Patterns of Mathematical Thought in the later Seventeenth Century

DEREK THOMAS WHITESIDE

Communicated by C. B. BOYER *and* I. B. COHEN

To Michael Hoskin, without whom this would probably never have come to exist, with deepest thanks

Contents

Foreword

The original impulse which led to the research incorporated in the following essay was the desire to probe into the philosophical basis of concepts, especially those of number, space and limit, which were the keystone of the immense proliferation of mathematical discoveries during the 17th century. With wider knowledge of the original texts, manuscript and printed, and through deeper appreciation of the complexities involved, that impulse became modified into a more restricted and concrete shape: the study of the particular mathematical forms which developed in the 17th century with emphasis on their interconnections rather than on their philosophical aspects.

There were many reasons for this change of interest. In particular, there exists a great richness of material bearing on developments in technique as against a paucity of anything which can be interpreted as original comment on underlying structure or methods of proof, and a change of viewpoint brought with it an immense increase in study-material. Moreover, to some extent I found myself captivated by the beauties and intricacies of solutions given to particular problems—the BROUNCKER continued fraction is an example—which in the original plan for study could have found place only as a set of unwieldy appendices. But, above all, I became convinced that by the 17th century mathematical structures had become too systematised and too remote from any possible physical origins to allow any further incursion of concepts from without, that mathematical development took place almost entirely within its own tight field, and that therefore extended discussion of a philosophical background, existing or postulated, would be largely irrelevant.

This is not to deny in any way the immense influence which mathematical technique had in other fields, and especially, at that time, in giving a precisely definable numerical basis to physical reality through closely tied concepts of spatial and temporal dimension, force, mass and weight, and that with the 17th century quantitative rather than qualitative examination of natural phenomena becomes significant. The crucial point here is that the mathematical structures set up to mirror aspects of physical reality were taken over whole, suitably and ingeniously interpreted but unmodified. Thus, NEWTON'S proof that a point traversing an elliptical path is directed to a focus by an attraction varying as the inverse square of its distance away from it is a strict deduction from purely mathematical premisses, elaborated for the most part in Greek times but with the novelty of a concept of geometrical fluxion due to NEWTON himself. A great deal may be said in extramathematical justification of the physical interpretation of this result—and was indeed said at great length at the close of the 17th century—but we can assess its mathematical importance and validity only within the very narrow conceptual frame within which it was evolved.

Along with this virtual rejection of a viewpoint which emphasises extramathematical aspects of mathematical advance, in the more technical, particularised discussion given I have neglected a prevailing fashion which sees mathematics as a mere handmaiden to the sciences, and the 17th century scientific achievement as a revolution in which scientific thought was freed from the largely sterile dominance of scholastic authority under a universal guiding principle of the primacy of theory induced from observed instances in phenomena. Though many historians are now willing to search out the tangled complexity which is the 17th century scientific achievement rather than reinforce a simplicity which it never had, to see the period as less original in thought than it claimed to be and vastly more indebted to previous centuries—in short, to strip away all the irritating mystique which has in the past surrounded the "scientific revolution"—to consider mathematical development in the context of its scientific influence seems too external a study. Rather, I have found myself returning to the detailed analysis of mathematical concepts which has, since MONTUCLA, been characteristic of the technical histories of mathematics.

In the two centuries since the *Histoire des mathématiques* was first published the technical historians have, through repeated revision and addition, gradually built up a store of hard fact together with exact reference tc original manuscript and text. Such amassing of incontrovertible detail is possible in mathematics, for centuries the model and inspiration of exact thought, —perhaps more so than in any other intellectual study—but the danger of such an approach is that our ideas on and evaluations of particular mathematical forms and periods of advance become solidified, that we continue to accept an undisputed historical fact as important when it is completely trivial. The great need is for the continued introduction of new approaches and fresh insights along with factual additions. It is perhaps fortunate, therefore, that with the rapid growth of mathematics the all-inclusive descriptive account of development, for so long the historians' ideal, is no longer possible. Today a growing importance, reflected in the increasing number of histories of particular mathematical concepts, is attached to the historical study of methods in mathematics, an approach in which details, rather than existing as the primary object of study, are chosen to highlight significant points and aspects of conceptual development. There, however, the tendency is to be imprecise—to tailor the niceties of historical development to an oversimplified interpretation of available fact, disregarding inconsistencies as unimportant if not trivial. The great problem would appear to be to isolate significant trends of development without denying—and leaving the way open to modification by—the richnesses, idiosyncracies and reduplications which seem concomitant with any widening of the boundaries of human experience, and within the context of 17th century mathematical advance I have tried to resolve it.

In this essay a detailed analysis is given of aspects of later 17th century mathematics, some of which—especially the calculus—have been extensively studied, while others—such as synthetic geometry—hardly at all. Wherever possible manuscript and original printed material has been used to give added insight into more familiar sources. The restriction of geographical area to Britain is made largely to give a workable study-field rather than to insist on the separateness of English mathematics in the period. In fact, of course, many of the English mathematicians received a training on the Continent—as JAMES GREGORY—or through Continental literature—as WALLIS and NEWTON—and English mathematics is to be characterised more by certain localisations of interest than as a separate entity. Further, in many cases it has been impossible to give a comprehensive discussion except by including details of non-English developments.

To some extent the verbal text is independent of the numerous examples included in it. These, however, do something more than illuminate the general themes developed—by their mutual dependence on each other for proof they impress the fact that mathematics had then become an integral structure. More often than not a sketch of the proof given—and where necessary a complete account—is inserted as well as a description of the result itself. With few exceptions historians have in the past considered it not very important to study outdated forms of proof, considering them—if at all—the subject matter of logic and preferring to substitute modern proofs. From the present viewpoint, however, the proof-structure is at least as important as the particular result obtained by it, and it becomes possible many times to see how the inadequacy or lack of proof-structures conditioned the development of whole classes of results. For the most part—notably in examining the method of exhaustion—where the original notations would seem to obscure ideas which can be clarified in appropriate symbolism, anachronistic notation is used. This concession to concise expression and to understanding was not made without hesitation, but rather than become involved in an intricate study of the modifying influence of symbolism it seemed preferable to substitute a cautious use of modern notation for the often unnecessarily cumbrous original.

One final personal remark may be not out of place. Working with a wide range of written and printed material, it is very tempting to base a final judgment on the written word alone (in the form of reference notes) without trying to recapture the thought which underlies it, to write mere textual criticism without attempting a wider view. The word, whether in print or manuscript, is there before us, pleasantly concrete and unchanging, fixed in form but for a possible dubious reading, misprint or contradictory alternative draft. Its existence is independent of any commentary we may choose to make on it, and it must therefore be treated with the utmost respect. In contrast, the thought which a word is designed and chosen to convey seems often a vague, fleeting and almost illusory thing, rough and inexact in the freshness of inspiration and so often seeming to escape the net of a precise definition. Indeed, the very independence of a word form with its attached layers of conventionally accepted meaning can make any adequate expression of the thought almost insuperably difficult. But we must try to go beyond the written word, accepting its inadequacy as a means of expression, and—since there can be no personal appeal to the author for clarification of a 17^{th} century text—make a leap into darkness, however considered, in the attempt to bridge the chasm between word and concept. From growing familiarity with the work, especially in manuscript, of individual writers and with the effort to see into their minds there appears gradually, along with the excitement of recreating a process of thought and the pleasure of seeing a way through some difficulty, a very complex web of impressions and convictions, barely tangible and ever ready to be broken, which it pleases us to see as the truth. To penetrate further into this process would be to enter on a study of the psychology of understanding and belief, but unless we use the intricate pattern of knowledge, often felt as much as intellectually perceived, which crystallizes out our criticism may often be inadequate. Not always may we be able fully to document some insight —though we must always try—and in the absence of a factual basis it can seem worthwhile to formulate hypotheses.

In conclusion, I am much indebted for material on NEWTON'S mathematical thought to original papers in the Portsmouth Collection deposited in the University Library, Cambridge. Other acknowledgements are made in footnotes to the text and, more generally, in the bibliography. I would like to acknowledge my debt to my thesis supervisor, Dr. M.A. HOSKIN, for the warmth of his encouragement at all times, and to Professor R.B. BRAITHWAITE, who sponsored me in the all-important first year of my research. Finally, I extend my thanks to the librarians of the University Library, Cambridge, of the Bodleian, of the British Museum and of Trinity College, Cambridge for the generous access to original documents and rare texts allowed to me.

I. The "mathematickal art": basic elements and philosophical attitudes

At each except the most primitive level mathematical thinking has been something more than a mere calculating routine whose only criterion of value is that it gives an answer to a problem. Since Greek times each succeeding generation has inherited an increasing bulk of concepts, techniques, unsolved problems and paradoxes, often mingled in a bewilderingly disordered way. Above all, at the beginning of the 17th century the inheritance was almost too rich and too confusing, compounded of elements from Greek, Arabic and medieval sources as well as from contemporary Europe which were part mathematics, part philosophy, part religion, part mysticism, part literature. It must have seemed at times an insuperable task to see a way through it all, but within a century that great mass of inconsistent elements had formed a richly suggestive amalgam which was the foundation for the more unified mathematical advances of modern times. Since the 17th century there has been no significant external influence on the growth of this European tradition of mathematics, and with its roots now spread throughout the world, none would seem possible—which takes from its colourful side perhaps, but adds immeasurably to its firmness and solidarity.

As a preliminary, however, to a discussion of certain aspects of the contributions of the 17th century to this tradition—particularised, though not absolutely, to the latter part of the century in time, and to the school of English mathematics which centred on Cambridge, Oxford and London in geographical location—an outline of the basis on which these achievements rested and depended is not out of place.

The clearer insight into proof structures and deductive procedures which has come with the vast elaboration of techniques of logical exploration in the last few decades now allows us to see certain tendencies as valuable and to ignore others as being merely the product of muddled thinking if not incomprehension. But its very success in exact symbolic formulation of most of the classical logical forms, its notational facilities which allow us to see the nature of a block to a process of thought and its axiomatic formulation of conditions which can remove such a block has, in one sense, made it difficult to see the value of outdated and inadequate forms of proof. In particular, since we are now able accurately to define in some suitable notation all the proof structures used in mathematics, we tend to judge past attempts at such definition by more or less the same standards, criticising a proof, perhaps, because an unstated axiom is used implicitly, or a deductive procedure because no exact definition of a limit procedure is formulated. We tend, too, to assume that mathematics has always been developed in abstraction from any model other than a logical one, forgetting, for example, that before the 19th century geometry was in part developed on the basis of conventional ideas of real physical space, and that it might in some ways be more fitting to see it as a theory of allowable transformations in space in the period before modern axiomatic treatments were developed. In fact, extramathematical ("psychological") considerations still play a large role in 17th century mathematical procedures, but thereby compensate for the apparent lack of rigour or loose assumptions rather than invalidate the proof forms used.

To one accustomed to the idea that exact proof-trees shall be set down in rigorous mathematical argumentation very few proofs of any kind in classical mathematics will be allowable, and certainly none were given in the 17th century on any but the most elementary numerical level. Rather, we would do well to criticize the form of a 17th century mathematical proof from the viewpoint that it is a psychologically satisfying sketch and no more. Such a proof does, in a very strong sense, prove a result which we find valuable and new (if only in the sense of not previously being seen logically to follow from the given structure), and in historical fact very often mirrors more adequately than a tight and rigorous modern form the thought-processes which led to its formulation. Mostly, too, it has a directness and immediacy—even a warmth and guilelessness—which is very often lacking in the cool surgical precision of its modern equivalent, and which is to be appreciated only through familiarity with 17th century mathematical writings. Perhaps the precision and rigour of the modern proof is obtainable only by sacrificing the lack of generality which is so often the basis for such feelings of immediacy, but it remains true that the particular results obtained by such methods seem largely justified at a heuristic level by the forms of deduction which were historically given for them and by which they were in most instances derived. It is unfortunate only that the plausible is not always true (or, at least, not probably true or false).

Those 17th century authors[1] who tried to make precise the nature of mathematics and mathematical argument for the most part accepted classical Greek theories of causality and proof. Partly this was due to the continued veneration of all things Greek, but the need for justification of deductive procedures had been felt from early Greek times. Whatever the debt to previous civilizations[2], Greek thinkers had squarely faced questions of mathematical existence, the nature of mathematical truth, the cogency of proof and its connection with the allied philosophical concept of causality; and the views of ARISTOTLE in his *Organon* and *Physics*, and to a lesser extent of PLATO in his *Republic*, but above all the model mathematical text of EUCLID'S *Elements* influenced attitudes to the nature of mathematics over the next two thousand years. ARISTOTLE'S main object[3] had been to codify something of the subtle and intricate way in which verbal

---

[1] In England especially BARROW, WALLIS, RAPHSON and NEWTON—specifically (BARROW) *LM*, given in 1664—1666 as the Lucasian lectures at Cambridge; (WALLIS) *MU* and *institutio logicae* ... which is virtually a university textbook on ARISTOTLE'S syllogistic canon, with medieval clarifications and additions of "fallacies" and "dilemmas"; (RAPHSON) *SR*; and (NEWTON) *AU*, especially preface and the introduction to the appendix *aequationum constructio linearis* (279*ff*.), and various drafts of an essay on proof-methods by analysis and composition in CUL Add. 3963.

