*, then his theorem was already a comment on it; but in a short paper of 1826 on the convergence problem he certainly showed his awareness of it. For he began that paper with a version of POISSON's convergence proof based on (22) to produce the FOURIER series (18); and then he remarked:

"The preceding series [(18)] can be very usefully employed in many circumstances. But it is important to show its convergence."[57]

CAUCHY's own proof followed; and while it was of considerably better mathematical calibre than POISSON's, it contained one vital flaw — the false assumption that if $u_n \to v_n$ as n tends to infinity, then $\sum_{r=1}^{\infty} u_r$ and $\sum_{r=1}^{\infty} v_r$ converge together. That this assumption *is* false was pointed out in a paper of 1829 on the convergence problem by the young DIRICHLET. In this masterpiece DIRICHLET showed the power of the new analysis in producing the famous sufficient "Dirichlet conditions" for the convergence of a FOURIER series to its function: that it may have

---

[54] A.-L. CAUCHY, "Note sur les séries convergentes ...", *C. R. Acad. Roy. Sci.*, **36** (1853), 454—459 = *Oeuvres*, (1) **12**, 30—36. For a detailed account of the introduction of modes of convergence, see my *Foundations*, ch. 6. The relevance of CAUCHY's theorem in the *Cours* is especially connected with one paper important in the development of modes of convergence: P. L. SEIDEL's "Note über eine Eigenschaft der Reihen, welche discontinuirlichen Functionen darstellen", *Abh. Akad. Wiss. Münich*, **7** (1847—49), math.-phys. Kl., 381—393. This paper (by a pupil of DIRICHLET!) dealt explicitly with that theorem in the light of discontinuous FOURIER series, and is more than likely to have been the (unmentioned) inspiration of CAUCHY's paper of five years later.

[55] A detailed description is given in my *Foundations*, ch. 5.

[56] S.-D. POISSON, "Mémoire sur la manière d'exprimer les fonctions ..." *Journ. Ec. Polyt.*, cah. 16, **11** (1820), 417—489 (pp. 422—424).

[57] A.-L. CAUCHY, "Mémoire sur les développements des fonctions en séries périodiques", *Mém. Acad. Roy. Sci.*, **6** (1823: publ. 1827), 603—612 (p. 606) = *Oeuvres*, (1) **2**, 12—19 (p. 14).

a finite number of discontinuities and turning values in an otherwise continuous and monotonic course.[58] And his proof was a development of a sketched argument in FOURIER's book of 1822: he took an alternative form to FOURIER's for the n^{th} partial sum of the series and applied to it a precise version of the proof that FOURIER had outlined.[59] Yet there was more than mathematics in DIRICHLET's paper, too, for during his visit to Paris in 1826 he formed such a close personal attachment to FOURIER that his work on the convergence problem was a personal homage in FOURIER's last years. However, he formed no close relationship to CAUCHY: as well as pointing out the error in CAUCHY's 1826 proof and finding general convergence conditions which, in allowing discontinuities in the function refuted CAUCHY's 1821 theorem, he reported in his paper a presumably verbal remark of CAUCHY's on his 1826 paper that:

"The author of this work himself acknowledges that his proof is defective for certain functions for which, however, convergence is incontestable."[60]

One can find CAUCHY's reaction to DIRICHLET's results if one looks carefully: in 1833 CAUCHY published in French at Turin a summarised version of all his 1820's text-books (based on the lectures that he had been giving there in Italian), and was careful to include his theorem from the *Cours d'Analyse* word for word.[61]

And so we return to BOLZANO and his Prague pamphlet. Is it any wonder that in an atmosphere like this CAUCHY made no acknowledgement to him? References were often not made (apart from honorific citations of the great names of the past), either between members of the Paris cliques or outside them; and even then they were some times double-meant. For example, when CAUCHY finally managed to get his 1814 paper on definite integrals and the 1815 paper on water-waves published in 1827 he introduced in 1825 some extra notes and footnotes to the texts and introduced fawning references to the powerful *secrétaire perpétuel* (FOURIER), especially with regard to his invention of the notation $\int_a^b$ to represent the definite integral; he also inserted attacks on the declining POISSON.[62] But there seems to me to be more specific reasons for CAUCHY's failure to acknowledge BOLZANO. He had appreciated the qualities of BOLZANO's work, and I think that he deliberately excluded references to an obviously obscure work in order to *prevent* its acquaintance by rivals such as POISSON and FOURIER (and perhaps others such as AMPÈRE). This is perhaps not a nice remark to make about CAUCHY but it is all too justified, and indeed CAUCHY's personality is worth our separate attention.

---

[58] P. G. LEJEUNE-DIRICHLET, "Sur la convergence des séries trigonometriques ...", *Journ. rei. ang. Math.*, 4 (1829), 157—169 = *Werke*, **1**, 117—132. DIRICHLET's contributions to the new analysis in this and other works (described in my *Foundations*, ch. 5), surpass in my view any other of CAUCHY's successors — including ABEL.

[59] See J. B. J. FOURIER[50], esp. art. 423.

[60] See P. G. LEJEUNE-DIRICHLET[58], 157 = *Werke*, **1**, 119.

[61] A.-L. CAUCHY, *Résumés Analytiques* (1833, Turin), 46 = *Oeuvres*, (2) **10**, 55—56.

[62] For CAUCHY's acknowledgements to FOURIER, see [37], 623 = *Oeuvres*, (1) **1**, 340; and [38], 194 (omitted from *Oeuvres*, (1) **1**, 197). For the attacks in [38] on POISSON, see pp. 187—188 = *Oeuvres*, (1) **1**, 189—191.

7. Cauchy's Personality

If CAUCHY was one of the greatest mathematicians of his time, he was one of the most unpleasant personalities of all time: a fanatic for Catholic and Bourbonist causes to the point of perversion, he had to prove his superiority at all times over even the weakest of his contemporaries and to publish a virtually continuous stream of work. He also wrote articles on education, the rights of the Catholic and Bourbon causes, and the reform of criminals, to supplement his mathematical output; but he never helped and even at times hindered his younger colleagues in their careers and work. A good example of this concerns a young man who wrote the following of him:

"Cauchy is a fool, and one can't find any understanding with him, although he is the mathematician who at this time knows how mathematics should be treated ... he is extremely catholic and bigoted ... ."

The writer was ABEL, in a letter sent to his friend HOLMBOE when, like DIRICHLET, he visited Paris in October, 1826.[63] Poor ABEL: he cannot have known how right he was, just as he did not understand the Parisian political situation. While in Berlin during the previous January, he had written a paper on convergence tests and their application to the binomial series which made important use of the new analysis: he had also spotted the weakness in CAUCHY's theorem of the *Cours d'Analyse* and made the first public mention of the point in a footnote to the paper.[64] Later in the same letter to HOLMBOE he remarked:

"I have worked out a large paper on a certain class of transcendental functions to present to the *Institut*. I am doing it on Monday. I showed it to Cauchy: but he would hardly glance at it. And I can say without bragging that it is good. I am very curious to hear the judgement of the *Institut* ... ."[65]

This was the paper which ushered in the transformation of LEGENDRE'S theory of elliptic integrals into his own theory of elliptic functions; and the story of its fate is only too characteristic of Parisian science and of CAUCHY. CAUCHY and LEGENDRE were the examiners: CAUCHY took it and, perhaps because of ABEL's footnote against his theorem, ignored it entirely: only after ABEL's death in 1829 did he fulfil a request to return it to the *Académie des Sciences*. It was finally published in 1841, when the manuscript vanished in sensational circumstances, to be rediscovered only in the 1950's. This story is well-known;[66] however, there is one aspect of it which has been little remarked upon but which shows the depths to which CAUCHY could sink. When ABEL's paper was in the press another Norwegian mathematician presented a paper to the *Académie des Sciences*

[63] *Niels Hendrik Abel. Mémorial publié à l'occasion du centenaire de sa naissance* (1902, Christiana), *Correspondance d'Abel* ..., 135 pp. (pp. 45 and 46) = *Texte original des lettres* ..., 61 pp. (pp. 41 and 42). Also in *Oeuvres* (ed. L. SYLOW & S. LIE), **2**, 259.

[64] N. H. ABEL, "Untersuchungen über die Reihe ...", *Journ. rei. ang. Math.*, **1** (1826), 311—329 (p. 316) = *Oeuvres* (ed. B. HOLMBOE), **1**, 66—92 (p. 71) = *Oeuvres* (ed. L. SYLOW & S. LIE), **1**, 219—250 (p. 225).

[65] In addition to the references in [63], we may add for this passage ABEL's *Oeuvres* (ed. B. HOLMBOE), **2**, 269—270.

[66] For a detailed account of this affair, see O. ORE, *Niels Hendrik Abel — mathematician extraordinary* (1957, Minneapolis), 246—261.

on elliptic functions. Cauchy was again an examiner, and his report contains the following words:

"Geometers know the beautiful works of Abel and of Mr. Jacobi on the theory of elliptic transcendentals. One knows that of the important papers ... one of them in particular was approved by the *Académie* in 1829, on the report of a commission of which Mr. Legendre was a part [Cauchy himself having been the other!], then crowned by the *Institut* in 1830, and that the value of the prize was remitted to Abel's mother. In fact this illustrious Norwegian, whom a project of marriage had determined to undertake a voyage in the depth of winter, unfortunately fell ill towards the middle of January 1829 and, in spite of the care that had been lavished on him by his fiancée's family, he died of phthisis on the 6th April, having been confined to bed for three months ....