[2] BRUINS, E. M., in: *On the system of Babylonian geometry.* Sumer **11** (1955): 44—49, developing ideas of F. THUREAU-DANGIN in his *Textes mathématiques babyloniens*, Leyden 1938, traces the beginning of a deductive system in Babylonian mathematics on the basis of extant texts containing area-formulas. Arguing that a concept of similarity and proportion is implicit in them and keeping in mind that no Babylonian words for such concepts as "angle" and "parallel" exist, he reconstructs a plausible proof of "PYTHAGORAS'" theorem connecting the sides of a (EUCLIDEAN) right triangle.

[3] *Cf.* J. LUKASIEWICZ: *Aristotle's syllogistic from the standpoint of modern formal logic*, Oxford, 1951; and J.M. BOCHENSKI: *Formale Logik*, München, 1956: 47—114.

language communicates meaning and especially the concept of propositional truth, and to that end in his *Organon* had developed a class-calculus theory of the syllogism. Elsewhere, but especially in the *Physics*, he had formulated views on number and infinity which were to influence medieval attitudes very strongly, and to be passed on to 17th century mathematicians through the scholastic commentaries rather than directly. PLATONIC viewpoints, after a lapse from favour in the later medieval period, became influential again with the Neoplatonist movement of the Renaissance, and most 17th century writers find PLATO's theory of ideal and real and the limits which his philosophy puts to sense-perception not unattractive. EUCLID, building on the work of EUDOXUS and other unknown systematizers, had restricted himself in the *Elements* to a specific programme which had for its ideal—if not wholly successfully carried out—an elaboration of elementary geometry on the basis of stated axioms (which were to be accepted as "self-evident") by deduction procedures which were those of any reasoned proof. The brilliance of his achievement made the *Elements* a model of mathematical reasoning and one still accepted as a guide throughout the 17th century, while the idea of axiomatic deductive proof, implicit only in the *Elements* but discussed explicitly in Greek, Arabic and European commentaries became an acceptable part of 17th century mathematical propaedeutic. Coalescing together in the 17th century, these three approaches to the nature of mathematics became a general eclectic attitude, differing to a greater or less degree with the individual exponent, but comprising well-defined elements. Mathematical reasoning was seen as a mental art rather than a physical one, with all the causal force and necessity and empirical unverifiability of a theoretical process, and mathematical creation took on a PLATONIC coat of inspiration from a divine intelligence, while a mixture of EUCLIDEAN axiomatics and ARISTOTLEAN syllogistic (in its developed scholastic forms) came to be accepted as a basis for practical reasoning.

Unfortunately, this fusion of classical theory seems to have been more a veneer of respectability than a living creative exploration of mathematical reality. Certainly, unlike the development of techniques of mathematical logic in the past century, it seems to have contributed nothing to mathematical advance, and is treated with mere casual respect by the professional mathematicians if not by outright impatience[4]. Typically BARROW[5] discusses the concept of mathematical proof and logical deduction, seeing the subject matter of mathematics as lying in the abstractions from the particular properties and affections of really existing phenomena—a process of abstraction not to be explained solely as a numerical induction from particular instances—and emphasising that mathematical structure must mirror that which exists as a basis for the real, perceptible world. Granted that the argument is put too baldly—BARROW, in fact, argues the case with the precision of a modern linguistic analyst, and very often in strange

[4] As the young JAMES GREGORY wrote: "I warn students of mathematics how futile is the attempt to promote mathematics by the aid of fictive philosophical reasons which are useful merely for influencing the common credulous throng; for in mathematics there is no logic except geometry, nor any philosophy which by geometry's help is not raised on infallible experiments" (see *VCHQ*: proemium: vi).

[5] *LM*: (1664): lectiones 4—8.

anticipation of his verbal fluency—there yet remains little for the practising mathematician but a faith on which to live, and certainly no guide to practical prosecution of the subject*. Moreover, the 17th century mathematician had faith enough in his own ability and the richness of the subject matter remaining to be explored not to be worried about the nature of proof and deductive cogency. For him, a series expansion showed the value and importance of mathematical investigation more than any inquiry into foundations, and it is significant that BARROW in his later work became interested in the creation of original mathematics to the exclusion of developed thoughts on the nature of mathematics[8].

It is easier and seemingly more worthwhile to inquire into the particular definitions and concepts which were accepted as basic and necessary in the study of more complex mathematical forms.

The concept of (positive) integer is fundamental in all numerical mathematics, and the standard way of introducing it is through a model in which some quantity, suitably defined, is used as a unit on which to measure ("count out") the rest. WALLIS gives a typical treatment in his *mathesis universalis*[9] suitably ordering

* This is not to deny that such a philosophical basis afforded very often a neat tie between metaphysics and psychology. The PLATONIC theory of sense-perception, insisting on the limits of observed reality and the supremacy of the ideal theoretical structure on which our perception of reality is based—as a flickering shadow cast by a fire on a wall, in PLATO's analogy, merely reflects the nature of the body casting the shadow—is attractively joined with the Aristotelian concept of actual (limit) infinity (which is strictly unobservable and so non-admissible), and potential (unboundedly large) infinity which mirrors a popular 17th century attitude that by suitably controlled experiment we can reach ever nearer to absolute truth. Closely allied was a growing feeling that physical space, structured mathematically, extended into infinity (a view itself justified by an appeal[6] to the concept of a free variable) —a scheme in which such conventional attributes of God as his being absolute, unknowable, all-including and all-pervading had a natural place. Indeed, on very much this basis is developed both the view of God in PASCAL's *Pensées* and the concept that God is to be equated with the whole of infinite space as a universal[7].

The view that mathematical structure should, in some way, mirror physical reality is, of course, basic to all schemes which apply mathematical techniques in analysis of the real, observable world. But there is something of the flatness and boredom of the obvious truth about it which can only be removed by making precise the nature of such contact of mathematical structure with observed reality, and on that point we would not expect to be enlightened by BENTLEY, RAPHSON and SAMUEL CLARKE (proponents of such views in the period in England) who are derivative in their mathematical ideas, however creative and provoking in the fields of philosophy, religion and literary criticism. As the development of symbolic methods in mathematics was to show, and especially the slow recognition of non-EUCLIDEAN geometries as admissible mathematical structures, the view that mathematics be tied always to existing perceived reality could become a block to conceptual expansion. (It is irrelevant that non-EUCLIDEAN structures were to be admitted into physical explanations at the close of the 19th century. During the period in which non-EUCLIDEAN concepts were rejected from mathematics on the basis that the parallel postulate was "self-evident" and necessary, perceived reality was accepted axiomatically as EUCLIDEAN.)

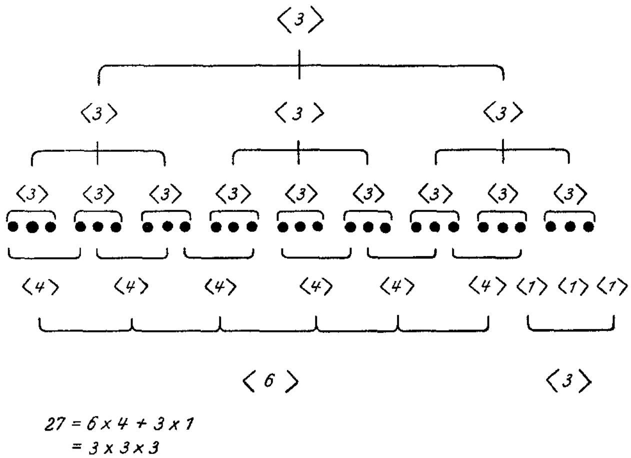
[6] For example, in RAPHSON's *SR*: cap. 3: 37—53: *de infinito abstracte considerato*.

[7] See A. KOYRÉ: *From the closed world to the infinite universe*, New York, 1957: *passim*: and MAX JAMMER: *Concepts of space*: Harvard, 1954: ch. 4: *The concept of absolute space*.

[8] Especially in his *LG*.

[9] *MU*: chs. 10ff.

the individual integers by $n < n+1$ (and allowing extension to negative integers by: $x = -a$ for $x+a=0$) and giving them conventional names, we can arrange them arbitrarily into sets, and then use the (ordered) integers to count these sets. So[10] WALLIS divides 27 units, numbered '1' to '27' into 3 sets each of 3 sets of 3 units, and again into 6 sets of 4 units with 3 units over. Clearly, definitions of addition and multiplication are immediate (together with their inverses, subtraction and division): When a set with λ units, $\langle\lambda\rangle$, is added to a second set with μ of the same units $\langle\mu\rangle$, the resulting set will have $\lambda+\mu$ of those units, and we denote it by $\langle\lambda+\mu\rangle$; and similarly we can divide some set of $\lambda\mu$ units into μ sets each of λ units (or λ sets each of μ units) or



$$\langle\lambda\mu\rangle = \langle\mu\rangle\times\langle\lambda\rangle$$
$$= \langle\lambda\rangle\times\langle\mu\rangle.$$

Denoting the set $\langle\lambda\rangle$ by λ, we can then codify the four admissible operations of arithmetic in the following rules, assumed if not stated explicitly in one form or another by all 17th century mathematicians:

$$\lambda+\mu = \mu+\lambda, \qquad \lambda\times\mu = \mu\times\lambda,$$
$$\lambda+(\mu+\nu) = (\lambda+\mu)+\nu, \qquad \lambda\times(\mu\times\nu) = (\lambda\times\mu)\times\nu,$$
$$\lambda\times(\mu+\nu) = \lambda\times\mu+\lambda\times\nu, \qquad [(\mu\times\nu)^\lambda = \mu^\lambda\times\nu^\lambda]$$

(where $\mu^\lambda = \mu\times\mu\times\mu\times\cdots\times\mu$, λ times). Further, there are three special integers, 0, 1, ∞ which satisfy these operation rules in an exceptional way:

$$\lambda+0 = \lambda, \qquad \lambda\times 1 = \lambda,$$
$$[\lambda+(-\infty) = -\infty], \qquad \lambda\times 0 = 0,$$
$$\lambda+\infty = \infty, \qquad \lambda\times\infty = \infty.$$

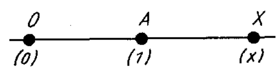
It is in these three integers that all the difficulty of the concept of integer lies. With the modern strict distinction between a set and its members, it is perhaps difficult to feel the confusion which arises when the distinction is not made. An element which does not exist cannot be used as a unit to count off the members of a set, and yet the null set $\langle 0\rangle$ is, in modern treatments, used to count off the members of a set. WALLIS in his introductory text on mathematics[11] takes up this point, and discourses at length on the difference between no quantity ("nullum") and the property of being no quantity, of being a member of the zero class ("nullitas"). Similarly, a careful distinction between one quantity ("unum",

[10] *MU:* ch. 11.
[11] *MU: ibid.*

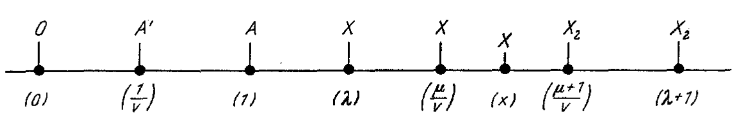
an element) and the property of being unity ("unitas") is drawn and used to resolve the medieval antimony (ARISTOTELIAN in origin) which argues that unity cannot be divisible when to divide it increases its number, which is absurd—an argument which confuses a number-element and a number-set of unity. Such difficulties are largely the result of verbal muddle, and, in the absence of a symbolic notation which can clarify them in an obvious way, number-mystical concepts of a type popular in logical texts of high scholasticism are easy to introduce but difficult to refute convincingly by verbal argument.

It is, however, significant that the integer was largely accepted in the period as a self-evident quantity whose importance lay in its being useful in computation and numerical mathematics generally, and WALLIS' detailed discussion is quite untypical*. The professional mathematician especially, in comparison with the rich and abundant consequences which he could draw from the concept, saw inquiry into its basis and the distinction between an element and a class as being, if not reprehensible on an intellectual level, as trivial as logic-chopping.

The case was different with the notion of a general real number and the theory of proportion built on it, which were widely seen for the genuinely subtle concepts they are. Unlike the integer, dealt with simply by defining it operationally, the standard 17th century introduction to the general real number was through the geometrical model of an (infinite) line-segment[12] though this is sometimes lightly disguised as a continuum of time ("duratio"). The fundamental idea is



that we can take a fix-point O on the given line-integral, the distance OA from which to a second fix-point A on the line is taken as a unit to measure the distance (x) from O to any third point X of the line. When OX is an exact multiple of OA, (x) will be an integer, and using this as a basis—in particular, the fact that the integers are naturally ordered by $-\infty < \cdots < -1 < 0 < 1 < \cdots < +\infty$—we can set up an equivalent order of a denumerable number of points X [where (x) is integral] of the line. Immediately, the way to order all points X of the line is suggested by the geometrical concept of "betweenness", and thereby the class of integers is seen as part of ("embedded in") the class of reals—or, on the model, line-segments OX which are of integral length (x) are part of the whole collection



of line-segments OX, (x) real. Specifically, a general segment OX of length (x) is such that x is defined uniquely by being between integers λ and $\lambda+1$: $\lambda < x < \lambda+1$, or, in the model, $X_1 \equiv (\lambda) < X < X_2 \equiv (\lambda+1)$. Further division of the unit-interval OA into ν parts (each of length $(1/\nu)$) allows a narrower inequality $\frac{\mu}{\nu} \leqq x \leqq \frac{\mu+1}{\nu}$ and on the model $X_1' \equiv \left(\frac{\mu}{\nu}\right) \leqq X \leqq X_2' \equiv \left(\frac{\mu+1}{\nu}\right)$ (where μ is the unique

* So BARROW in his *lectiones mathematicae*, probably a Cambridge equivalent to WALLIS' introductory lectures, skims lightly over the concept.

[12] Compare WALLIS *MU*: ch. 14; BARROW *LG*: lectio 1.

integer such that $\lambda\nu \leqq \mu, \mu+1 \leqq (\lambda+1)\,\nu$ which has $X_1' \leqq X \leqq X_2'$). Finally by choosing a sufficiently narrow measuring-interval $OA' \equiv (1/\nu)$, we can find points on the line which approximate to X with any desired accuracy but which still satisfy the ordering (inequality) $X_1' \leqq X \leqq X_2'$. (When, for some integer ν, $x = \mu/\nu$, we have a general rational point X.) It was, of course, an achievement of early Greek mathematics to show that—assuming the constructions of EUCLIDEAN geometry, and especially "PYTHAGORAS'" theorem on the sides of right triangles which defines its metric—points on the line exist which cannot be measured in the ordering by any length (μ/ν), μ, ν integers*, and the EUDOXIAN formulation which, as given in EUCLID Bk. 5, overcame the difficulty[13] was accepted in 17th century treatments as standard. (BARROW thinks its subtlety great enough to devote almost the whole of part 3 of *LM* to its explication[14] and on the Continent ARNAULD in his *Elémens*[15] discussed it at equal length, if less thoroughly.) On the geometrical model the two complementary forms of the EUDOXIAN definition** of real number seem more heuristically plausible than in an abstract symbolism, and it is in this way that BARROW introduces them in his *lectiones mathematicae*. In this formulation

$$\alpha = \beta \quad \text{if} \quad (m, n)\left(\alpha \gtreqless \frac{m}{n} \cdot \equiv \cdot \beta \gtreqless \frac{n}{m}\right);$$

and

$$\alpha > \beta \quad \text{if} \quad \begin{cases} (E m', n')\left(\alpha > \dfrac{n'}{m'} \geqq \beta\right), & \text{or} \\ (E m'', n'')\left(\alpha \geqq \dfrac{n''}{m''} > \beta\right); & \end{cases}$$

which on the model straightforwardly expresses the coincidence, or separateness, of the points (α), (β): the points (α), (β) are separate, or otherwise, according as we can, or cannot, find a third point (n/m) which lies between them, and if we can find such a point then, say $(\alpha) > (n/m) > (\beta)$, this defines $(\alpha) > (\beta)$ and conversely. The EUDOXIAN definition can then be used, as in the *Elements*, to prove all the

* Specifically, EUCLID Bk *10* shows that, for n a non-square integer, $(\sqrt{n})$ is such a point.

** Two reals $\alpha \equiv a/b$, $\beta \equiv c/d$ are *equal* if for all integers m, n, $ma = nb$ if and only if $mc = nd$; and *unequal* if we can find integers m', n' such that $m'a \geqq m'b$ while $m'c < n'd$ (or, equivalently, we can find integers m'', n'' such that $m''a > n''b$ while $m''c \leqq n''d$).

[13] Modern research suggests that the discovery of such "irrationales"—numbers which cannot be defined as the ratio between two integers and so cannot have any ratio at all in the Greek sense—occasioned a crisis in the 5th century BC, and that a first inadequate way out of the difficulty was by a continued-fraction approximation approach, later discarded when the improved EUDOXIAN definition was introduced. *Cf.* K. VON FRITZ: *The discovery of incommensurability by Hippasos of Metapontum.* Annals of Math. **48** (1945): 242—264; O. BECKER: *Eudoxon-Studien,* I—IV: Quellen und Studien zur Geschichte der Math. B **2** (1933): 311—333, 369—387; **3** (1936—): 370—388; and B.L. VAN DER WAERDEN: *Ontwakende Wetenschap (Science awakening):* Groningen, 1954: chs. 4, 5. Traces of the early continued-fraction theory have been found in Arabic commentators—*see* E.B. PLOOIJ: *Euclid's conception of ratio and his definition of proportional magnitudes as criticized by Arabian commentators,* Rotterdam, 1950.

[14] *LM* pt. 3 (1666): lectiones 3—8.

[15] ARNAULD, A.: *Nouveaux élémens de géométrie* ... Paris, 1667, 1683.

other properties of reals. So BARROW[16] gives the proof that, where α, β, γ, δ are reals, then $\alpha>\beta\cdot\equiv\cdot\gamma>\delta$ where $\alpha/\beta=\gamma/\delta$: for otherwise we could find integers m, n such that $m\alpha>n\beta$ with $m\gamma\leqq n\delta$, which defines $\alpha/\beta>\gamma/\delta$ ($\alpha>\beta, \gamma\leqq\delta$ implies that $\alpha/\beta>1\geqq\gamma/\delta$).

The most significant property of the real number is that it satisfies the operational scheme for integers[17] and the importance of placing it on a rigid basis is that the whole of analysis restricted to real functions can be developed—if not with advantage—by suitable definitions using that operational scheme as foundation*. Much of 17th century mathematical work was carried out, if not rigorously, very much in modern style, with suitable introduction of number bases (which implicitly contain a concept of successor function when systematic notation is used to denominate them) and even, as we shall see later, of simple functions. However, along with such analytical treatment, many developments were still made using the restricted but equivalent form of proportion theory, especially in geometry—a theory perhaps unjustly treated by recent writers[18].

Apparently the theory, like so many aspects of 17th century mathematics, Greek in origin[19], had developed as an offshoot of the concept of ratio (defined most generally between two reals). In particular by PYTHAGOREAN times two proportions (ἀναλογία) had been introduced to relate integers (and, by extension, reals) a, b, c, d, *viz*:

the arithmetic proportion (A) $(a, b; c, d)$ defined by

$$a-b=c-d,$$

and

the geometric proportion (G) $(a, b; c, d)$ defined by $a/b=c/d$.

Closely related are the three means:

arithmetic mean $(AM)\,(a, c)=b$ when $a-b=b-c$,

geometric mean $(GM)\,(a, c)=b$ when $\frac{a}{b}=\frac{b}{c}$

and

harmonic mean $(HM)\,(a, c)=b$ when $\frac{1}{a}-\frac{1}{b}=\frac{1}{b}-\frac{1}{c}$.

(It is an immediate consequence that $(AM)\times(HM)=(GM)^2$, or that (GM) is a geometric proportional between (AM) and (HM).) In later Greek mathematics other proportions[20] of theoretical rather than practical importance** had been

* And was so developed before and during the 17th century, apart from the small attention given to complex numbers in the theory of equations.

** It is provable that all such proportions can be defined in terms of (A) and (G): a fact which mirrors the two basic arithmetical operations of $\pm$, $\times/\div$.

16 *LM* (1666): lectio 8: 322.

17 See above.

18 For example C. B. BOYER *Proportion, equation, function: three steps in the development of a concept.* Scripta Mathematica **12** (1946): 5—13.

19 There is little accurate evidence, but the late Greek authority NICOMACHUS in his *Εἰσαγωγὴ ἀριθμητική (Introduction to Arithmetic)*: transl. by M.L. D'OOGE: New York, 1926: 151*ff.* and his commentator IAMBLICHUS (ed. PISTELLI: Leipzig, 1894: 103*ff.*) credit the PYTHAGOREANS with the arithmetic and geometric proportions.

20 MICHEL, P.-H.: *De Pythagore à Euclide.* Paris, 1950: pt. 2: ch. 1, III.

developed, but it was above all the geometrical proportion which, basic in the EUDOXIAN definition of reals, remained important in mathematics, and to a lesser degree also the arithmetic proportion. Over the centuries operations permissible in connection with it were codified, and by the end of the medieval period emphasis was placed on the especial importance of *the operation* ":", where $a/b=c/d$ (or, equivalently, $a:b=c:d$), then $b:a=d:c$ (invertendo), $a:c=b:d$ (permutendo) $(a+b):b=(c+d):d$ (componendo), $(a-b):b=(c-d):d$ (dividendo) and $a:(a-b)=c:(c-d)$ (convertendo). Using these operations and one final main theorem that, where $(GM)\,(a_1, a_2, \ldots, a_n)=(a_1\, a_2 \ldots a_n)^{1/n}$ and $(AM)\,(a_1, a_2, \ldots, a_n)=\frac{a_1+a_2+\cdots+a_n}{n}$ are generalized geometrical and arithmetical means, then $(GM) \leqq (AM)$, a powerful theory of inequalities can be built up which are the equivalent of corresponding inequalities in real number theory. So, for instance, there follows at once $(HM)\,(a_1, a_2)<(GM)\,(a_1, a_2)<(AM)\,(a_1, a_2)$, $a_1 \neq a_2$, an inequality extensively used in geometrical texts of the period. In general, a surprising number of important mathematical developments arose on the basis of proportion theory, and HUYGENS[21], JAMES GREGORY[22] and BARROW[23] made notable use of it in refining approximations to the length of the circle-arcs. The typical proportion proof has a delightful symmetry, and its elegance, no doubt, was one reason for its continued use. Further, there seems no reason why proportion theory could not be extended by the introduction of suitable definitions to cover most of the ground treated in classical mathematics by free-variabled polynomials, though admittedly the extension would be unwieldy. It is, however, important to notice that the proportion theory was superseded not as being theoretically inadequate but as cumbrous at a practical level. In comparison with the computational facility of polynomial theory (which lent itself to computations with the decimal-base Hindu-Arabic numerals) treatment by proportions seemed relatively difficult and not worth the time needed to learn its manipulations. Its arguments are indeed tricky*, and it is significant that BARROW in his edition of ARCHIMEDES (written perhaps about 1665) rewrites the ratio theory proof forms of the original Greek text in the free variable notation which was passing into accepted use, and indeed, when faced at one point in the text with a particularly involved form, cannot believe it the way of ARCHIMEDES' original discovery and supposes the method of analysis used much nearer to the modern

* In particular, the inadequacies of verbal treatment made the distinction between a multiple of a ratio $(\lambda \times (a/b))$ and the corresponding power $((a/b)^{\lambda})$ very tricky. Many medieval texts fall into the error of confusing the two[24]—an error repeated in the 17th century in the *opus geometricum* of GREGORY ST. VINCENT[25], an immensely detailed work which had as its main aim the proof of the impossibility of analytical quadrature of the circle.

[21] In *de circuli magnitudine inventa*. Leyden, 1654.

[22] A particularly fine example comparing a limit-sequence with the limit-sum of a geometrical progression is given *in extenso* in ch. 5 (taken from his *VCHQ*).

[23] Especially lectio 11: appendix of his *LG*.

[24] See, for example, Richard SWINESHEAD: *liber calculationum*. Venice, 1520: tract 11: *de loco elementi* (36rb—38va) *passim*.

[25] *opus geometricum quadraturae circuli et sectionum coni*. Antwerp, 1647, especially Bk. 11: prop. 53 (1132ff.); and the criticism by HUYGENS in *theoremata de quadratura hyperboles, ellipsis et circuli ex dato portionum gravitatis centro*. Leyden, 1651: app. *Ἐξέτασις cyclometriae ... Gregorii à Sancto Vincentio ...* = *HO* **11**: 315—337.

form*. Yet the theory on any evaluation was more than the minor branch of elementary mathematics which it has today become.