"Before completing this report where we have often had to recall the works of Abel, it appears to us proper to dispel an error which is already quite widespread. It has been supposed that Abel died in misery, and this supposition has been the occasion for violent attacks directed against scholars from Sweden and from other parts of Europe. We would want to believe that the authors of these attacks will regret that they expressed themselves with such vehemence, when they read the Preface of the ... *Oeuvres d'Abel*, recently published in Norway by Mr. Holmboe, the teacher and friend of the illustrious geometer. They will see there with interest the flattering encouragements, the expressions of esteem and admiration that Abel received from scholars during his life, particularly from those who occupied themselves at the same time as he with the theory of elliptic transcendentals ... ."[67]

In fact Cauchy must have known that, while preparing his 1839 edition of Abel's works, Holmboe had tried without success to obtain the 1826 manuscript from the *Académie des Sciences* and that its publication in 1841 was due only to the fact that he had raised the matter to governmental level. Anyone capable of writing in this manner, knowing the negative role played by himself in the matter under discussion, would hardly think twice about borrowing from an unknown paper published in Prague without acknowledgement.

But how unknown *was* Bolzano's paper?

---

[67] A.-L. Cauchy, "Rapport sur un mémoire de M. Broch, relatif à une certaine classe d'intégrales," *C. R. Acad. Roy. Sci.*, **12** (1841), 847—850 = *Oeuvres*, (1) **6**, 146—149. Abel's paper was then appearing as "Mémoire sur une propriété générale d'une classe très-étendue de fonctions transcendantes", *Mém. prés. Acad. Roy. Sci. div. sav.*, (2) **7** (1841), 176—264 = *Oeuvres* (ed. L. Sylow & S. Lie), **1**, 145—211. Broch's paper appeared as "Mémoire sur les fonctions de la forme

$$\int x^{s-y\varrho-1} f(x^\varrho)\, R(x^\varrho)^{+\frac{s}{x\varrho}}\, \partial x\text{''},$$

Journ. rei. ang. Math., **23** (1846), 145—195 and 201—242: we note the five-year delay, and the fact that its publication was *not* in the journal of the *Académie* to which it had been assigned. Cauchy's report (with Liouville as co-signatory but certainly not author!) prefaced the paper on pp. 145—147: he was referring in the above quotation to the "Notice sur la vie de l'auteur" that Holmboe put in his edition of Abel's *Oeuvres*, **1**, v—xiv. At the end of that edition Holmboe included a selection of his letters from Abel, and we note from [63] and [65] that he did *not* include Abel's remark on Cauchy.

8. The Availability and Familiarity of Bolzano's Work

We have mentioned several times that BOLZANO's achievements anticipated specifically the work of the WEIERSTRASS school in the 1860's, and it was they who first brought BOLZANO's mathematical publications[68] to general attention at that time. DU BOIS REYMOND, CANTOR, HANKEL, HARNACK, HEINE, SCHWARZ, STOLZ — they formed perhaps the most talented group ever to work on foundational problems in analysis, and they all had a deep interest in the history of their subject. I do not know which of them first came across BOLZANO's writings: the first to make a reference in print was HANKEL in 1871,[69] but SCHWARZ was the one most interested in these questions and it was he who around that time named WEIERSTRASS's theorem on the existence of a limiting value of an infinite closed sequence of values the "BOLZANO-WEIERSTRASS theorem", in view of BOLZANO's theorem on the existence of an upper limit in his 1817 pamphlet which we quoted in section 2.4.[70] WEIERSTRASS's group were then studying continuous non-differentiable functions, rational and irrational numbers, and the early ideas of set theory, on all of which BOLZANO had preceded them; and so it had tended to be assumed (posterior wisdom again) that in his own day BOLZANO was not read at all. Without any doubt his works were not widely available — for proof of this, we need only mention that it is today extremely difficult to find copies of any of them. But it would be a mistake to assume that because they appeared as pamphlets they could not have become widely familiar. On the contrary, at that time the publication of pamphlets was a common method of issuing scientific literature and indeed avoided the notorious delays of academy journals: CAUCHY for example, always anxious for rapid publication, put some very important work into pamphlets and lithographs, and even published his own mathematical journal during two periods of his life.[71] There seems to have been a well organised trade in the sale of such material, based on the catalogues of book shops designed especially for scientific and intellectual circles: it was by these means, for example, that BOLZANO in Bohemia managed to learn of and obtain the current literature. So we may presume that the work was in reasonably fluid circulation — and surely especially in Paris, the scientific centre of the age. CAUCHY himself reveals this in his own writings. Although his refer-

[68] Apart from the *Beweis* and the works listed in [29] and [34], BOLZANO also published *Betrachtungen über einige Gegenstände der Elementargeometrie* (1804, Prague) = *Schriften*, **5** (1948, Prague), 9—49; and *Beiträge zu einer begründeten Darstellung der Mathematik. 1. Lieferung* (1810, Prague) = (1926, Paderborn). (No other parts published.) BOLZANO's friend F. PŘIHONSKÝ posthumously published his *Paradoxien des Unendlichen* (1851, Leipzig): there have been various re-issues and translations of this work, including an English edition (1950, London).

[69] H. HANKEL, "Grenze", *Allg. Enc. Wiss. Künste*, sect. 1, pt. 90 (1871, Leipzig), 185—211: see pp. 189, 209—210. The first major study was by STOLZ, as "B. Bolzanos Bedeutung in der Geschichte der Infinitesimalrechnung", *Math. Ann.*, **18** (1881), 255—279 (and corrections in **22** (1883), 518—519).

[70] See K. SCHWARZ, "Zur Integration der partiel Differentialgleichung $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$", *Journ. rei. ang. Math.*, **74** (1872), 218—253 (p. 221) = *Abhandlungen*, **2**, 175—210 (p. 178).

[71] See his *Exercises des Mathématiques* (4 vols. and 1 instalment: 1826—30, Paris), and *Exercises d'Analyse et de Physique Mathématique* (4 vols: 1840—47, Paris). They appear respectively in his *Oeuvres*, (2) **6—9**; and (2) **11—14**.

27*

ences were often not always given, they show that he was abreast of current writings in all European languages, and not only the most prominent authors, books and journals: there are also references to little known material. In the *Cours d'Analyse*, for example, he referred to a pompous little tract of 1820 published in London on rules of signs in the theory of equations,[72] which was at least as obscure as BOLZANO's pamphlet. In fact, BOLZANO had given his paper two opportunities for publication, for not only did he issue it as a pamphlet in 1817, but — with the same printing — inserted it into the 1818 volume of the Prague Academy *Abhandlungen*.[73] That journal was available in Paris: indeed, the *Bibliothèque Impériale* (now the *Bibliothèque Nationale*) began to take it with *precisely the volume containing Bolzano's pamphlet*.[74] So here is at least one plausible possibility for CAUCHY to have found a copy of BOLZANO's paper, quite apart from the book-trade: he could have noticed a new journal in the library's stock and examined it as a possible course of interesting research.

We turn now from the availability to the familiarity of BOLZANO's works. We have seen that they were not widely circulated, although probably more so than might be imagined; but apart from that I feel that an important factor in the apparent indifference of his contemporaries was a *lack of understanding* of what he had achieved. Since his important results were so far ahead of its time, only a genius of CAUCHY's type and magnitude could bring them to the realisation they deserved (and of which their creator was probably incapable). We can appreciate this point better if we return to ABEL. There is no reference to any of BOLZANO's works in ABEL's writings, and seemingly no direct influence either, even though they had both written on the binomial series; but ABEL had certainly read some BOLZANO, for he expressed great admiration for him in a notebook and hoped to meet him in Prague during his European tour.[75] I suspect that several mathematicians were in ABEL's position: impressed by BOLZANO's work, but unable to take it further themselves.[76] But without doubt there were, unfortunately, many who never discovered it at all. This, therefore, is a situation in marked contrast to CAUCHY's works, which were read by everybody — including BOLZANO.

[72] P. NICHOLSON, *Essay on involution and evolution; containing a new accurate and general method of ascertaining the numerical value of any function* ... (1820, London). CAUCHY's reference is in the *Cours*, 500 = *Oeuvres*, (2) **3**, 409: he also wrote a number of papers on this subject in the 1810's, but with an interest towards structural properties (permutations, *etc.*) rather than in the foundations of analysis. For commentary, see H. WUSSING, *Die Genesis des abstrakten Gruppenbegriffes* (1969, Berlin) esp. pp. 61—66.

[73] See the references in [3].

[74] The present call mark of this volume is R. 15200 in the *Départment des Imprimés*. There is no record of its readers, neither does it contain any annotated markings or corrections. The only other copy of the work known to me in Paris is in the holding of the journal by the *Muséum Nationale d'Histoire Naturelle* — a source hardly likely to have been used by CAUCHY. The copy has no revealing annotations on it.

[75] See L. SYLOW, "Les études d'Abel et ses découvertes," [62], 59 pp. (pp. 6 and 13); and K. RYCHLIK, "Niels Hendrik Abel a Čechy", *Pok. mat. fys. astron.*, **9** (1964), 317—319.

[76] LOBACHEWSKY also knew BOLZANO's 1817 pamphlet on the roots of a continuous function: see B. L. LAPTIEV, "О библиотечных записах книг и журналов, выданных Н. И. Лобачевскому", Усп. Мат. Наук, **14** (1959), pt. 5, 153—155.

9. The Personal Relations between Bolzano and Cauchy

That CAUCHY read BOLZANO's 1817 pamphlet is the subject of our conjecture; but that BOLZANO read CAUCHY's *Cours d'Analyse* is beyond question, for in an important manuscript of the 1830's on analysis he referred to CAUCHY's as one of the recent formulations of continuity in his own style.[77] By then of course, BOLZANO's ideas had gained much publicity through CAUCHY's book, which itself had been published at Königsberg in 1828 in a German translation which may well have been the version that BOLZANO read. Yet there was never a priority row between the two over their common ideas. This is, however, not surprising. In the first place, BOLZANO was no CAUCHY, incessantly anxious for publication and his "rights"; and in addition he was already a controversial figure in Bohemia on account of his progressive views on society and religion. Thus, even if he had wanted to stage a priority row from Bohemia against the great CAUCHY in Paris, he would have found it especially difficult. But I would suggest that there is still another reason why BOLZANO did not promote such a row; namely, that he probably *never noticed* the correspondence of ideas — or at least their significance — when he read the *Cours d'Analyse*. For the *Cours* is a large book, nearly 600 pages in length; and almost all of it is CAUCHY, applying BOLZANO's germinal ideas to one analytical problem after another. But the ideas themselves and the direct points of correspondence appear only here and there in its course, and could easily be missed in the general context.