Before, however, the theory fell into disuse many mathematicians were beginning to realize the close analogy, pointed by proportion theory, which exists between the operations of $\pm$ and $\stackrel{\times}{\div}$, and which jumps to the eye when we set down standard results in parallel columns:

$$\begin{array}{ll} \lambda+\mu=\mu+\lambda, & L\times M=M\times L, \\ \lambda(\mu+\nu)=\lambda\mu+\lambda\nu, & (M\times N)^L=M^L\times N^L, \\ \lambda+0=\lambda, & L\times 1=L, \\ (AM)\,(\lambda,\mu)=\tfrac{1}{2}(\lambda+\mu), & (GM)\,(L,M)=(L\times M)^{\frac{1}{2}}, \\ (A)\,(\lambda,\mu{:}\nu,0)\cdot\!\equiv\!\cdot\,\lambda-\mu=\nu-0, & (G)\,(L,M){:}N,1)\cdot\!\equiv\!\cdot\,L\div M=N\div 1, \end{array}$$

to which we can add the arithmetical and geometrical progressions[27]

$$\begin{array}{ll} (AP)\,(\lambda,\mu{:}\varkappa)\cdot\!\equiv\!\cdot\,\lambda+\varkappa\times\mu, & (GP)\,(L,M{:}K)\cdot\!\equiv\!\cdot\,L\times M^K, \\ \varkappa=0,1,2,\ldots. & K=0,1,2,\ldots. \end{array}$$

Thus, a result on the left side becomes a corresponding theorem on the right where the operations $\pm$, $\stackrel{\times}{\div}$ pass into $\stackrel{\times}{\div}$ and power exponents, and $\lambda, \mu, \nu, 0$ become $L, M, N, 1$. We recognize the mapping as logarithmic—where $\lambda=\log(L)$, $\mu=\log(M)$, $\lambda+\mu=\log(L\times M)$, maps into $L\times M$—and isomorphic**, but we do not have to know the precise nature of the correspondence to feel the similarity of pattern and a full realization of its existence is everywhere in the period. So it was by analyzing the conditions under which $(\lambda+\mu)\leftrightarrow(L\times M)$ that NAPIER set up his canon of logarithms[28], but that was only a beginning. We find a little later that the correspondence is used virtually to set up dual theories (which are isomorphic by the mapping), one of which is considered in detail while the other is merely sketched in. As LEIBNIZ, on a theoretical level, puzzled over the similarity of the two proportion-concepts, arithmetical and geometrical $((A)\leftrightarrow(G))$[29], JAMES GREGORY gave many of the propositions of his *VCHQ* in dual form[30] and MENGOLI in his *geometria speciosa* used the uniqueness of the isomorphism to develop a rigorous basis for the logarithm on the model of the EUDOXIAN

* "... This is the exact equivalent of the proportion deduced by ARCHIMEDES (and, to insert a general remark, it reveals sufficiently the sort of analysis he used; for that he arrived at the result through application of those various compositions, divisions, alternations and inversions he produces is almost beyond belief: and, if he did so, it must be supposed by chance rather than by any design that he came on the true solution, and that this happened time after time can scarcely be believed)."[26]

** In fact, between the interval $[-\infty, +\infty]$ and $[0, \infty]$.

[26] *Archimedis opera* ...: 33; commenting on ARCHIMEDES: *Sphere and cylinder:* Bk. 2: prop. 5.

[27] See ch. 5.

[28] See ch. 3.

[29] Compare KARL BOPP: *Drei Untersuchungen zur Geschichte der Mathematik.* Schriften der Straßburger Wiss. Gesellschaft in Heidelberg, No. 10. Berlin and Leipzig, 1929: 2 (5—18): *Leibniz, Arnauld und de Nonancourt*, especially 11 *ff.*

[30] For example, prop. 21 $\leftrightarrow$ prop. 22, prop. 24 $\leftrightarrow$ prop. 25 under the mapping.

definitions of equality and inequality for reals[31]. But perhaps most important of all are the dual forms in which exhaustion proofs can be given, likewise isomorphic (and used implicitly by ARCHIMEDES himself in his various works[32]).

Clearly, the way was open for a general viewpoint on algebraic structure, but especially on isomorphic invariance. That it did not happen has no simple explanation—partly, perhaps, the resistance of accepted ideas is to be blamed, but it seems a more important hindrance was the sudden outpouring in the latter half of the 17th century of a mass of numerical formulae and infinite sequences which tended to draw the attention (and creative effort) of the few mathematicians of sufficient maturity to build such a theory of abstract mathematical structure.

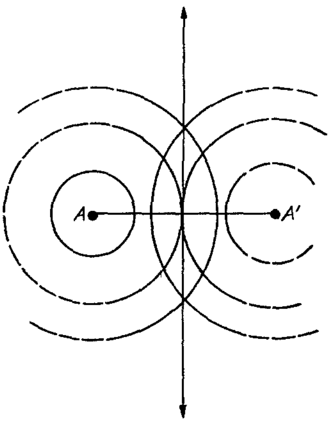


Fig. 1

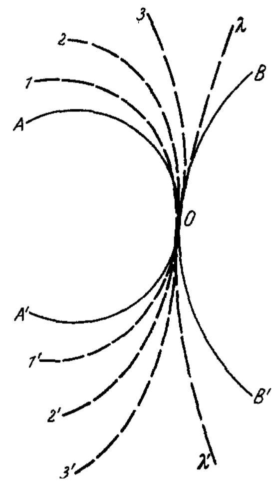


Fig. 2

The comment is general. The 17th century had bequeathed to it, especially from Greek sources, a very rich collection of valuable remarks on points in mathematics which it very willingly repeated but developed little. So, for example, the concept of continuity was still universally treated as an unanalyzable concept, to be expounded ostensively in some suitable model, or to be elaborated metaphorically. Thus NICOLAUS MERCATOR, in his introduction to the (anonymous) elementary geometry text *Euclidis elementa geometrica*[33] conceives the image of a stone dropped into a still pond, with ripples spreading out from the impact point in ever-widening circles, to introduce a real-number measure into geometry. Each point on a generating circle will, by its motion, traverse a continuous line-interval, and the set of concentric circles will cover the plane of the pond's surface in a (polar) coordinate system. Further, two stones dropped into the pond simultaneously will generate two separate concentric circle sets, corresponding members of which meet in two points which will each generate a half-line (from

[31] *geometria speciosa*, Bologna, 1659: especially bks. 4, 5.
[32] See ch. 9.
[33] *Euclidis elementa geometrica, novo ordine ac methodo fere demonstrata, una cum introductione brevi qua magnitudinum ortus ex genuinis principiis et ortarum affectiones ex ipsa genesi derivantur.* Londini, 1678: 16.

O in the diagram). The example is repeated (though not exactly) by RAPHSON[34], but neither attempts to abstract any general principles, relying exclusively on an intuitive concept of continuous variation (Fig. 1).

Nor is there any real advance in—what might seem the exception—WALLIS' treatment of the problem of the horn angle*[35], where he sketches in a treatment using ARCHIMEDES' lemma[36]: for $a < b$ there is some n (finite) with $na \geqq b$ (a, b homogeneous magnitudes), and which therefore gives, in effect, the necessary restriction for rejection that the angle measure be a real-number continuum. Indeed, this recourse to ARCHIMEDES' lemma is only the second of six arguments, the last of which, where he argues that to admit the horn angle would be to deny the optical properties of conics, is an incomprehending *petitio principii*. In fact, allowing n-order differentials of the curves, it is possible consistently to define λ-sections of the horn angle AOB which, measured conventionally as the rectilinear angle between the two (coincident) tangents to the curves AOA', BOB' at O, is indefinitely small (and so zero in the limit). WALLIS, basing so much of his argument on an uncritical appeal to experience, would never allow as meaningful the concept of λ-secting an angle of zero magnitude; and even in applying ARCHIMEDES' lemma introduces it on the same, unmodified viewpoint of an angle as generated by the continuous motion of a line around a fix-point, attaching a unique number out of the interval $[0, 2\pi]$—or $[0, \infty]$, allowing the concept of periodicity. Apparently he does not realize that ARCHIMEDES' axiom is a postulate to be denied at will (Fig. 2).

This same lack of rigour in basic definitions is probably a root reason why such general concepts as function had still to be treated abstractly from geometrica models at the end of the century, when a wealth of particular functions[37] had been found. So JAMES GREGORY in *VCHQ* had tried to apply DESCARTES' concept of an analytical construction[38] to the quadrature of a conic segment, seeking to show that such quadrature is impossible if we restrict ourselves to sequences starting from areas of rational-measure. Since any area of rational-measure is definable by (an infinity of) sequences of analytical operations from any other area of rational-measure, it suffices to show the impossibility of quadrature in the case of a single sequence of analytical operations performed on any given areas of rational-measure. It was easy enough for GREGORY[39] to

* In general, the angle between two (continuous) curves at a point where they share a common tangent.

[34] *de spatio reali:* 44ff. The concept of a concentric circle generator-system is, of course, Greek—compare PROCLUS' commentary on EUCLID Bk. 1 (French transl. by P. VER EECKE, Bruges, 1948), *passim*.

[35] In his *de angulo contactus et semicirculi tractatus*, printed in *operum mathematicorum pars altera*, Oxford 1656 and republished with a defence in appendix to his *Algebra*, 1685 ·≡· opera **2** (1693): 605—630; 631—634 respectively.

[36] DIJKSTERHUIS: *Archimedes:* 146*ff.*, see lambanomena 5 of *Sphere and cylinder* **1**.

[37] Especially those defined by limit-sequences: see ch. 5 *passim*.

[38] That is, the construction of a number to be obtained from given numbers by any combination of $\pm$, $\times$/$\div$ and root-extraction—*cf.* Descartes: *Géométrie:* ·≡· *Discours* ... app. p. 237.

[39] See ch. 5. GREGORY's construction of the sequence he gives in *VCHQ* is a generalization of theorems known widely in the 16th century—see TROPFKE$_2$ **4** (1923): 218—222.

define such a sequence starting with a circumscribed or inscribed polygon whose limit is a general sector of a conic, and he seems to have thought that the fact that the sequence was infinite was sufficient to show the non-analytical nature of the sequence[40]. The inadequacy of the reasoning becomes clear when we have a firm control over limit-processes, and rigorous proofs of non-analytical quadrature had to wait for stricter formulation of an analytical process in terms of the zeros of the general algebraic polynomial. (Two converging sequences which bound a given number (within an estimable error-range) are sufficient to evaluate the number to any required approximation by ratios, but we cannot, without further precision of the converging sequences, assume that an infinite sequence defines an irrational number, and certainly not a transcendental one, as GREGORY suggests.)

One final aspect of 17th century mathematics is of a general importance—the power which an adequate notation gives of emphasising and crystallizing thought-patterns in a significant way. It is a common, but none the less important, remark that general calculus forms are not found historically till usable notations were developed to express their intricate concepts, and that the general symbolic treatments beginning in the 18th century (and synthesized in the 19th century in such concepts as the CAUCHY-RIEMANN limit-sum integral) were dependent on simplification and generalization of 17th century techniques; and the point is true in general that notational improvement and conceptual mathematical advance are concomitant.

Not all symbolisms were of course significant in that they gave new insight into existing concepts[41], for they were frequently introduced to make for easier comprehension by simplifying the visual layout. This was, indeed, the explicit reason given for their introduction in the 17th century, one which receives concrete expression in BARROW'S compressed and cleared-up university texts of EUCLID (1674), and ARCHIMEDES, APOLLONIUS and THEODOSIUS (1675), typifying a general movement which sought to substitute simplified, more adequate notation for the clumsy verbal—and heuristically implausible—Greek treatment. But implicitly and at a deeper level such notational introductions often fixed concepts on the outer borders of existing knowledge. Thus, the theory of continued fractions—for instance, recursive definition of convergents—developed with the notation which formed it[42]; and the convergent analytical sequence given by JAMES GREGORY[43] for deriving approximations to the area and arc-length of

40 It is, of course, a necessary but insufficient condition. BARROW, in *LM* (1666) **1**: 175, falls into a similar but opposite error, arguing that, since it is possible to set up an ARCHIMEDEAN lines a sequence of circumscribing and inscribing regular polygons S_n, s_n whose common limit is a general circle arc, then the general circle arc must be rational when S_0, s_0 (and so all S_i, s_i) are.

41 For example, the simple transition from the ratio-forms of WALLIS' $q:d$ and BARROW'S $\pi:\delta$ into the modern constant π.

42 Convergents to unit continued fractions were first defined by DANIEL SCHWENTER in his *geometriae practicae novae et auctae tractatus*, Nuremberg, 1618: **1**: 58—59, and developed in his *deliciae physico-mathematicae*, Nuremberg, 1636: 111 ff. and are given for the general continued fraction, apparently derived by numerical induction from the observed pattern of the first few convergents, by WALLIS in *AI*: prop. 191: scholium.

43 In *VCHQ* and in extended form, in *EG*, *passim*.

the circle and rectangular hyperbola depended in large part on the ambiguous matrix form in which it is clothed. Again, many examples exist in the period which reveal how notational lack could prove a block to advance. BRIGGS surely failed to give the general binomial expansion[44] because he used the inadequate BOMBELLI ring notation for free variable (which had the triple function of distinguishing powers of the same variable, different variables and place-value in decimal expansions); while the lack of a symbol for the cross-ratio of four points (together with the fact that the harmonic case, -1, was treated separately in view of its importance in conic theory)[45] retarded development of a separate projective treatment of geometry till the 19th century; and WALLIS' inability to see the term by term equivalence of the inverse of the BROUNCKER continued fraction for $\square(=4/\pi)$ with LEIBNIZ' limit sum-sequence for $\pi/4$ is a failure to apply the recursively-given general convergents to a continued fraction developed in his *AI*[46], one which reflects the inadequacy and complexity of his notation.

In summary, we can say that basic concepts were not investigated in the 17th century with any insight, but that an adequate basis for mathematics, accepted as a matter of practice, did exist which was little different, if at all, from that explored in Greek and medieval times. The 17th century is, in mathematics, a period of rapid advance using valid but tenuously defined concepts as a basis for a rich and varied technical achievement. The greatness of that achievement is to be evaluated by a detailed study not of what 17th century mathematicians thought but of the evolving pattern of what they did, and to that end the rest of this essay is an attempt to isolate significant trends in that achievement.

II. Universal arithmetick and specious algebra

Throughout the 17th century algebraic studies were largely restricted to their traditional field of the theory of equations.[1] In particular much attention was still given to the cubic and quartic equations for which general algebraic reductions had been given in the 16th century, and a quite disproportionate amount of time was spent in developing geometrical constructions for their real roots as the cut-points of two conics.[2] What from the conceptual viewpoint is significant in all this is not the detail of the techniques evolved to deal with par-

---

44 See ch. 4.

45 See ch. 6.

46 See note 42, and compare WALLIS' letter to COLLINS, 16 September 1676. Rigaud **2**: 598–600. Specifically, where $\Phi_i = 1 + \frac{1^2}{2+}\,\frac{3^2}{2+}\cdots+\frac{(2i-1)^2}{2}$, $X_i = \sum_{1\leq j\leq i}\left[(-1)^{j-1}\times\frac{1}{2j-1}\right]$, then $X_i = \frac{1}{\Phi_i}$ for each i, and their common limit as i becomes infinite is $1/\Phi$ ($=1/\square$ in WALLIS' notation) $= X = \frac{1}{4}\pi$.

1 The word "algebra" derives etymologically from AL-HWARIZMI's 9th century treatise *hisab aljabr w'almuqabalah (de restauratione et de appositione)*, on restoration and reduction of equations.

2 The definitive treatment of this was given by PHILIPPE DE LA HIRE in *La construction des équations analytiques*, part **3**: 297–452 of his *Nouveaux élémens* ... Paris, 1679. LA HIRE showed that the real roots of any cubic or quartic could always be found by the meets of a circle and a parabola. NEWTON, in an appendix (written probably about the same time) to *AU*, likewise devotes much space to the subject.

ticular equations[3] but the general methods which were introduced both to define the equation and to solve it. Above all we owe to this elaboration of equation theory the general real (and complex) variable.

The development of the concept of variable is very closely tied up with the notation used to express it and the slow progress towards an adequate symbolism is mirrored in the prolonged difficulties over free and bound variable forms. It was on this basis that NESSELMANN[4] tentatively established a division of algebra into rhetorical (where the proof is purely verbal and non-symbolic), syncopated (where systematic abbreviation of the verbal forms occurs) and fully symbolic, operational forms. But the variable is something more than its mere symbolic denotation and NESSELMANN'S classification is perhaps a little narrow and rigid, and certainly arbitrary. Logically it seems natural to classify a variable by its range, a basis widely adopted in the 17th century in the view that algebra is a "universal" arithmetic, a systematisation not only of equation theory but of all arithmetical equations—as COLIN MACLAURIN was to state it[5]: "Algebra is a general method of computation by certain signs and symbols, which has been contrived for this purpose and found convenient. It is called an Universal Arithmetick, and proceeds by operations and rules similar to those in Common Arithmetick, founded upon the same principles." In short, algebra was defined as the generalisation of numerical arithmetic which retains the basic operations of $\pm$, $\stackrel{\times}{\div}$ and has variables ranging over the interval $[-\infty, +\infty]$* such that when numerical values are substituted for the variables (consistently), there results a theorem of arithmetic.

This viewpoint crystallises centuries of developing ideas—the concept of substitution-variable is as old as Diophantus and is found widely in the works of the medieval "calculators",—the final generalisation from substitution-variables to fully free variables, which we can connect suitably one to another and so use to define a general structure, came only with the general systematisation of equation theory which began with VIETA[6] and BOMBELLI[7]. To VIETA is due the first distinction between the single substitution-variable and the general free variable when he differentiates between numerical algebra, a mere series of substitution-instances compacted in a formula, and specious algebra, where we use the limitations of a defined algebraic structure to derive a "canon", a method of deriving particular solutions. Possibly VIETA himself would have

* Though for a long time the variable, say x, was allowed only to range over the positive interval $[0, \infty]$, and $x \in [-\infty, 0]$ was introduced by defining $x = -y$, y positive and ranging over $[0, \infty]$. This has been seized upon as a significant point, but in fact is easily held in mind and would be troublesome at an early stage only.

[3] For a good summary of these techniques see TROPFKE$_3$ **3** (1937): B: *Die Gleichungen:* 22—235.

[4] Compare G. H. F. NESSELMANN: *Versuch einer kritischen Geschichte der Algebra.* 1: *Die Algebra der Griechen.* Berlin, 1842: 301—306.

[5] COLIN MACLAURIN: *Treatise of algebra.* London 1748: part 1, ch. 1, 1—2.

[6] VIETA, in fact, developed the first adequate, usable symbolism for the free variable in a series of works beginning with *canon mathematicus, seu triangularis, cum adpendicibus* ... Paris, 1579. Compare FRÉDÉRIC RITTER: *François Viète, inventeur de l'algèbre moderne: Notice sur sa vie et son œuvre.* Paris, 1875.

[7] In his manuscript *Algebra,* printed in entirety for the first time in E. BORTOLOTTI: *L'algebra opera di Rafael Bombelli di Bologna,* Bologna, 1929, BOMBELLI systematised the whole of the 16th century Italian algebraical achievement.

wished to restrict specious algebra expressly to the techniques which examine the zeros of the polynomial $\Phi(x)$—the classical problem of equation theory, in short—but specious algebra soon became identified with universal arithmetick, and together these were seen as defining a general "analytical" approach to mathematics. So NEWTON, introducing his Lucasian lectures in the 1670's,[8] writes that "Computation is carried out either by pure numbers, as in common arithmetic, or by variables ("species"), as is the habit of the analyst."

The 17th century mathematicians themselves saw the great triumph of this analytical method in its applications to geometry and the general treatment of such traditional concepts as "curve" (defined from the time of DESCARTES' *Géométrie* as a point-set limited by some "relatio" which exists between co-ordinate line-lengths).[9] Consciousness of the new freedom afforded by universal algebra acted as an inspiration even where its method was not directly applicable and led to a widespread search for general treatments and a balancing dissatisfaction with particular cases—an attitude summed up by JAMES GREGORY in the preface to *GPU*: "It has been observed by geometers of our century that mathematics was ill divided by the ancients into geometry, arithmetic etc. ... and that a better division is into the universal and the particular. The universal part of mathematics treats of the common proportion which is to be abstracted from all species of quantity ...: the particular part of mathematics is divided into geometry, ... which is merely the universal part of mathematics restricted to the figuration (figura), into arithmetic, which is the same universal mathematics restricted to number, into statics, the same restricted to motion, and so on." (In the sequel, he sees the universal part of geometry as comprising in part the equivalence transforms, "transmutations", to which we can subject given geometrical configurations.)

An interesting objection, not untypical of the age, to allowing algebraic forms into mathematics was raised by BARROW[10]: "Perhaps someone will perchance marvel ... why I have not spoken of algebra or the analytical faculty ... Because to be sure analysis (understood as intimating something distinct from the propositions and rules of geometry and arithmetic) seems to belong to mathematics no more than to physics, ethics or any other science. For this is merely a part or species of logic, or a manner of using reason in the solution of questions and in the finding or proof of conclusions, and of a kind not rarely made use of in all other sciences. Therefore it is not a part or species but rather the servant of mathematics; and no more is synthesis, which is a manner of demonstrating theorems opposite and converse to analysis." Here, of course, BARROW is arguing for a rapidly dating, predominantly Greek viewpoint on mathematics, but his objection points the fact that, more than a notational or numerical advance, the introduction of free variable was a logical one: what is new is not that an adequate symbolism or a suitably widened range has been given to the variable, but that the logical restrictions on the variable, unexpressed notationally till the 19th century, can themselves delimit a mathematical structure. BARROW's remark is

[8] These lectures were, of course, printed as *AU*: compare preface, 1—2.
[9] See ch. 7.
[10] *LM* (1664): lectio 2: 31—32.

significant not in that he was unwilling to accept what seemed an extramathematical logic into mathematics—the trend of the age was wholly against it, and BARROW himself must have seen the question as purely academical—but that he realized that the basis of the new universal arithmetic lay in new concepts which were later to receive such names as quantification of the variable, dummy variable, tied variable, range of variance, domain of a function and functional form. One could, of course, use the new algebra without consciously being aware of the underlying subtleties, and in historical fact few 17th century mathematicians (on the whole, eminently practical and not prone to worry over logical niceties) had the minimal logical training necessary to appreciate them. DESCARTES, JAMES GREGORY and, to some extent, NEWTON had a feel of the logical basis together with BARROW, but it was LEIBNIZ with his years of study of classical, medieval and 17th century logical treatments who first began to consider the concept of function in the abstract.

Refinement in the concepts which were introduced roughly and readily in the 17th century was a slow process which lasted well into recent times. To the extent that it created an undue respect for particular results and formulas and that it needlessly obscured many generalisations, the slowness of recognition of the logical basis of algebra was a main conditioning factor in the sterility of much of 18th century mathematics—one could not hope for insight when MACLAURIN could approach the problem only by analogy:[11] "In geometry the representations are more natural, in algebra more arbitrary. The former are like the first attempts towards the expression of objects, which was by drawing their resemblances; the latter correspond more to the present use of language and writing." Yet a solid, usable logical basis exists explicitly in 17th century algebraic studies, however little understood, and is to be appreciated rather through detailed examination of particular techniques evolved. For that reason, in the remainder of this chapter—though the strictly algebraic are not to be separated from related geometrical approaches—certain aspects will be considered of interesting applications of free and bound variable forms which, in abstraction from particular contexts, can be shown to illuminate each other.

As we have said, much of the mathematical effort in the period—and particularly the new analytical study of geometrical concepts—was still reducible in one way or another to the derivation and solution of an equation between variables. So were solved many of the problems of astronomy and of applied mathematics in general, though in many cases the reduction was not immediately obvious. For example, WREN proposed and solved[12] a problem which had originally suggested itself in finding the distance of a comet's (supposed rectilinear) path from the earth: Given four coplane lines BA, BF, CG, DH, to find a fifth $GHAF$ which cuts these such that the respective segments AF, AG, AH are in

---

[11] *A treatise of algebra*. London, 1748: ch. 1, 2:2.