This view is strengthened when we consider their personal relations. There was no meeting between the two in the 1810's or 1820's, for CAUCHY was in France and BOLZANO in Bohemia; but after the fall of the Bourbons in 1830 CAUCHY exiled himself, firstly to Italy, and then, between 1833 and 1835, to Prague to assist in the education of the son of the dethroned King CHARLES X. The tone of BOLZANO's reaction to CAUCHY's visit to Prague, in a letter he sent to his friend PŘIHONSKÝ in August, 1833, indicates quite clearly that he had had no contact with CAUCHY of any sort and that he suspected no *direct* use of his results by CAUCHY:

"The news of Cauchy's presence [in Prague] is uncommonly interesting for me. Among all living mathematicians today he is the one whom I esteem the most and to whom I feel the most akin; I owe to his inventive spirit some of the most important proofs. I ask you very much to recommend me to him and to say that I would have travelled now straight to Prague to make his personal acquaintance, if I — after what you tell me of his appointment — could not hope for certain that I will meet him at the end of September, ..."[78]

There were in fact a few meetings, for BOLZANO described them in a letter of December, 1843 to FESL:

"Cauchy, the mathematician, was ... in the years 1834 or 1835 ... in Prague, where we met a few times during the few days that I was accustomed to spend at that time (at Easter and in the autumn) in Prague. After my departure I let

[77] B. BOLZANO[7], in *Schriften*, **1** (1930, Prague), 15: see also p. 94.

[78] See E. WINTER (ed.), "Der böhmische Vormärz in Briefen B. Bolzanos an F. Přihonský (1824—1848)", *Veröff. Inst. Slav., Dtsch. Akad. Wiss. Berlin.*, **11** (1958), 306 pp. (p. 156).

Kulik deliver to him (1834) an essay filling a single quarto sheet which I had drafted for Cauchy sometime in French, on the famous mathematical problem of the rectification of curves, because I rightly feared that he would find the "Paper on the three problems of rectification, planing and cubing" published in 1817[79] too comprehensive and difficult. Early last year, as I was looking through some issues of Cauchy's writings[80] bound with the usual coloured wrappers, and [turned to] the lists of works announced on the back, I noticed with astonishment a small note by him on the same subject, that he had published as a lithograph in Paris in 1834 (therefore presumably only after he had read my little essay). Naturally I would be very eager to read the note ... ."[81]

Eventually BOLZANO managed to obtain a copy of the paper: in fact it came through FESL who pointed out to him that it had been written in 1832 rather than 1834 and so could not be related to his essay, and that it treated the subject in a quite different way. BOLZANO admitted this in an acknowledgement to FESL in May, 1844,[82] and it is quite clear that in this case there was not even a correspondence of ideas; but on the foundations of analysis a very different situation seems to have applied. One would dearly like to know the content of their conversations; but if BOLZANO ever wondered even for a moment that CAUCHY had read his 1817 paper before writing the *Cours d'Analyse*, I imagine that he would have been pleased rather than annoyed. For when he wrote that paper, he had known then that it was a significant work which would probably not reach the audience that it deserved; and so he had ended its preface with a plea to the scientific community which I believe CAUCHY accepted:

"... I must request ... that one does not overlook this particular paper because of its limited size, but rather examine it with all possible strictness and make known publicly the results of this examination, in order to explain more clearly what is perhaps unclear, to revoke what is quite incorrect, but to let succeed to general acceptance, the sooner the better, what is true and right."[83]

10. Epilogue

My conjecture has aroused considerable adverse criticism before publication, and will doubtless receive much more now: thus to minimise the possibility of misunderstandings of this paper, a few points may be worth stressing.

1. Part of my purpose has been to describe some of the extra-intellectual aspects of Parisian mathematics; and whether or not my conjecture is correct

[79] The reference for this work is given in [29].

[80] Presumably the *Exercises d'Analyse* listed in [71].

[81] See I. SEIDERLOVÁ, "Bemerkung zu den Umgängen zwischen B. Bolzano und A. Cauchy," *Čas. pěst. mat.*, **87** (1962), 225—226.

[82] See [81]. CAUCHY's paper, read to the *Académie des Sciences* on the 22nd October, 1832, was the "Mémoire sur la rectification des courbes et la quadrature des surfaces courbes", *Mém. Acad. Roy. Sci.*, **22** (1850), 3—15 = *Oeuvres*, (1) **2**, 167—177; but in the publisher's lists in the *Exercises d'Analyse* it is described as an 11-page lithograph of 1832, which was its first publication. I do not understand why BOLZANO thought that it had been published in 1834.

[83] B. BOLZANO, *Beweis*, end of preface.

I am firmly convinced that rivalries of the type of which I have given some examples played an important role in Parisian mathematics, and so I have tried to bring to the attention of historians of this period the kinds of historical problem that they will have to face in interpreting its literature. In addition, the theory of "limit-avoidance" is an historical tool which appears to be some use in one form or another in investigating the development of analysis and the calculus in this and other periods.

2. I cannot stress too strongly that in characterising Cauchy's genius as responsive to exterior stimuli I am trying to *describe* rather than *decry* the depth and extent of his originality. Without any question he and GAUSS were the major mathematicians of the first decades of the nineteenth century: thus his work has to be given especial attention by historians. It is of course not my position that CAUCHY would *never* give references without intending a double meaning, but I do think that in his writings, *and equally in those of his "colleagues"*, questions of this type do need to be borne very carefully in mind. With regard to BOLZANO's pamphlet, it is possible that CAUCHY, the busy and active research mathematician and professor at three Paris colleges, simply did not bother to mention it or even forgot that he had read it (though personally I would not regard this explanation as sufficient). My case would be much strengthened by documentary evidence of some kind: CAUCHY did leave a *Nachlass* containing mathematical manuscripts and correspondence, for it was used by VALSON when preparing his excessively admiring biography of CAUCHY,[84] but unfortunately it was kept in the family and there is reason to think that, like his library, it has now been lost.

3. I remarked that CAUCHY was familiar with European languages: in the case of German, it is perhaps worth mentioning explicitly (from a number of examples) that he examined in 1817 a manuscript in German sent in to the *Académie des Sciences*,[85] and that he reviewed MÖBIUS's *Der barycentrische Calcul* in 1828.[86] We may also record another "coincidence of ideas" with obscure German writing strikingly similar to the case of BOLZANO's pamphlet. In April 1847, GRASSMANN, then a schoolmaster at Stettin, sent to CAUCHY two copies of his 1844 *Ausdehnungslehre*, but he never received any acknowledgement; however between late 1847 and 1853 CAUCHY published a number of papers on a theory of "clefs algébriques" which basically used the same sort of ideas and even some almost identical notation.[87] I offer no judgement here on the matter:

[84] C.-A. VALSON, *La vie et les travaux du Baron Cauchy* (2 vols.: 1868, Paris): see esp. vol. **2**, viii—x.

[85] See *Procès-Verbaux des séances de l'Académie tenues depuis la fondation jusqu'au mois d'août, 1835* (10 vols: 1910—22, Hendaye), **6**, 210. I may remark here that these volumes are an invaluable source of historical insight into the period 1795—1835, when the rivalries were at their height. They give the minutes of all the private meetings of the *Académie des Sciences*, which the participants can hardly have expected to be published!

[86] A.-L. CAUCHY, *Bull. Univ. Sci. Ind.* [Ferrusac], *Sci. math. phys. chim.*, **9** (1828), 77—80. Not in the *Oeuvres*.

[87] For the references and some discussion of the affair, see M. J. CROWE, *A history of vector analysis* (1967, Notre Dame and London), 82—85 and 106. CROWE's last reference in his [63] is inaccurate and in fact misleading; it should be "Mémoire sur les clefs algébriques", *Exercises d'Analyse et de Physique Mathématique*, **4** (1847, Paris), 356—400 = *Oeuvres*, (2) **14**, 417—460.

I merely record it as another example of the kind of historical problem which surrounds the great achievements of the Parisian mathematicians of the time, when Paris was the centre of the scientific world and CAUCHY's achievements among its principal adornments.

Index of Names

We list here the names and dates of persons mentioned in the main text.

D'ALEMBERT, JEAN LE ROND (1717—1783)
AMPÈRE, ADRIEN MARIE (1775—1836)
ARBOGAST, LOUIS FRANÇOIS ANTOINE (1759—1803)
BERNOULLI, DANIEL (1700—1782)
BESSEL, FRIEDRICH WILHELM (1784—1846)
BIOT, JEAN BAPTISTE (1774—1862)
DU BOIS REYMOND, PAUL DAVID GUSTAV (1831—1889)
BOLZANO, BERNARD PLACIDUS JOHANN NEPOMUK (1781—1848)
BOREL, EMILE FELIX EDOUARD JUSTIN (1871—1959)
CANTOR, GEORG FERDINAND LUDWIG PHILIPP (1845—1918)
CARNOT, LAZARE NICOLAS MARGUERITE (1753—1823)
CAUCHY, AUGUSTIN-LOUIS (1789—1857)
CHARLES X, KING (1757—1836)
DIRICHLET, PETER GUSTAV LEJEUNE- (1805—1859)
EULER, LEONHARD (1707—1783)
FESL, MICHAEL JOSEPH (1786—1864)
FOURIER, JEAN BAPTISTE JOSEPH (1768—1830)
GAUSS, KARL FRIEDRICH (1777—1855)
GRASSMANN, HERMANN GÜNTHER (1809—1877)
HANKEL, HERMANN (1839—1873)
HARNACK, CARL GUSTAV AXEL (1851—1888)
HEINE, EDUARD HEINRICH (1821—1881)
L'HUILIER, SIMON ANTOINE JEAN (1750—1840)
HOLMBOE, BERNT MICHAEL (1795—1850)
JACOBI, CARL GUSTAV JACOB (1804—1851)
KULIK, JAKOB PHILIPP (1793—1863)
LACROIX, SYLVESTRE FRANÇOIS (1765—1843)
LAGRANGE, JOSEPH LOUIS (1736—1813)
LAPLACE, PIÈRRE SIMON (1749—1827)
LEGENDRE, ADRIEN MARIE (1752—1833)
LEIBNIZ, GOTTFRIED WILHELM (1646—1716)
MÖBIUS, AUGUST FERDINAND (1790—1868)
MONGE, GASPARD (1746—1818)
NEWTON, ISAAC (1642—1727)
POISSON, SIMÉON DENIS (1781—1840)
PŘIHONSKÝ, FRANZ (1788—1859)
RIEMANN, GEORG FRIEDRICH BERNHARD (1826—1866)
SCHWARZ, KARL HERMANN AMANDUS (1843—1921)
STOLZ, OTTO (1842—1905)
TAYLOR, BROOKE (1685—1731)
VALSON, CLAUDE ALPHONSE (1826— ?)
WEIERSTRASS, KARL THEODOR WILHELM (1815—1897)

Enfield College of Technology
Middlesex
England

(Received February 17, 1970)

Did Cauchy Plagiarize Bolzano?

H. FREUDENTHAL

1. Introduction

1. In an elaborate erudite paper* I. GRATTAN-GUINNESS has put forward a case that CAUCHY plagiarized BOLZANO:

In Section 2, he discusses why *if* CAUCHY plagiarized BOLZANO, he did it so badly,

In Section 3, he presents a new limit concept which he calls "limit avoidance",

In Section 4, he mentions some facts from analysis before CAUCHY'S time,

In Section 5 he claims that CAUCHY could not have written a so "utterly untypical" work as his *Cours d'Analyse* of 1821 without having been inspired by somebody else,

In Section 6–7 he analyzes the quarrels among French mathematicians around 1800 and CAUCHY'S bad character so as to explain psychologically *why* CAUCHY plagiarized BOLZANO,

In Section 8 he discusses whether CAUCHY could have read BOLZANO,

In Section 9 he deals with the personal relations between CAUCHY and BOLZANO.

Here I wish to discuss the specific question set as the title of this paper, whether CAUCHY plagiarized BOLZANO, a question not considered directly by GRATTAN-GUINNESS.

I have to apologize that I am not well enough acquainted with the *chronique scandaleuse* of the French Academy to follow GRATTAN-GUINNESS there. On the other hand I entirely agree with him that a historian is obliged to read *between* the lines**, though I think it just as important to read the lines themselves. In history of mathematics it is also a good idea to understand the mathematics involved.

The question set as the title of the present paper can be put more precisely by asking

whether CAUCHY read BOLZANO,
whether CAUCHY could have learned new things from BOLZANO,
whether these things were so important that he should have cited BOLZANO.

---

* I. GRATTAN-GUINNESS, "Bolzano, Cauchy and the New Analysis of the Early Nineteenth Century", Archive for History of Exact Sciences **6** (1970), 372–400.

** p. 387, 17.

It is no sacrilege to ask such questions, even the last one. False ascriptions are a tradition in mathematics; twice I have met opposition when I refuted such ascriptions*.

2. The Style of Cauchy's Text-Books on Calculus**

CAUCHY is credited with having laid the first solid foundations of what is now called Analysis or Calculus. Though this is true, it is not the whole truth, and in a certain sense it is a misleading statement. It is true that mathematicians learned from CAUCHY's *Cours d'Analyse* and other text-books what continuity and convergence were and how to test for them, how to be careful with TAYLOR series and how to estimate their remainders, how to avoid pitfalls when multiplying and rearranging series, how to deal with multivalued functions, how to define differential quotients and integrals, how to be careful with improper and singular integrals, and that they found there the first example of the powerful method that later became standard in analysis and recently has come to be called "epsilontics".

To know what was new in CAUCHY's textbooks on Calculus, we had better listen to his own words, in the Introduction to his Cours d'Analyse***:

> Quant aux méthodes, j'ai cherché à leur donner toute la rigueur qu'on exige en géométrie, de manière à ne jamais recourir aux raisons tirées de la généralité de l'algèbre. Les raisons de cette espèce, quoique assez communément admises, surtout dans le passage des séries convergentes aux séries divergentes, et des quantités réelles aux expressions imaginaires, ne peuvent être considérées, ce me semble, que comme des inductions propres à faire pressentir quelquefois la vérité, mais qui s'accordent peu avec l'exactitude si vantée des sciences mathématiques. On doit même observer qu'elles tendent à faire attribuer aux formules algébriques une étendue indéfinie, tandis que, dans la réalité, la plupart de ces formules subsistent uniquement sous certaines conditions, et pour certaines valeurs des quantités qu'elles renferment. En déterminant ces conditions et ces valeurs, et en fixant d'une manière précise le sens des notations dont je me sers, je fais disparaître toute incertitude; et alors les différentes formules ne présentent plus que des relations entre les quantités réelles, relations qu'il est toujours facile de vérifier par la substitution des nombres aux quantités elles-mêmes. Il est vrai que, pour rester constamment fidèle à ces principes, je me suis vu forcé d'admettre plusieurs propositions qui paraîtront peut-être un peu dures au premier abord. Par exemple, j'énonce dans le chapitre VI, qu'*une série divergente n'a pas de somme;* dans le chapitre VII, qu'*une équation imaginaire est seulement la représentation symbolique de deux équations entre quantités réelles;* dans le chapitre IX, que, *si des constantes ou des variables comprises dans une fonction, après avoir été supposées réelles, deviennent imaginaires, la notation à l'aide de laquelle la fonc-*

* GRATTAN-GUINNESS remarks (p. 398, 5 f.b.) that his "conjecture has aroused considerable adverse criticism before publication". In his lecture on this subject before an audience of mathematicians rather than historians that I attended, it was his mathematics rather than his thesis on CAUCHY that aroused opposition.

** CAUCHY, Oeuvres (2) **3—5**.

*** CAUCHY, Oeuvres (2) **3**.

tion se trouvait exprimée, ne peut être conservée dans le calcul qu'en vertu d'une convention nouvelle propre à fixer le sens de cette notation dans la dernière hypothèse; & c. Mais ceux qui liront mon ouvrage reconnaîtront, je l'espère, que les propositions de cette nature, entraînant l'heureuse nécessité de mettre plus de précision dans les théories, et d'apporter des restrictions utiles à des assertions trop étendues, tournent au profit de l'analyse, et fournissent plusieurs sujets de recherches qui ne sont pas sans importance. Ainsi, avant d'effectuer la sommation d'aucune série, j'ai dû examiner dans quels cas les séries peuvent être sommées, ou, en d'autres termes, quelles sont les conditions de leur convergence; et j'ai, à ce sujet, établi des règles générales qui me paraissent mériter quelque attention.

The "generality of algebra" meant that what was true for real numbers, was true for complex numbers, too, what was true for convergent series, was true for divergent ones, what was true for finite magnitudes, held also for infinitesimal ones. Today it is hard to believe that mathematics ever relied on such principles, and since differentials now are only an uneasy remainder of the pre-CAUCHY period, we readily identify CAUCHY's renovation with the progress from "infinitesimal" methods to epsilontics, in spite of CAUCHY's own, much broader, appreciation, by which *all* metaphysics was barred from mathematics. The next generation of mathematicians, who had been brought up with the *Cours d'Analyse*, and the generations after WEIERSTRASS, CANTOR and DEDEKIND, who knew which course the development of analysis was due to take after CAUCHY, put the stress differently than CAUCHY and his generation would have done; at that time, and even more today, people would not properly understand what it meant if you told them that CAUCHY abolished "the generality of algebra" as a foundation stone of mathematics.

I. GRATTAN-GUINNESS has been puzzled by the "untypical" character of CAUCHY's work on Calculus as compared to his production before 1821. It is indeed puzzling. But GRATTAN-GUINNESS might have added that it is untypical even if compared with CAUCHY's work after 1821. The strange thing is that in his research papers CAUCHY never lived up to the standards he had set in his *Cours d'Analyse*. Though he had given a definition of continuity, he never proved formally the continuity of any particular function. Though he had stressed the importance of convergence, he operated on series, on FOURIER transforms, on improper and multiple integrals, as though he had never raised problems of rigor. In spite of the stress he had laid on the limit origin of the differential quotient, he developed also a formal approach to differential quotients like LAGRANGE's. He admitted semi-convergent series and rearrangements of conditionally convergent series if he could use them. He formally restricted multivalued complex functions of x as $\log x$, $\sqrt{x}$, and so on, to the upper half plane, but if he could use them in the lower half plane, he easily forgot about this prescription. CAUCHY looks self-contradictory, but he was simply an opportunist in mathematics, notwithstanding his dogmatism in religious and political affairs. He could afford this opportunism because, with the background of a vast experience, he had a sure feeling for what was true, even if it was not formulated or proved according to the standards of the *Cours d'Analyse*.

Why, then, was the *Cours d'Analyse* so different from his other work? Not because it was more fundamental, but because it was a textbook, in which he not only communicated his results but also made explicit his background experience. CAUCHY was not a lover of foundational research like BOLZANO, but to teach mathematics to beginners, he had to analyze and to present the techniques implicit in his background. A similar situation is common today, when a modern teacher of mathematics will make explicit his logical habits, even though he is not a logician.