[12] About 1661—see WALLIS: *opera* 2 (1693): 455—462. (Latin version of 1673 only) *Algebra:* cap. 105. A similar problem is treated by NEWTON in *AU:* Prob. 30: *cometae in linea recta uniformiter progredientis positionem cursus ex tribus observationibus determinare*, while both are akin to APOLLONIUS' studies in *de sectione rationis* (ed. HALLEY), Oxford, 1704.

a given ratio, say $1:m:n$. Taking $BC=c$, $BD=d$; $AE=x$, $EF=y$ and $BE:EF=f:1$, $CM:MG=g:1$, $DN:NH=h:1$ (which defines the lines BF, CG, DH in fixed position with regard to BA), we have where GM, HN, FE are perpendicular to BA, $GM=my$, $HN=ny$, $CM=gmy$, $BE=fy$, $DN=hny$, $AM=mx$ and $AN=n$; or $EM=x(m+1)$, $EN=x(n+1)$, $CE=c+fy$, $DE=d+fy$, so that

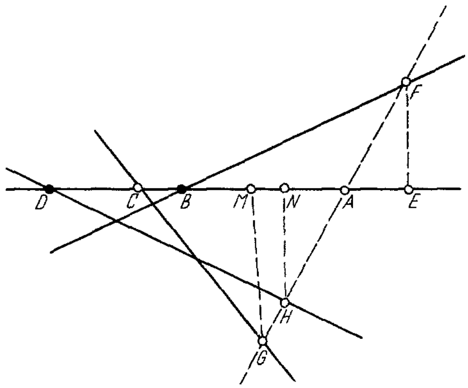


Fig. 3

$$EM(=CE-CM)=x(m+1),$$
$$=c+y(f-gm)$$
and
$$EN(=DE-DN)=x(n+1),$$
$$=d+y(f-hn).$$

We have then two simultaneous linear equations in x and y, and standard reduction gives a solution.

Such problems, some medieval in origin, are to be found in large number in all the algebras of the period[13], but more important was the growing consciousness of the values of "indeterminate" equations—general polynomial forms in one or more variables. The concept is basic to analytical geometry in that a polynomial form can, when a suitable coordinate system is defined, be seen as a model of the point-set of an (algebraic) curve, but the application was made when already the polynomial form had been developed in equation theory as an independent general algebraic structure, and especially in the special case of a single variable $\Phi(x)=\sum_{0\leq i\leq n}(a_i x^i)$.

Above all, through its origin in theory of equations, a great emphasis had been put on finding the zeros of a polynomial, on isolating a root and if possible finding its value. On that basis and particularly in 16th century Italy[14] there had grown a proliferation of results for removed in many cases from practical application, incorporating general methods of reduction and the synthesizing of standard procedure for whole classes of polynomials. In particular, it had become accepted that a linear equation always has a real root (which may be non-positive and so unacceptable on a particular view of mathematical reality); that a quadratic may have two real roots or none (in which case we can, if we

[13] Compare G. KINCKHUYSEN: *Algebra oste stelkonst*, Harlem 1661; R.H. RAHN: "Teutsche Algebra", Zurich, 1659 (which had a popular English translation by T. BRANKER, London 1668); J. PELL, who published little himself but whose pupils BRANKER, RHONIUS, LITTLEBURY and others printed many of his problems; but above all J. KERSEY: *The elements of ... Algebra ...* London, 1673, and J. WALLIS: *A treatise of Algebra both historical and practical ...* London, 1685, with many additions in the Latin translation of *opera* 2 (1693): 1—482.

[14] Though, of course, the use of equations in solving problems is at least as old as the Babylonians of the third millennium B.C., and many standard results on linear and quadratic equations had been formulated in Greek times (and independently in India, China and Japan before Western ideas penetrated there). Further particular examples of higher polynomials had been treated in Arabic texts—for instance, the solution of the cubic by intersecting conics—and medieval mathematicians such as FIBONACCI had developed successful numerical techniques. See TROPFKE *op. cit.* (note 3).

wish, extend the range of the root and allow two (conjugate complex) ones so that the quadratic has always two roots); that a cubic may have three real roots or only one (or always three if we allow the possibility of a conjugate complex pair); and similarly for quartics (which may have, likewise, four, two or no roots or always four). The big block to extending polynomial concepts beyond the quartic had been that no standard reduction of root-isolation techniques to those of lower-degree polynomials had been found (and, of course, none is possible: quintic and higher polynomials are, in general, irreducible). A second hindrance to general treatment lay in the conventional practice of distributing particular polynomial forms on either side of the equation so that each coefficient is positive, which confuses the suggestive denotation of a polynomial as a finite sum-sequence, $\sum_{0 \leqq i \leqq n} (a_i x^i)$, ordered by powers of the variable, x—a concept further distorted by those proportion-theory treatments which found it convenient to set the zero of the polynomial $\sum_{0 \leqq i \leqq n} (a_i x^i) = 0$, in proportion-form as $\lambda(x) : \mu(x) = \nu(x) : o(x)$, where λ, μ, ν, o are polynomials such that $\lambda \times o - \mu \times \nu = K \times \sum_{0 \leqq i \leqq n} (a_i x^i)$.[15]

However, particularly through the influence of VIETA, the modern form of denotation had been more widely accepted by the early 17th century, and we find ideas on the general polynomial forthcoming in more rapid sequence. Particularly with the introduction of the curve point-set, we find the concept of a polynomial having a root which is enumerable, approximately if not exactly, being transformed into the concept of a polynomial form having a specifiable number of zeros (its roots) equal in number to its degree[16], of which the real zeros are represented on the geometrical model by the meet of the curve $y = \Phi(x) = \sum_{0 \leqq i \leqq n} a_i x^i$ with the right-line $y = 0$. In some ways the geometrical model offered no immediate guidance, and in particular seemed to suggest no way of isolating real and complex roots from abstract consideration of the polynomial form: but adequate techniques were quickly developed in DESCARTES' rule of signs[17], which gave upper bounds to the number of positive roots, and more spectacularly, in NEWTON's rule, given in his *AU*[18], which states upper bounds for the number of

[15] Compare C. B. BOYER: *Proportion, equation, function: three steps in the development of a concept.* Scripta Mathematica **12** (1946): 5–13.

[16] Significantly TROPFKE$_3$ **3**; 175 cites PETER ROTHE in his *arithmetica philosophica*, Nuremberg, 1609, as the first to state generally that the n^{th} degree polynomial can have up to n real roots, and ALBERT GIRARD in his *Invention nouvelle en l'algebre*, Amsterdam, 1629, as stating firmly that the n^{th}-degree polynomial has exactly n roots, real or complex.

[17] Given in outline in *Géométrie*, Bk. **3**: 373, but already in fact, stated in T. HARRIOT: *artis analyticae praxis*, London, 1631.

[18] *AU:* part 2, ch. 2: 241 ff.: *de forma aequationis.* NEWTON gives no proof, and this is, in fact, extremely difficult. Despite several attempts in the 18th century the first rigorous treatment was developed by J. J. SYLVESTER in the 19th century using complex analytical techniques—see J. J. SYLVESTER: *On an elementary proof and generalisation of ... Newton's hitherto undemonstrated rule for the discovery of imaginary roots,* Proc. London Math. Soc. **1** (1865): 1–16, =*Collected mathematical papers*, Cambridge: **2** (1904): 498–513; and compare H. W. TURNBULL: *The mathematical discoveries of Newton,* London, 1945: 49–51. The only way NEWTON could reasonably have found his rule with the techniques at his disposal would seem by a RAMANUJAN-type induction over numerical instances, or over the lower orders of polynomials (the lower genera of algebraic curves).

positive and negative roots (and so, complementarily, a lower bound for the number of complex ones) in a much tighter way than DESCARTES' rule. But on the whole geometrical curve and algebraic polynomial yielded a rich store when studied together.

So, arising naturally from the abstract study of the polynomial is the realization that the roots can be expressed as homogeneous functions of the roots*—an idea seemingly original with GIRARD (stated for the general polynomial)[19], but given independently by JAMES GREGORY[20] and NEWTON[21] who both use the fact to express the sum-powers of the roots, $\Sigma(a_i^\lambda)$ $(\lambda = 1, 2, 3, 4, \ldots)$, in terms of the polynomial's coefficients. Elsewhere in *AU*[22] NEWTON gives simple applications to geometrical problems, particularly to the problem of drawing conics through a specified number of fixed points to touch fixed lines. Perhaps most interesting, however, is an example which occurs in a draft of his *enumeratio*[23] and which appears to have been the basis for certain of the geometrical properties of cubics developed there.

Here NEWTON begins[24] by deriving several known results on conics from the 2-degree polynomial (general quadratic form), using the expansion of the coefficients in terms of the roots. Though no proofs are given, the approach is clear. Consider some conic defined by the five points O, O', P, P', A, where OA is parallel to PP', and let a third parallel YXY' be drawn meeting the conic in Y, Y' and OBO' in X. Taking abscissa $OX = x$ and (in general oblique) ordinates $\begin{matrix} XY = y_1 \\ XY' = -y_2 \end{matrix}$, suppose the representing equation of the conic to be $y^2 - y(ax+b) + (\lambda x^2 + \mu x + \nu) = 0$. Then, assuming a suitable sense to the lines, $X \to O$ has $x = 0$, $y = 0$ or $-OA$, or $y_1 y_2 = \nu = 0$, and $y_1 + y_2 = b$, $= -OA$; $X \to B$ (the meet of OO', PP') has $x = OB$, $y = BP$ or $-BP'$, or $y_1 y_2 = \lambda OB^2 + \mu OB$, $= -BP \cdot BP'$, with $y_1 + y_2 = a\,OB + b$, $= BP - BP'$; and finally

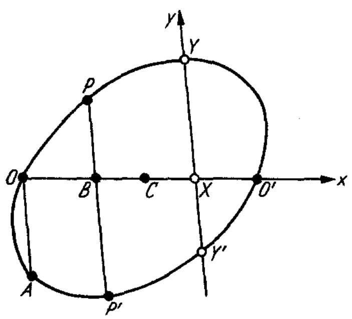


Fig. 4

* Briefly, where

$$\prod_{0\le i\le n} (x - a_i) \equiv \sum_{0\le i\le n} (\alpha_i x^i),$$

$$\alpha_0 = 1,\ \alpha_1 = -\sum_{1\le i\le n} (a_i)\ \alpha_2 = +\sum_{\substack{1\le i\le n \\ 1\le j\le n}\,\big|\, i\neq j} (a_i a_j), \ldots,\ \alpha_n = (-1)^n a_1 a_2 \ldots a_n.$$

19 *Invention nouvelle en l'algebre* (*op. cit.*, note 16): Def. 11: ciii. Compare H. BOSMANS: *Albert Girard et Vieta à propos de la théorie de la "syncrèse" de ce dernier.* Ann. Soc. sc. de Bruxelles: **45** (Louvain, 1926): 34 ff.

20 GREGORY sees it as an obvious thing, giving, in a letter to COLLINS of 26 May 1675 (·≡· GREGORY *TV:* 302—204), $\Sigma(a_i^\lambda)$ in terms of the coefficients of a 7th-degree polynomial, $\lambda = 1, 2, \ldots, 7$, with the remark: "... It is no hard matter to give the rule whereby to continue this in infinitum; for it is so in all equations ..."

21 *AU:* appendix: *de transmutationibus aequationum:* 251—252, to be dated in the 1670's by manuscript drafts in the Portsmouth Collection and the original Lucasian lectures (of which a copy is deposited in Cambridge University Library) (Dd. 9.68).

22 For example, in problems 28, 58, and 61.

23 CUL Add. 3961: 19R—23V, especially 19V—20R, to be dated about 1695, partially published in W. W. R. BALL: *Newton's classification of cubic curves*, Proc. London Math. Soc. **22** (1891): 104—143, appx. 1: 132—140, especially 85—88.

24 19V—20R.

$X \to O'$ has $x = OO'$, while one value of y is zero, so that $y_1 y_2 = \lambda OO'^2 + \mu OO' = 0$. These conditions are sufficient to evaluate all the unknowns, and so, taking $\frac{BP - BP'}{-OA} = \frac{BC}{OC}$ (to define a (unique) point C in OO') $a = -\frac{OA}{OC}$ and we can rewrite the defining equation as $y^2 - \frac{OA}{OC}(x - OC) - \frac{BP \cdot BP'}{BO \cdot BO'} x(OO' - x) = 0$. In particular, we have shown that $\frac{BP \cdot BP'}{BO \cdot BO'} = \lambda$*, constant, true for all conic chords PP' parallel to OA, which is "NEWTON's theorem"[25] for the conic; while $OA = 0$ ($A \to O$, and so OA tangent at O) yields $y^2 - \frac{BP \cdot BP'}{BO \cdot BO'} x(OO' - x) = 0$, which in the form $y^2 : x\ (OO' - x)$ $(= YX^2 : OX \cdot XO') = BP^2 : OB \cdot BO'$ is APOLLONIUS' defining "symptom" for the general conic.