There is at least one work of CAUCHY, his theory of determinants of 1812*, which shows the same "untypical" features; it is not to be wondered at that for a long time this was the only textbook on determinants. The most "untypical" CAUCHY of all, however, is found in his marvellous first communication on Elasticity of 1822**, which by its conceptual style towers high above the usual algorithmic swamp in which he moves.

Certainly, one has to be careful with stylistic arguments. If CAUCHY'S work had come down to us anonymously, by stylistic arguments we might attribute the *Cours d'Analyse*, the introduction to elasticity, and the remainder of his scientific work to at least three different CAUCHYS; on account of content we might even attribute his work on complex functions also to at least three CAUCHYS, so as to account for the strange phenomenon of periodic amnesia: often he asserts propositions he had recognized as wrong a short time before*** and for 26 years he seems to have forgotten the most important paper he wrote in this field****.

CAUCHY did not live *in vacuo*. He was moved by work of others, and though he made lavish acknowledgements to work of others, we can never be sure whether he cited all sources of his inspiration. By his own testimony we know that LEIBNIZ was inspired to his discoveries in Calculus by work of PASCAL which actually was only weakly related to what LEIBNIZ himself finally achieved; even according to modern standards LEIBNIZ could hardly have been obliged to cite PASCAL on these grounds. In any case from LEIBNIZ' publications we could not guess who among LEIBNIZ' predecessors was the most influential.

To tell from mere stylistic arguments that CAUCHY'S *Cours d'Analyse* must have been inspired by essentially other sources than those on complex functions or hydrodynamics, is an utterly dangerous conclusion. I have spent so much time on it because the difference of style between the *Cours d'Analyse* and other work of CAUCHY is indeed striking, and because I. GRATTAN-GUINNESS confesses that this feature was the starting point of his investigation.

---

* CAUCHY, Oeuvres (2) **1**, 91–169. (Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs...) See also Oeuvres (2) **1** 64–90. (Mémoire sur le nombre de valeurs qu'une fonction peut acquérir.)

** CAUCHY, Oeuvres (2) **2**, 300–304.

*** *E.g.* the conditions for development into a series of partial fractions in CAUCHY, Oeuvres (2) **7**, 324–362, and (1) **8**, 55–64, or multivalued functions in CAUCHY, Oeuvres (1), **8**, 156–160 and (1) **8**, 264.

**** A. L. CAUCHY, Mémoire sur les intégrales définies prises entre des limites imaginaires, Paris 1825, 4°, 68 pages. Reprinted in Bull. sci. math. **7** (1874), 265–304; **8** (1875), 43–55, 148–159; due to be reprinted in CAUCHY, Oeuvres (2) **15**.

3. Bolzano's Pamphlet of 1817

The first theorem of BOLZANO's pamphlet* is what is now called CAUCHY's convergence theorem; since a theory of real numbers is lacking, its proof can be nothing but a sham. We will come back to this point.

The next theorem is usually described as the theorem on the existence of the lowest upper bound of a bounded set of real numbers; in fact the only bounded sets considered are lower classes as used in DEDEKIND cuts, so that it would be better to term it the theorem on the *existence of the cut number*. From old times this existence has been used implicitly or explicitly. It was BOLZANO's great idea to prove it. The proof, using a sequence of dichotomies and the "Cauchy convergence criterion", is correct.

The third theorem is about continuous functions f and ϕ with $f(\alpha)<\phi(\alpha)$ and $f(\beta)>\phi(\beta)$; it states the existence of an intermediate x where $f(x)=\phi(x)$. Continuity had been defined in the preface in a perfectly modern way. The theorem is derived by considering the subset of y such that $f(x)<\phi(x)$ for all $x\leqq y$ and by applying the preceding theorem to it. Again it is a merit of BOLZANO to have recognized the idea to prove it.

The last theorem asserts the existence of a real root of a polynomial between two points where its values are of opposite sign.

As compared to CAUCHY's work, BOLZANO's pamphlet is clumsily written and partially confused. BOLZANO has no term for convergence, and none for the limit of a sequence; he always circumscribes the convergence to a certain limit by the sentence that defines this property. Of course he has no term for lowest upper bound either. His terminology is unusual; a sequence of functions is called a *veränderliche Grösse*, and a single function a *beständige Grösse*. The CAUCHY convergence criterion is formulated for a sequence, not of numbers, but of functions, and the property that is formulated, is, in fact, uniform convergence although BOLZANO draws no conclusion from it (*e.g.* with respect to continuity); the criterion is actually applied to numerical sequences only**. The proof of this criterion is worse than faulty, it is utterly confused and not at all related to the thing to be proved. At that time it was, indeed, hard to understand that such a theorem could not be proved without an underlying theory of real numbers; recently published papers of BOLZANO show that later he became aware of this fact.

This failure does not prevent the pamphlet from being a marvellous piece of work; the proofs of the other theorems are correct.

4. The Common Ideas in Bolzano and Cauchy

I am borrowing the titles of this section and of the subsections 1–5 from I. GRATTAN-GUINNESS; his remarks in the corresponding section will be analyzed here.

* B. BOLZANO, *Rein analytischer Beweis des Lehrsatzes, dass zwischen je zwey Werthen, die ein entgegengesetztes Resultat gewähren, wenigstens eine reelle Wurzel der Gleichung liege* (1817), Prague = Abh. Königl. Böhm. Gesell. Wiss. (3) **5** (1814–1817; publ. 1818), 60 p. — Also in: OSTWALD's Klassiker No. 153, ed. PH. E. B. JOURDAIN.

** This is dissimulated in I. GRATTAN-GUINNESS' quotation, where the hypothesis of the theorem is replaced with a provisional announcement taken from another section of the pamphlet.

4.1. Continuity of a Function. BOLZANO's and CAUCHY's definitions are equivalent. BOLZANO's is far better; it is modern (though instead of δ and ε he uses ω and Ω); the succession of the quantifiers is correct and clear. CAUCHY's definition uses the language of infinitesimals (an infinitely small increase of the variable produces an infinitely small increase of the functions); even the succession of the quantifiers is not clear in this formulation.

It is hard to explain how CAUCHY, if borrowing the definition of continuity from BOLZANO, could have presented it in deteriorated form; later on such occurrences are explained by I. GRATTAN-GUINNESS as instances of CAUCHY's failure to fathom the depth of BOLZANO's thought. There is, however, not the slightest reason to assume that CAUCHY learned the concept of continuous function from BOLZANO, since it was already instrumental in CAUCHY's* treatise of 1814 on complex functions (the Cauchy integral theorem):

> *Solution.* — Si la fonction $\varphi(z)$ croit ou décroit d'une manière continue entre les limites $z=b'$, $z=b''$, la valeur de l'intégrale sera représentée, à l'ordinaire, par
>
> $$\varphi(b'')-\varphi(b').$$
>
> Mais, si, pour une certaine valeur de z représentée par Z et comprise entre les limites de l'intégration, la fonction $\varphi(z)$ passe subitement d'une valeur déterminée à une valeur sensiblement différente de la première, en sorte qu'en désignant par ζ une quantité très petite, on ait
>
> $$\varphi(Z+\zeta)-\varphi(Z-\zeta)=\Delta,$$
>
> alors la valeur ordinaire de l'intégrale définie, savoir,
>
> $$\varphi(b'')-\varphi(b')$$
>
> devra être diminuée de la quantité Δ, comme on peut aisément s'en assurer.

To within a formal definition the full-fledged idea of continuity is presented not only here; it is also the main idea underlying the introduction of the CAUCHY principal value of singular integrals, which provided CAUCHY's approach to his integral theorem. There can be little doubt that here was CAUCHY's point of departure to continuity.

I. GRATTAN-GUINNESS claims that in 1821 CAUCHY did not know that continuity did not imply differentiability, while BOLZANO knew it. There is no proof for the second claim, and in the light of the role continuity plays in CAUCHY's treatise of 1814, the first claim is ridiculous.

4.2. Convergence of a Series. In the case of the Cauchy convergence criterion CAUCHY's formulation is much better than BOLZANO's. If CAUCHY ever read BOLZANO, and even if he did not understand his confused exposition, the possibility can hardly be excluded that he guessed what BOLZANO meant and consequently arrived at an improved version. Of course, this is no proof that it really happened this way. CAUCHY prepares the announcement of his criterion by a fine heuristic approach which, undoubtedly, is his own**; when reading his exposition, one can

* CAUCHY, Oeuvres (1) 1, 402–403.

** CAUCHY, Oeuvres (2) 3, 115–116.

imagine him standing at the blackboard, explaining that for a sum $\sum u_n$ to converge, it does not suffice that the u_n converge to 0, nor does it suffice that the $u_n + u_{n+1}$ converge to 0, nor does it suffice that the $u_n + u_{n+1} + u_{n+2}$ converge to 0, and so on, and that in order to get convergence of the sum you have rather to make all these expressions arbitrarily small by choosing n large.

In today's mathematics this is so natural an approach that one feels little need to ask who invented it, yet in the historical setting the CAUCHY convergence criterion looks like a premature discovery. In fact, if we expect a great many applications of the CAUCHY convergence criterion in CAUCHY's work, we are likely to be disappointed. It is applied at essentially two places:

First, to justify the majorant method of convergence proofs (if $|a_n| \leqq |c_n|$ for almost all n, and if $\sum |c_n|$ converges, then $\sum a_n$ converges), which in the particular case of a geometrical series as a majorant, is the foundation of CAUCHY's famous "*Calcul des limites*" in power series and differential equations,

Second, to prove the convergence criterion on alternating series (if the $|a_n|$ are such that $a_n a_{n+1} \leqq 0$, $|a_n| \geqq |a_{n+1}|$, and $\lim a_n = 0$, then $\sum a_n$ converges).