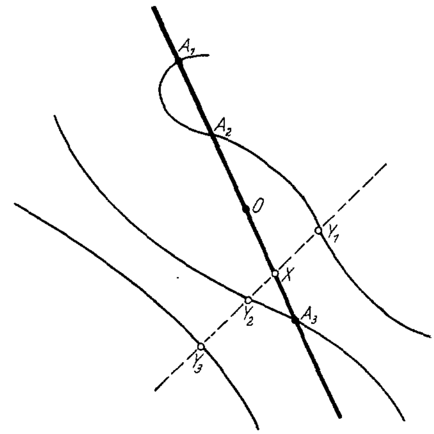


Fig. 5

Extension to the cubic[26] is similar. Let a cubic be cut by the line OA in three (real) points A_1, A_2, A_3, and with respect to some fix-point O on OA and coordinates $OX = a$, $XY = y$ (where XY, inclined at some fixed angle to OA, meets the cubic in Y_1, Y_2, Y_3**) take the representing equation of the cubic by

$$y^3 - y^2(ax + b) + y(rx^2 + sx + t) - (\lambda x^3 + \mu x^2 + \nu x + \pi) = 0.$$

For $X \to$ each of A_1, A_2, A_3, one corresponding value of y is zero, or for $\varrho = OA_1$, OA_2, OA_3 successively $y_1 y_2 y_3 = 0$, $= \lambda\varrho^3 + \mu\varrho^2 + \nu\varrho + \pi$. These are sufficient to define μ, ν, π in terms of λ by $\mu = -\lambda(OA_1 + OA_2 + OA_3)$, $\nu = +\lambda(OA_1 \cdot OA_2 + OA_2 \cdot OA_3 + OA_3 \cdot OA_1)$ and $\pi = -\lambda \cdot OA_1 \cdot OA_2 \cdot OA_3$, so that $\lambda x^3 + \mu x^2 + \nu x + \varrho \equiv (x - OA_1) \cdot (x - OA_2) \cdot (x - OA_3)$. Finally, for X at a general point, $x \doteq OX$, $y = XY_1$ or XY_2 or XY_3, so that $y_1 y_2 y_3 = \lambda \cdot (OX - OA_1)(OX - OA_2)(OX - OA_3) = XY_1 \cdot XY_2 \cdot XY_3$, or $\frac{XY_1 \cdot XY_2 \cdot XY_3}{XA_1 \cdot XA_2 \cdot XA_3} = \lambda$, constant, which is "NEWTON's" theorem for transversals in fixed directions from a point to a cubic, and is clearly generalisable immediately to the n-degree curve.

Implicit in NEWTON's treatment is the counterpart of the analytical theorem that a n-degree polynomial has just n zeros—*viz:* the idea that a line given a

---

* Specifically, $\nu = O$, $\mu = -\lambda OO'$; $b = -OA$,

$$a = \frac{BP - BP' + OA}{OB} = \frac{OA}{OB}\left(\frac{BP - BP'}{-OA} - 1\right),$$

and finally

$$-BP \cdot BP' = \lambda OB(OB - OO'), \quad \text{or} \quad \lambda = \frac{BP \cdot BP'}{BO \cdot BO'}.$$

** NEWTON, in fact, to simplify geometrical calculation has, as with the conic, O coincident with one of A_1, A_2, A_3.

[25] In fact, APOLLONIUS: *Conics:* Bk 3: prop. 17.

[26] 20Rff.

general position meets an n^{th} degree algebraic curve in just n points (if all are real)—and this is basic in a well-known lemma in PM[27]: "There is no oval figure whose area, cut off by right lines at pleasure, can be found generally through equations whose dimensions and number of terms is finite." [It is not quite clear what NEWTON means by his "oval figure" *(figura ovalis)*, but it has been taken by commentators in general as some simple continuous closed curve]: "... Within an oval let there be given some point around which as pole there revolves perpetually a line [with uniform motion]*, while at the same time in that line a moving point goes out from the pole, proceeding always with a speed proportional to the [square] of the line contained within the oval. By this motion the point will describe a spiral with infinite gyrations [round the pole]."[28] NEWTON argues that this spiral is[29] "but one simple curve and irreducible to further curves" and then introduces the idea of defining its nature by considering its meet with a line given in general position: since the spiral, so defined, makes an infinite number of ever-increasing gyrations round the pole, the number of these meets will be infinite, and, further, the "law" and "calculus" for each meet-point will be the same. In amplification NEWTON supposes that some defining equation $\Phi(x)=0$ exists, which gives each distance, x, of the meet of spiral and the given line from some fix-point on that line, and that therefore the function $\Phi(x)$ is unique (giving *all* intersections of the line and the spiral).

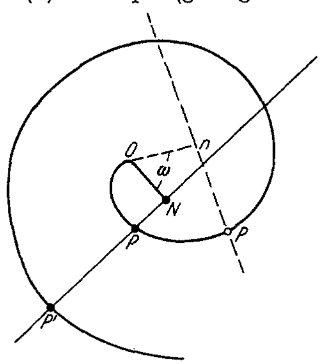


Fig. 6

More exactly, let us suppose that ON is the perpendicular to the line from pole O, and that the line in rotating through some angle ω round O has some intersection P with the spiral pass into a new intersection p; and that $\Phi'(x,\omega)=0$ is the equation which defines the distance of the intersection p from the same fix-point on the line. After one whole revolution $(\omega=2\pi)$, P will pass into a second meet P' of the original fix-line with the spiral, so that $\Phi(x)=0$ has a common root with $\Phi'(x, 2\pi)$. Similarly $\Phi(x)=0$ has a zero equal to one of each $\Phi'(x, 2\lambda\pi)$, $\lambda=1,2,3,\ldots,n$, and we conclude that eventually (if $\Phi(x)$ is of finite degree) we exhaust all zeros of $\Phi(x)$ by identifying them with a zero of each of the $\Phi'(x, 2\lambda\pi)$; and so justify its existence (so that we have $\Phi(x)$ unchanged by revolutions of the line through multiples of 2π). On this preliminary basis (not given as rigorously as stated, but verbally) NEWTON argues that, since the spiral in its infinite ever-increasing revolution round the pole must cut the line in an

* I use square brackets to denote corrections and additions from PM_2 (1713).

27 PM_1: Bk 1: lemma 28: 105–107 (with corrections from PM_2).

28 Clearly by simple stretching transforms (continuously defined) along lines through the pole we can reduce the oval to a circle which has the spiral pole for its centre, and NEWTON'S argument seems to be merely the inverse generalisation. In the circle (which we can see as the canonical case) the spiral becomes ARCHIMEDEAN with (polar) representing equation $r=a^2\vartheta$, where r is the positive distance of a general point on the spiral from the pole, a the radius-length of the circle and ϑ the radian-measured angle of rotation.

29 PM_2 (1713).

infinite number of points, $\Phi(x)$ cannot be an algebraic equation of finite degree (and so the spiral is likewise not representable by a polynomial of finite degree).

Thus far NEWTON shows a deep insight, but he applies the argument in an unjustifiable way by considering the spiral whose point-distance, r, from the pole is given by $r = \frac{1}{2\pi} a^2\vartheta$ where ϑ is the angle of revolution[30]. Following his argument we argue: since the area of the corresponding sector OAQ of the "oval" (here taken for simplicity as the canonical circle of radius a whose centre is the pole of the spiral) is $a^2\vartheta$ [31], the distance r can be used to represent the area of the "oval" sector*: but the (Cartesian) representing equation of the spiral, which meets the fix-line above in an infinite number of points, must be of infinite degree, and so correspondingly the general circle segment cannot be represented by a polynomial of finite degree. The argument is plausible, greatly subtle and involved, but the conclusion wrong[32], and remarkable for its deft intermanipulation of concepts derived from the abstract theory of the polynomial and from a corresponding geometrical model.**

Much of 17th century work on polynomials was, however, not concerned with such theoretical existence considerations, but remained concerned with the pre-eminently practical viewpoint of producing refined methods of approximating to the roots of equations. With such an attitude the testable results had priority over rigour of method—whether, on physical substitution of a particular value in a polynomial form, a zero was produced (or near enough). So we find a wide variety of numerical methods introduced without pretension to rigour or theoretical justification in many cases. Most, in fact, depend on some adaptation of a basic principle—to be formalized rigorously with respect to a tightly defined concept of continuous function by BOLZANO in the 19th century—that where $\Phi(x)$ is a polynomial form continuous in the interval $x \in [a, b]$ such that

* Since they differ only by the factor of $1/2\pi$.

** The fallacy, never previously pointed out to my knowledge, lies in the uncritical representation of circle-sector area by line-length. Restricting our attention to the (infinite number of) meets of OA, with the spiral, say $A i$, $i = 1, 2, 3, \ldots$, what each length OA_i represents is not the simple area of the circle centre O and radius a, but this *same* area taken i times; and, in general, where $A_1\widehat{O}B_1 = \vartheta$ and OB_1 meets the spiral in successive points B_i, OB_i measures the circle area taken $\left(i + \frac{\delta}{2\pi}\right)$ times. It is evident that the infinite gyration of the spiral expresses the *periodicity* of the general angle of OB_i with $OA_i = 2i\pi + \vartheta$, and has nothing to do with the circle-area which remains invariable, a basic undefined quantity.

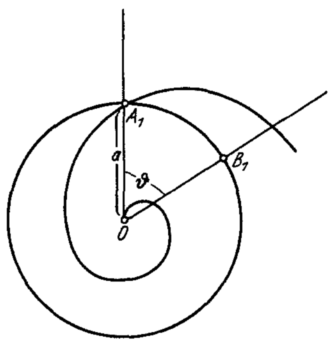


Fig. 7

[30] I take, for simplicity, the canonical form of the oval by the circle of radius r whose centre is the pole (see note 28).

[31] Which means that the spiral will pass through the meet of the circle with the tangent to the spiral at the pole (A_1 in the next diagramm).

[32] So H. BROUGHAM and E. J. ROUTH in their *An Analytical view of Sir Isaac Newton's Principia*, London, 1855: 72—74 give the counter-example of the closed oval $y^m = x^{(n-1)m} \cdot (a^n - x^n)$, m, n even integers, which has an exact quadrature.

$\Phi(a) \leqq 0 \leqq \Phi(b)$, there is at least one zero for x between a and b: that is, for $\vartheta \in [0, 1]$, $\Phi(a+\vartheta(b-a))=0$ is true[33]. The practical use of this is that by taking a and b closer and closer we can approximate to a zero of $\Phi(x)$ with more and more accuracy (this is done in an immediate way by splitting the interval $[a, b]$ into (smaller) subintervals) so that for one, say $[a', b']$, at least it will be true $\Phi(a') \leqq 0 \leqq \Phi(b')$, with $a \leqq a'$, $b \geqq b'$.

However, simplicity of application is a keynote of a numerical method, and early techniques proved very cumbrous to apply[34], and perhaps NEWTON'S modification and simplification of VIETA'S approach[35] was the first practicable approximation-method. Thus, in his example, $\Phi(y) \equiv y^3 - 2y - 5 = 0$ (an example later to become a standard test for the efficacy of numerical methods) we see $\Phi(2) = -1 < 0 < \Phi(3) = 16$, and so by the continuity postulate there is a zero at some $y \in [2, 3]$. Take $y = 2 + p$, or $\Phi'(p) \equiv -1 + 10p + 6p^2 + p^3 = 0$. From this point NEWTON takes the linear approximation $-1 + 10p \approx 0$, $p \approx 0.1$, and the process repeats by $\varrho = 0.1 + q$; $0.061 + 11.23q \approx 0$ or $q \approx -0.0054$, $q = -0.0054 + r$; $0.005416 + 11.162r \approx 0$, or $r \approx 0.00004852$, which NEWTON considers sufficiently exact. The method extends easily to the two-variables polynomial $\Phi(x, y) = 0$, yielding an appropriate series expansion for y, $y = X(x)$ (where X satisfies $\Phi(x, X(x)) \equiv 0$).*

This process (after the first stage) of using linear approximation typifies a standard problem of finding general ways of "iterating" a polynomial zero, of systematizing ways of deriving successive approximations in a recursive way without the troublesome task of deriving a new polynomial form at each successive step (which is necessary in NEWTON'S method). A first step was taken independently by JAMES GREGORY[36] and MICHAEL DARY[37] in solving equations of the form $x = \Phi(x)$:

---

* Here, of course, convergence has to be considered, and while NEWTON shows himself familiar with the implicit theoretical restrictions, the lack of rigour makes exact justification difficult.

[33] The assumption is not, of course, original with the 17th century, but used in numerical methods given by such 16th century mathematicians as STEVIN, BURGI and VIETA—see TROPFKE$_3$ 3: 157—159.

[34] PELL in mid-century could quote WARNER on VIETA'S approximation method as saying that "to attempt the same (finding of a polynomial root) in VIETA'S method [is] work unfit for a Christian and more proper to one that can undertake to remove the Italian Alps into England ..." (quoted by COLLINS in a letter to OLDENBURG of the early 1670's—see RIGAUD (C) 1: 247—248).