As soon as these two criteria have been established, the reader of the *Cours d'Analyse* may readily forget about the CAUCHY convergence criterion.

This is not to be wondered at since there was not any other essential use of the CAUCHY convergence criterion up to the rise of the direct methods of the variational calculus at the turn of the 19^{th} century. The majorant method and the criterion on alternating series as algorithmic tools were just what mathematicians in CAUCHY's time, and even later, needed. The CAUCHY convergence criterion with its much more involved logical structure, lacked this algorithmic appeal. CAUCHY's work in analysis would not have looked different if he had never formulated the CAUCHY convergence criterion and, instead, had accepted the principle of the majorant method and the criterion on alternating series as obvious truths which did not need a proof, just as, for instance, he accepted without argument that the endpoints of a nested sequence of intervals, shrinking to zero, had a limit*.

From CAUCHY's time up to the end of the 19^{th} century the CAUCHY convergence criterion was an expression of logical profundity rather than a practical tool. This is what I meant when I characterized the CAUCHY convergence criterion as a "premature discovery"—a characterization which at the same time means a praise of its discoverers.

I. GRATTAN-GUINNESS could have made a relatively strong point against CAUCHY out of the argument that the CAUCHY convergence criterion fitted less into CAUCHY's work than anything else. Strangely enough he did not. Though he challenged CAUCHY's originality in much weaker cases, he did not do so in this one, which would have been the strongest.

Though I cannot exclude the possibility that CAUCHY borrowed his convergence criterion from BOLZANO, I stress that I do not see any indication that he actually did so.

---

* CAUCHY, Oeuvres (2) **3**, 379; in the proof of the theorem of the intermediate zero of a continuous function.

4.3. Bolzano's Main Theorem. The theorem on the vanishing of a continuous function between two points where its values are of opposite sign is still less fundamental to CAUCHY's Calculus. It is almost self-evident that such a pure existence theorem did not mean much at that time. In CAUCHY's *Cours d'Analyse* it stands in the classical constructive context of solving numerical equations, particularly in connection with a method of LEGENDRE*, cited by CAUCHY**. The theorem itself had long been known. BOLZANO's and CAUCHY's merit is to have proved it. I. GRATTAN-GUINNESS' statement that CAUCHY's proof uses a condensation argument is far off the mark if by "condensation argument" he means what is usually understood by this term. His claim that CAUCHY's proof

> seems very much like an unrigorous version of the intricate proof developed in BOLZANO's paper

is as wrong as can be. The most convincing though somewhat lengthy way to refute this claim is to quote CAUCHY himself***:

> **Théorème I.** — *Soit $f(x)$ une fonction réelle de la variable x, qui demeure continue par rapport à cette variable entre les limites $x=x_0$, $x=X$. Si les deux quantités $f(x_0)$, $f(X)$ sont de signes contraires, on pourra satisfaire à l'équation*
>
> $$f(x)=0 \tag{1}$$
>
> *par une ou plusieurs valeurs réelles de x comprises entre x_0 et X.*
>
> *Démonstration.* — Soit x_0 la plus petite des deux quantités x_0, X. Faisons
>
> $$X-x_0=h,$$
>
> et désignons par m un nombre entier quelconque supérieur à l'unité. Comme des deux quantités $f(x_0)$, $f(X)$, l'une est positive, l'autre négative, si l'on forme la suite
>
> $$f(x_0),\quad f\left(x_0+\frac{h}{m}\right),\quad f\left(x_0+2\frac{h}{m}\right),\quad \ldots,\quad f\left(X-\frac{h}{m}\right),\quad f(X),$$
>
> et que, dans cette suite, on compare successivement le premier terme avec le second, le second avec le troisième, le troisième avec le quatrième, etc., on finira nécessairement par trouver une ou plusieurs fois deux termes consécutifs qui seront de signes contraires. Soient
>
> $$f(x_1),\quad f(X')$$
>
> deux termes de cette espèce, x_1 étant la plus petite des deux valeurs correspondantes de x. On aura évidemment
>
> $$x_0<x_1<X'<X$$
>
> et
>
> $$X'-x_1=\frac{h}{m}=\frac{1}{m}(X-x_0).$$

* M.-A. LEGENDRE, Essai sur la théorie des nombres. Supplément, février 1816, § III.

** CAUCHY, Oeuvres (2) 3, 381.

*** CAUCHY, Oeuvres (2) 3, 378–380.

Ayant déterminé x_1 et X' comme on vient de le dire, on pourra de même, entre ces deux nouvelles valeurs de x, en placer deux autres x_2, X'' qui, substituées dans $f(x)$, donnent des résultats de signes contraires, et qui soient propres à vérifier les conditions

$$x_1 < x_2 < X'' < X',$$
$$X'' - x_2 = \frac{1}{m}(X' - x_1) = \frac{1}{m^2}(X - x_0).$$

En continuant ainsi, on obtiendra: 1° une série de valeurs croissantes de x, savoir

(2) $$x_0, \quad x_1, \quad x_2, \quad \ldots;$$

2° une série de valeurs décroissantes

(3) $$X, \quad X', \quad X'', \quad \ldots,$$

qui, surpassant les premières de quantités respectivement égales aux produits

$$1 \times (X - x_0), \quad \frac{1}{m} \times (X - x_0), \quad \frac{1}{m^2} \times (X - x_0), \quad \ldots,$$

finiront par différer de ces premières valeurs aussi peu que l'on voudra. On doit en conclure que les termes généraux des séries (2) et (3) convergeront vers une limite commune. Soit a cette limite. Puisque la fonction $f(x)$ reste continue depuis $x = x_0$ jusqu'à $x = X$, les termes généraux des séries suivantes

$$f(x_0), \quad f(x_1), \quad f(x_2), \quad \ldots,$$
$$f(X), \quad f(X'), \quad f(X''), \quad \ldots$$

convergeront également vers la limite commune $f(a)$; et, comme en s'approchant de cette limite ils resteront toujours de signes contraires, il est clair que la quantité $f(a)$, nécessairement finie, ne pourra différer de zéro. Par conséquent on vérifiera l'équation

(1) $$f(x) = 0,$$

en attribuant à la variable x la valeur particulière a comprise entre x_0 et X. En d'autres termes,

(4) $$x = a$$

sera une *racine* de l'équation (1).

Cauchy's proof is simply a faithful description of the naive procedure for solving equations numerically (the title of this *Note* is "*Sur la résolution numérique des équations*"). The only sophistication is that the length of the unit interval is replaced by a more general h, and the 10 of our decimal system by a general basis m.

The proof is not a version of Bolzano's and it is as rigorous as a proof can be. The only correct remark I. Grattan-Guinness made is that Bolzano's proof is intricate; it goes by way of the existence of the least upper bound of a bounded set (or rather the existence of the cut number); once this existence is presumed, BOLZANO's proof is more elegant than CAUCHY's.

Anyhow there is not the slightest need to suppose that CAUCHY took his proof from BOLZANO. The idea, however, that such a theorem needed a proof and could be proved, may well have come from BOLZANO. The title of BOLZANO's pamphlet could have been enough to inspire CAUCHY to prove the theorem even if he never read the pamphlet itself.

Of course this does not prove that CAUCHY ever saw BOLZANO's pamphlet.

4.4. Bolzano's Lemma. The corner stone in I. GRATTAN-GUINNESS' case that CAUCHY plagiarized BOLZANO, is the following argument: In his *Cours d'Analyse*, instead of the limit concept, which would have been sufficient, CAUCHY used the concept of upper limit, which was not needed, simply because he found it in BOLZANO's pamphlet. If this were true, it would, indeed, prove convincingly that CAUCHY knew BOLZANO's pamphlet.

It was pointed out to I. GRATTAN-GUINNESS that his statement here rests on a few mathematical errors. In I. GRATTAN-GUINNESS' paper we now find a text (section 2.4), which, mathematically and historically, is wrong, as I will show in all details; further, attached to this text, footnote 24, which in fact invalidates the main text, and which is wrong in itself. I will now analyze this paragon of confusion.

As I explained, BOLZANO proved in his pamphlet the existence of the least upper bound of bounded sets of a special kind (DEDEKIND lower classes). I. GRATTAN-GUINNESS quotes BOLZANO's text and then continues:

> with this extraordinary theorem came another new idea into analysis, completely untypical of its time: the upper limit of a sequence of values.

Speaking of upper limit rather than of least upper bound could be a terminological deviation, since for a long time usage here was unsettled. It is certain, however, that I. GRATTAN-GUINNESS means "upper limit" since he refers to a sequence rather than to a set or a lower class, and since he continues with a reference to a convergence test of CAUCHY, the $\sqrt[n]{u_n}$-criterion for the convergence of $\sum u_n$ (with positive u_n). Here, indeed, the upper limit (that is, in modern terms, the largest accumulation value) is needed and is used. I. GRATTAN-GUINNESS says that the term of upper limit is

> ... not to be found explicitly in Cauchy's *Cours d'Analyse*, but instead we have there a frequent use of phrases like "...the largest value of the expression..."

This is entirely wrong. At one of the places alluded to by I. GRATTAN-GUINNESS we read*

> Cherchez la limite ou les limites vers lesquelles converge, tandis que n croît indéfiniment, l'expression $(u_n)^{1/n}$ et désignez par k la plus grande de ces limites, ou, en d'autres termes la limite des plus grandes valeurs de l'expression

* CAUCHY, Oeuvres (2) 3, 121.

dont il s'agit. La série (1) sera convergente si l'on a $k<1$, et divergente si l'on a $k>1$.

At another place*:

> Cherchez la limite ou les limites vers lesquelles converge, tandis que n croît indéfiniment, l'expression $(\varrho_n)^{1/n}$. Suivant que la plus grande de ces limites sera inférieure ou supérieure à l'unité, la série (3) sera convergente ou divergente.

The alternative definition is here repeated in the proof of the theorem:

> Considérons d'abord le cas où les plus grandes valeurs de l'expression $(\varrho_n)^{1/n}$ convergent...

It is difficult to say which one of the two definitions was operative, since the proofs do not use the explicit value of the upper limit but only its being <1 (or >1), that is, the existence of an U such that $(u_n)^{1/n}<U<1$ for almost all n $\big((u_n)^{1/n}>U>1$ for infinitely many $n\big)$. Contrary to I. GRATTAN-GUINNESS' statement the term of upper limit *(la plus grande de ces limites)* is explicit in CAUCHY'S text. On the other hand the plural form and the context "*la limite des plus grandes valeurs de l'expressions*" clearly show that this is not CAUCHY'S terminology for the upper limit as suggested by I. GRATTAN-GUINNESS' quotation "the largest value of the expression..." Cut out this way from CAUCHY'S text by I. GRATTAN-GUINNESS, it is meaningless because it does not allow the hidden quantifiers to be traced.

It does not matter too much what artificially isolated pieces of a text mean if the text is globally clear; in the present case it is not far-fetched, and it is in agreement with the global text to assume that "la plus grande valeur" applies to a finite set, to wit the set of $(u_n)^{1/n}, \ldots, (u_{n+k})^{1/n+k}$, and the plural is to indicate that all such sets are considered.

I. GRATTAN-GUINNESS continues:

> As with continuity of a function, CAUCHY was revealingly only partially aware of the significance of the idea; for he used it only as a tool for developing the proofs of his particular theorems and not as a profound device for investigating more sophisticated properties of analysis. Therefore it would be especially surprising if it were CAUCHY'S own invention...

Everybody who is not a stranger to calculus knows that there is no other use of upper limits than just those theorems where CAUCHY used them. Even today they provide an unusual and ineffective device. The conclusion that it was not CAUCHY'S invention because he used it too little is consequently mistaken. I. GRATTAN-GUINNESS still suggests that CAUCHY took this tool from BOLZANO. When he wrote that sentence, he certainly believed that this tool was in BOLZANO'S pamphlet. Probably he was misled by the so-called BOLZANO-WEIERSTRASS Theorem on the existence of an accumulation point for an infinite bounded set of numbers, which can be proved by showing the existence of the upper limit.

* CAUCHY, Oeuvres (2) 3, 235.

BOLZANO's name in this context, however, is an honorific rather than an historic epithet as is HEINE's name in "HEINE-BOREL theorem"*.

CAUCHY did not use the notion of upper limit more often than he did, because he could not**, and he did not take it from BOLZANO, because it was not in BOLZANO's pamphlet. There is no doubt that I. GRATTAN-GUINNESS now knows these facts, but instead of cancelling the whole section, he has nullified it in a footnote:

> There is a distinction between BOLZANO's introduction of an upper limit and CAUCHY's "largest value of the expression..." in that CAUCHY actually used the *Limes* of a sequence... while BOLZANO defined the upper limit... but we cannot interpret this distinction as intentional in BOLZANO's and CAUCHY's time...

First, neither did CAUCHY use the term "largest value of the expression" nor did BOLZANO speak of upper limits. According to modern terminology the terms are upper limit (or limit superior) and least upper bound (or cut number), respectively. Second, CAUCHY does not use the limit but the upper limit—I. GRATTAN-GUINNESS seems still not to grant that these are different things. Third: Both BOLZANO's and CAUCHY's concepts of least upper bound and upper limit, respectively, were introduced on purpose because in the given context neither of them could use any other concept.

The fact that at first I. GRATTAN-GUINNESS did not notice this distinction, does not entitle him to claim that BOLZANO and CAUCHY could not make it. They did not have to, because they were confronted with different situations, and it is no use asking whether they would have made the distinction if there had been some need to do so.

To summarize, at this point there is no influence of BOLZANO on CAUCHY visible.

4.5. The Real Number System. I. GRATTAN-GUINNESS says:

> In the course of proving this Lemma as well as in other parts of his paper BOLZANO had recourse to extended considerations of real numbers regarding the rational or irrational limiting values of sequences of certain finite series of rationals...

On the contrary:

> CAUCHY wrote just once on the real number system: it was in the *Cours d'Analyse*, where he gave a superficial exposition of the real number system. The initial stimulus for this work was foundational questions concerning the representation of complex numbers; but he took the development of the ideas well into BOLZANO's territory, twice including the remark that "when *B* is

* HEINE first recognized the importance of uniform convergence, but he did not formulate covering properties.

** Even a concept like the least upper bound was not of any importance for the mathematics of the CAUCHY era. Such concepts become instrumental only with the direct methods of the variational calculus at the end of the 19th century, in particular after HILBERT's salvation of DIRICHLET's principle.

> an irrational number one can obtain it by rational numbers with values which are brought nearer and nearer to it"—merely a remark on a property of the real numbers and not as a *definition* of the irrational number... Once again CAUCHY did not fully appreciate the depth of BOLZANO's thought; and yet it is clear from his partial success that he was aware of BOLZANO's ideas rather than from his partial failure that he was ignorant of them.

It is hard to believe, but the truth is just the other way round. It is true that neither BOLZANO nor CAUCHY defined real numbers (in later investigations BOLZANO tried to do so). There is, however, nothing in BOLZANO's pamphlet that justifies the sentence quoted. There are no "extended considerations on real numbers...", there is not *any* consideration of real numbers and not even anything that could be misunderstood as such by somebody unaccustomed to reading mathematics. What I. GRATTAN-GUINNESS writes is a pure invention. The terms "rational" and "irrational" do occur once, in § 8, when, using as an example the decimal development of $\frac{1}{9}$, BOLZANO warns the reader against believing that the limit of a sequence of different rational numbers must be irrational.

On the contrary, CAUCHY's occupation with real numbers in the *Cours d'Analyse* is hatefully misrepresented. CAUCHY, though not defining real numbers, at least defines the algebraic and exponential operations on real numbers; starting from the rational numbers, where they had been defined directly, he extends the definitions to the real numbers by continuity. In this context he twice uses the fact that real numbers can be obtained as limits of rational ones. These are not isolated remarks as I. GRATTAN-GUINNESS claimed, but rather a deliberate use of this property in a meaningful context.

In any case CAUCHY wrote in the *Cours d'Analyse* much more on real numbers than BOLZANO did in his pamphlet (which was nothing). What could CAUCHY learn at this point from BOLZANO? What was the "depth of BOLZANO's thought" that CAUCHY could not fathom? The bare Nothing or the fact that 0.111... is rational? Where did he trespass into BOLZANO's territory, if this territory consisted of Nothing or of the fact that 0.111... was rational?

4.6. Summary as to the Common Ideas in Bolzano and Cauchy.

1. The idea of continuity, common to them both, was arrived at by each of them independently.

2. The CAUCHY convergence criterion was formulated by each of them; it is possible that CAUCHY took it from BOLZANO, though it can easily be explained as an original invention of CAUCHY's.

3. The theorem on the intermediate value of a continuous function had long been known as a more or less obvious proposition. The idea to prove it may have come to CAUCHY when he read the title of BOLZANO's pamphlet if he ever did. His proof is different from BOLZANO's.

4. As regards upper limits and least upper bounds, there is no common element.

5. On real numbers BOLZANO's pamphlet contains nothing, while CAUCHY in his *Cours d'Analyse* developed a theory of operations with real numbers.

In section 2 I explained how the *Cours d'Analyse* rested on a much broader basis of ideas than the few CAUCHY could have borrowed from BOLZANO's pamphlet. Therefore I. GRATTAN-GUINNESS' insinuating question*

What would have happened if CAUCHY had *not* read BOLZANO?

is irrelevant. The present section shows that there is even little if any cause to ask the other insinuating question**

But if CAUCHY owed so much to BOLZANO, why did he not acknowledge him?

Before analyzing his answer on this question, we shall cast a glance at his section 3.

5. Limit-Avoidance

I quote I. GRATTAN-GUINNESS' new limit definition***:

> When we speak of "introducing the concept of a limit" into analysis, we are actually introducing limit-avoidance, where the limiting value is *defined* by the property that the values in a sequence avoid that limit by an arbitrarily small amount when the corresponding parameter [the index n or the sequence s_n of n-th partial sums, say, or the increment α in the difference $(f(x+\alpha)-f(x))$ for continuity] avoids its own limiting value (infinity and zero in these examples). The new analysis of BOLZANO's pamphlet and developed in CAUCHY's text-books was nothing else than a complete reformulation of the whole of analysis in limit-avoidance terms...

No, no, and no. BOLZANO and CAUCHY knew better than I. GRATTAN-GUINNESS what was convergence and what was continuity. It is true there are bad 19^{th} century textbooks where you can find such silly definitions, but this was neither BOLZANO's fault nor CAUCHY's****.

6. Cauchy's Character

To explain *why* CAUCHY plagiarized BOLZANO, I. GRATTAN-GUINNESS writes a story about what he calls the Paris clique of mathematicians. No doubt he has studied that *chronique scandaleuse* better than anybody else. But if the secrets of that society are as relevant to understanding the history of mathematics as he suggests, why does he wrap himself in veils of mystery rather than disclose them? Why does he concoct a pompous story from plain historical facts and unfathomable allusions?

Whoever has studied CAUCHY's work knows how chaotic it is. A proposition is stated, then refuted, only to be stated once more; a procedure is severely criticized, only to be applied successfully at the next opportunity; for no reason

* p. 383, 12 f.b.

** p. 387, 5.

*** p. 378, 13 f.b. — 5 f.b.

**** When I. GRATTAN-GUINNESS lectured at the Utrecht Mathematical Colloquium everybody protested. An hour later people thought they had convinced him. It is a pity they had not done so.

notations are changed back and forth. No, I. GRATTAN-GUINNESS says, stating a certain apparently wrong theorem was a strategic move in the secret game of the Paris clique. As long as I do not know the secret information on which such conclusions must be based, I cannot challenge them*.

A critic is on a safer ground when I. GRATTAN-GUINNESS gives his sources. To prove that CAUCHY took sides in the quarrels of the "Paris clique" (which is utterly improbable) he mentions, in the same work, "fawning references to the powerful secrétaire perpétuel (FOURIER)" and "attacks on the declining POISSON"**. Any one who checks the sources will find that neither is the reference to FOURIER fawning nor is POISSON attacked. The first reads

> si l'on désigne avec M. FOURIER avec $\int_{x'}^{x''} f(x)\,dx$ l'intégrale définie, prise entre les limites $x=x'$, $x=x''$...

and it is the style in which such acknowledgements have been made a thousand times by mathematicians. At the second place quoted we find CAUCHY, rather than attacking POISSON, explaining why he had overlooked certain consequences of his theory which had meanwhile been discovered by POISSON.

To understand what citations mean for mathematicians, it would be worthwhile to make a statistical study of them, say around CAUCHY. Isolated examples are of little value. At the very period when, according to I. GRATTAN-GUINNESS, CAUCHY had reasons to fawn FOURIER and to attack POISSON, he used the introduction to his *Cours d'Analyse* to extend his thanks to LAPLACE and POISSON, who had advised him to publish his courses, and at the end of the same introduction he acknowledged the good counsel he had received from POISSON, AMPÈRE and CORIOLIS. Should we interpret these acknowledgments, too, as attacks?

It is well known that CAUCHY was a strange fellow, and to prove it, there is no need to invent strange stories about him. The strangest is his quixotic conduct after the July revolution of 1830, when as a lone paladine he followed his king to his exile court in Prague. He was a religious and political dogmatic who often exhibited an appalling lack of human relations.

---

* A characteristic pomposity is the remark in footnote 85 that the *Procès verbaux des séances de l'Académie tenues depuis la fondation jusqu'au mois d'août* 1835 (10 vols; 1910–22, Hendaye) "are an invaluable source of historical insight into the period 1795–1835, when the rivalries were at their height. They give the minutes of all the private meetings of the *Académie des Sciences*, which the participants can hardly have expected to be published!"

In fact, there is little that might be regarded as sensational to be found in the *Procès verbaux*. The style is the same as that of the later *Comptes Rendus*; the greater part is routine business. The meetings were not private but public. All spontaneous remarks were afterwards carefully edited or omitted; the oral text is better reflected by the newspaper reports.

** CAUCHY, Oeuvres (1) **1**, 340 and 189–191; another source mentioned is not accessible to me.

The adjectives "powerful" and "declining" are melodramatic stereotypes. There has never been any *secrétaire perpétuel* who was not powerful, but I doubt whether FOURIER was more so than his predecessors or successors. Facing a powerful *secrétaire perpétuel*, POISSON, too, needed an adjective though it is a pity that I. GRATTAN-GUINNESS hit on one that is so trivially mistaken as is "declining".

There is a story about CAUCHY and a manuscript of ABEL. In 1826, when his first important work had yet to appear, ABEL visited Paris. A few times he met CAUCHY, who at that period was interested only in mathematical physics. In Paris ABEL wrote the famous work he presented to the French Academy in October 1826. In 1829 he died. In the late thirties the editor of his *Oeuvres*, who knew about the manuscript, tried to get it back from the Academy, but it could not be found. Suddenly, in 1841, the text of the manuscript appeared in print in a publication of the Academy, though, strangely enough, the manuscript itself was still lost.

This trackless manuscript has always been an exciting feature in the melodramatic life of ABEL, who according to the stories died in misery, oblivion, and disappointment. (It has long been known that this story is untrue*.)

In such a story a villain is needed. According to old LEGENDRE, ABEL's paper was illegible, so the referees, CAUCHY and himself, could not read it. Even today it is commonly believed that the manuscript was lost by CAUCHY's neglect. In 1922 a copy of CAUCHY and LEGENDRE's report on ABEL's paper, dated 29 June 1829, was discovered**; it proved that CAUCHY's account of his role in the story was correct. It is obvious that CAUCHY had no further business with ABEL's manuscript, since after the July revolution of 1830 he went abroad and did not return before 1838. The academician LIBRI, however, who to annoy other people, had invented the main facts in ABEL's melodramatic life, got some business with ABEL's paper; in any case he read the proofs, though according to him without the manuscript. LIBRI was a mediocre mathematician who became famous by his sudden departure to London in 1848, when he was accused of having over many years stolen from the French public libraries a million's worth of rare books and manuscripts. Thus it was not too far-fetched to look into LIBRI's estate in the Moreniana library in Florence. Finally, in 1952, VIGGO BRUN did so, and he found ABEL's manuscript***. A written explanation of it by LEGENDRE had been published in World War II**** but had not been noticed. It reads†:

> Ce Mémoire a été mis d'abord entre les mains de M. Le Gendre qui l'a parcouru, mais voyant que l'écriture étoit peu lisible et les caractères algébriques souvent mal formés, il le remit entre les mains de son confrère, M. Cauchy avec prière de se charger du rapport. M. Cauchy distrait par d'autres affaires et n'ayant reçu nulle provocation pour s'occuper du Mémoire de M. Abel, attendu que celui-ci n'était resté que peu de jours à Paris après la présentation de son Mémoire à l'Académie, et n'avait chargé personne de suivre cette affaire auprès des commissaires, M. Cauchy, dis-je, a oublié pendant très long temps le Mémoire de M. Abel dont il étoit dépositaire. Ce n'est que vers

* Read VIGGO BRUN's debunking paper in Journal r. u. angew. Math. **193** (1954), 239–249.

** D. E. SMITH, *Amer. Math. Monthly* **29** (1922), 394–5. Among my autographs, 29. Legendre and Cauchy sponsor Abel. — It is in agreement with the Procès verbaux (*cf.* footnote *, p. 389).

*** See footnote *.

**** G. CANDIDE, Sulla mancata pubblicazione, nel 1826 della celebre Memoria di Abel. Tip. Editr. "Marra" di G. Bellone, Galatina 1942, XX.

† Journ. r. u. angew. Math. **193**, 244–245.

le mois de mars 1829, que les deux Commissaires apprirent, par l'avis que l'un d'eux réçut** d'un savant d'Allemagne, que le Mémoire de M. Abel, qui avait été presenté à l'Académie, contenait ou devait contenir des résultats d'analyse fort interessants, et qu'il était étonnant qu'on n'en eût pas fait de rapport à l'Académie. Sur cet avis M. Cauchy rechercha le Mémoire, le trouva et se disposait à en faire son rapport; mais les Commissaires furent retenus par la considération que M. Abel avait déjà publié dans le Journal de Crelle une partie de son Mémoire présenté à l'Académie, qu'il continuerait probablement à faire paraître la suite, et qu'alors le rapport de l'Académie, qui ne pouvait être que verbal, deviendrait intempestif*.

Dans cet état de choses nous apprenons subitement la mort de M. Abel, perte très fâcheuse pour les sciences, et qui paraît maintenant rendre le rapport nécessaire pour conserver s'il y a lieu, dans le receuil des savants étrangers, un des principaux titres de gloire de son auctor**.

This unveils the mystery around ABEL's manuscript. It is not unusual that referees neglect their task, in particular, if they are not interested in the subject or if it is the work of a virtually unknown author, though I agree that CAUCHY was usually more careful. Delays of 10–15 years in printing treatises accepted by the French Academy were not unusual either; every publication needed a royal authorization. In ABEL's case it may have played a role that the essential part of the manuscript had already been published in "Crelle's Journal".

I. GRATTAN-GUINNESS' report on this event is a distortion of the story as it is known now. He omits all evidence that is in favour of CAUCHY, and he falsifies two points***:

First he claims that the neglected manuscript

> ...was the paper which ushered in the transformation of LEGENDRE's theory of elliptic integrals into his own theory of elliptic functions...

to add one more melodramatic feature. The paper on elliptic functions was published in Crelle's Journal. The manuscript in question was about "ABEL's theorem"; an extract also appeared in Crelle's Journal.

Second, he claims:

> CAUCHY took it and, perhaps because of ABEL's footnote against him, ignored it entirely: only after ABEL's death in 1829 did he fulfil a request to return it to the *Académie des Sciences*.

The reader can check that this is in all essentials contrary to LEGENDRE's report. If I. GRATTAN-GUINNESS is in the possession of secret information that refutes LEGENDRE's report, he should reveal his sources. Meanwhile I am entitled to consider LEGENDRE's report as correct.

---

* The procedure of a formal report was applied only to manuscripts; printed pieces submitted to the Academy were given a *rapport verbal.*

** *Sic.*

*** p. 393.

I. GRATTAN-GUINNESS continues:

> ...there is one aspect of it which has been little remarked upon but which shows the depths to which CAUCHY could sink.

The evidence I. GRATTAN-GUINNESS produces for CAUCHY's moral downfall is an exposé of 1841, where CAUCHY first praises ABEL and then refutes the story that ABEL died in misery. We now know that CAUCHY's exposé is correct.

I. GRATTAN-GUINNESS does not explain in what CAUCHY's downfall consisted, but anyhow it was a downfall and

> ...anyone capable of writing in this manner, knowing the negative role played by himself in the matter under discussion, would hardly think twice about borrowing from an unknown paper published in Prague without acknowledgment.

Anyone? Maybe. But CAUCHY was someone.

Mathematical Institute
Rijksuniversiteit
Utrecht

(Received February 1, 1971)