[35] First made public in a letter to Collins of 20 June 1674 (see RIGAUD (C) 2: 362—365), but given generally in his first letter to LEIBNIZ in 1676 (compare OLDENBURG-LEIBNIZ, 26 July 1676, = GERHARDT (B). 1: 179—192, especially 183—185). The method appears widely in the manuscript drafts in the Portsmouth Collection, and is given in the printed *de quadratura curvarum* and HORSLEY'S manuscript collection, *geometria analytica* (see HORSLEY: *Newtoni opera* 1: 391 ff.), along with his famous "parallelogram" rule for dealing with the two-variabled polynomial $\Phi(x, y) = 0$.

[36] Reported in a letter to COLLINS of 2 April 1674—see GREGORY TV; 278—279, and compare 394—395. GREGORY considers the equation $b^{n-1}c = b^{n-2}(b+c)x - x^n$, which reduces to $x' = \dfrac{x^n}{b^{n-2}(b+c)}$ by the substitution $x = x' + \dfrac{bc}{b+c}$.

[37] Who considers the equation $x^p = ax^q + n$, $p > q$ (or $x = (ax^q + n)^{1/p}$) in a letter to NEWTON of 15 August 1674 (RIGAUD (C) 2: 365—366).

choosing some close first approximation x_0, the approximation sequence $x_1, x_2, \ldots$ is found by $x_{i+1} = \Phi(x_i)$ a simple and plausible method though (neither hint of restrictions which are necessary for the sequence to converge to a limit*). A more general procedure was found by NEWTON[38] which, though first published in 1685,[39] received a very full treatment by JOSEPH RAPHSON[40] (though the unnecessary restriction to an algebraic n-degree polynomial is made). Interestingly, while RAPHSON bases his development on a cumbersome variable substitution, he gives virtually a TAYLOR expansion. Consider the n-degree polynomial $\Phi(x) = 0$ and some suitably close approximation to a root X: RAPHSON defines a new variable x' by $x = x' + X$, substitutes in the polynomial and expands to derive the equivalent of $(0 =)\ \Phi(x' + X) = \Phi(X) + x'\,\Phi'(X) + \frac{x'^2}{2!}\Phi''(X) + \frac{x'^3}{3!}\Phi'''(X) + \cdots + \frac{x'^n}{n!}\Phi^{(n)}(X)$, or, since x' is small in comparison with X (by choice of a suitable X), we can take $\Phi(X) + x'\,\Phi(X) = 0$, very nearly, or $x = X + x' = X - \frac{\Phi(X)}{\Phi'(X)}$, and in general $x_{i+1} = x_i - \frac{\Phi(x_i)}{\Phi'(x_i)}$.[41] (The simple justification by appeal to the corresponding geometrical representation is not made[42].)

An intriguing application of the general continuity principle $\big(-\Phi(a) \leqq 0 \leqq \Phi(b)$ implies $\Phi(a + \vartheta(b-a)) = 0$ for at least one $\vartheta \in [0, 1]$—where $\Phi(x)$ is restricted to being a 2-degree polynomial—$\big)$ was made by BROUNCKER in his work on the general FERMAT equation $n\alpha^2 + 1 = \beta^2$, where α, β are restricted to being integers

---

* Justification on the geometrical model is immediate: the sequence defines the root by generating a continuous broken line between curves $y = x$, $y = \Phi(x)$, parallel to ordinate and abscissa alternately, which converges to their meet—that is, at the point such that $y = x = \Phi(x)$.

[38] Apparently some time in 1675, perhaps in pondering over DAVY's letter (see note 37), but given in a letter to COLLINS of 24 July 1675 (RIGAUD (*C*) **1**: 372), where he iterates $A^{1/n}$ by $x_{i+1} = \frac{1}{n}(n-1)\,x_i + \frac{A}{n \times x^{n-1}}$.

The early 15[th] century mathematician JAMSID AL-KĀSĪ seems to have sketched in the first stage of NEWTON's formula in his *Miftah al-Hisab (Key to arithmetic)* Bk. **1** (*cf.* Russian translation by B.A. ROSENFELD and A.P. YUŠKEVIČ, Moscow, 1956). Briefly, to derive $A^{1/n}$ AL-KĀSĪ takes the equation $x^n - A = 0$ and a first approximation $x_0 = [A^{1/n}]$ (the first integer smaller than $A^{1/n}$), and $A = x^n = (x_0 + x')^n$ yields $(x_0^n - A) + n\,x_0^{n-1}x' \approx 0$ or $x' \approx \frac{A - x_0^n}{n\,x_0^{n-1}}$, or $x = A^{1/n}, \approx x^0 + \frac{A - x_0^n}{n\,x_0^{n-1}}$. This is, of course, the first stage of the NEWTON-RAPHSON iteration (though AL-KASI does not iterate, content to take that as his approximation) and is equivalent to the above.

[39] In WALLIS' *Algebra:* 338.

[40] In his *analysis aequationum universalis seu ad aequationes resolvendas methodus generalis ... ex nova infinitarum serierum deducta*, London 1690; of which an outline is given by WALLIS in the Latin edition of his *Algebra*—see WALLIS *opera* **2** (1693): 396—397.

[41] The NEWTON-AL-KĀSĪ root approximation recursion follows by taking $\Phi(x) \equiv x^n - A$.

[42] Specifically, where $y = (\Phi x)$, $\frac{\Phi(x)}{\Phi'(x)} = y\frac{dx}{dy}$ is the subtangent. Significantly, the first rigorous treatments of the method were elaborated on ideas derived from this geometrical approach—compare J.R. MOURRAILLE: *Traité de la réduction des équations ...* Pt. **1**: Paris, 1768; J. FOURIER: *Analyse des équations déterminées ...* Paris, 1818; and F. CAJORI: *Fourier's improvement of the Newton-Raphson method of approximation.* Bibliotheca mathematica$_2$ **11** (1910—1911): 132—137.

and n is a (non-square) integer.[43] Neither WALLIS nor BROUNCKER realized at first the force of the restriction to integer solutions, and BROUNCKER, using the identity $(\lambda^2-n)^2+n(2\lambda)^2=(\lambda^2+n)^2$, contented himself with the rational solution $n\cdot\left(\frac{2\lambda}{\lambda^2-n}\right)^2+1=\left(\frac{\lambda^2+n}{\lambda^2-n}\right)^2$. However, a little later BROUNCKER told WALLIS of FERMAT'S insistence on integer solutions, and WALLIS derived a recursive rule for an infinity of solutions, given one, from BROUNCKER'S rule by taking $\lambda=r/s$*. There remained the problem of deriving particular solutions systematically, and this was solved by BROUNCKER.[44]

To exemplify his method, consider $\Phi_1(A)\equiv A^2-13a_0^2=1$. The continuity rule gives $\Phi_1(3a_0)<1<\Phi_1(4a_0)$, or $A=3a_0+\vartheta a_0$, for some $\vartheta\in[0,1]$. Taking $A=3a_0+a_1$, we deduce $1=-4a_0^2+6a_0a_1+a_1^2\equiv\Phi_2(a_1)$, in which $\Phi_2(a_1)>1>\Phi_2(2a_1)$ and we take $a_0=a_1+a_2$. Similarly

$$
\begin{array}{lll}
1=\Phi_3(a_1)\equiv 3a_1^2-2a_1a_2-4a_2^2: & \Phi_3(a_2)<1<\Phi_3(2a_2): & a_1=a_2+a_3;\\
1=\Phi_4(a_2)\equiv -3a_2^2+4a_2a_3+3a_3^2: & \Phi_4(a_3)>1>\Phi_4(2a_3): & a_2=a_3+a_3;\\
1=\Phi_5(a_3)\equiv 4a_3^2-2a_3a_4-3a_4^2: & \Phi_5(a_4)<1<\Phi_5(2a_4): & a_3=a_4+a_5;\\
1=\Phi_6(a_4)\equiv -a_4^2+6a_4a_5+4a_5^2: & \Phi_6(6a_5)>1>\Phi_6(7a_5): & a_4=6a_5+a_6;\\
1=\Phi_7(a_5)\equiv 4a_5^2+6a_5a_6-a_6^2: & \Phi_7(a_6)<1<\Phi_7(2a_6): & a_5=a_6+a_7;\\
1=\Phi_8(a_6)\equiv -3a_6^2+2a_6a_7+4a_7^2: & \Phi_8(a_7)>1>\Phi_8(2a_7): & a_6=a_7+a_8;\\
1=\Phi_9(a_7)\equiv 3a_7^2-4a_7a_8-3a_8^2: & \Phi_9(a_8)<1\leqq\Phi_9(2a_8); & a_7=a_8+a_9;
\end{array}
$$

and finally $1=\Phi_9(2a_8)\equiv a_8^2$, which is solved by taking $a_8=1$. So, working backwards $a_7=2$, $a_6=3$, $a_5=5$, $a_4=33$, $a_3=38$, $a_2=71$, $a_1=109$, $a_0=180$, $A=649$: $13\cdot(180)^2+1=(649)^2$—and the same procedure is to be used in the case of any n". We notice that this is, implicitly, a continued fraction expansion of $\sqrt{13}$: $\frac{A}{a_0}=3+\frac{1}{\frac{a_0}{a_1}}$, $\frac{a_0}{a_1}=1+\frac{1}{\frac{a_1}{a_2}},\ldots,\frac{a_6}{a_1}=1+\frac{1}{\frac{a_1}{a_8}}$, or $\frac{A}{a_0}=\sqrt{13+\frac{1}{a_0^2}}\approx\sqrt{13}$, $=3+\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{6+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{2}=\frac{649}{180}$. BROUNCKER went on to notice[45] that the sequence Φ_i can, in fact, be continued indefinitely:

$$
\begin{array}{lll}
1=\Phi_9(a_7)\equiv 3a_8^2-4a_8a_9-3a_9^2: & \Phi_9(a_8)<1\leqq\Phi_9(2a_8): & a_7=a_8+a_9;\\
1=\Phi_{10}(a_8)\equiv -4a_9^2+2a_9a_{10}+3a_{10}^2: & \Phi_{10}(a_9)\geqq 1>\Phi_{10}(2a_9): & a_8=a_9+a_{10};
\end{array}
$$

---

* Specifically, $n\left(\frac{2rs}{r^2-ns^2}\right)^2+1=\left(\frac{r^2+ns^2}{r^2-ns^2}\right)^2$, which yields an integer solution if $r^2-ns^2=\pm1$.

43 This work carried out in partial collaboration with WALLIS in 1657, was published in (*CE*) *commercium epistolicum de quaestionibus quibusdam mathematicis nuper habitum*, Oxford, 1658, and summarised in ch. 98 of WALLIS' *Algebra* (1685): 363–372, = *opera* 2 (1693); 418–426. BROUNCKER had received the problem from FERMAT at the beginning of September 1657 (*cf. CE* No. 8: FERMAT'S "scriptum" is set in appendix) though FERMAT had originally posed the problem to FRÉNICLE in February 1657 (FERMAT *OE* 2: 333ff.). A general discussion is in H. KONEN: *Die Geschichte der Gleichung* t^2-Du^2-1, Leipzig, 1901, and E.E. WHITFORD: *The Pell equation*, New York, 1912; especially 47–58; and compare J.E. HOFMANN: *Neues über Fermats zahlentheoretische Herausforderungen von 1657*: Abh. der Preuß. Akademie der Wissenschaften (1943), Nr. 9, Berlin 1944.

44 About November 1657.

45 *CE*: No. 19.

$1 = \Phi_{11}(a_9) \equiv 9_{10}^2 - 6a_{10}a_{11} - 4a_{11}^2$: $\Phi_{11}(a_{10}) \leqq 1 < \Phi_{11}(2a_{10})$: $a_9 = a_{10} + a_{11}$;
$1 = \Phi_{12}(a_{10}) \equiv -4a_{11}^2 + 6a_{11}a_{12} + a_{12}^2$.

We notice now that $\Phi_{12}(\lambda) \equiv -4a_{11}^2 + 6a_{11}\lambda + \lambda^2$, $\equiv \Phi_2(\lambda)$ if we replace a_{11} by a_1, and so the cycle repeats itself: $\Phi_i \equiv \Phi_{i+10k}$, $k = 1, 2, 3, \ldots$. This, of course, mirrors the periodicity of the continued fraction for $\sqrt{13}$:

$$13 = \lim_{l\to\infty}\left(3:\ 1, 1, 1, 1, 6;\ 1, 1, 1, 1, 6; \ldots, \frac{a_l}{a_{l+1}}\right) = (3:\ \overline{1, 1, 1, 1, 6}).$$

Further, noting with BROUNCKER that $\Phi_6(1) = -1$, $13\left(3 + \frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1+}\,\frac{1}{1}\right)^2 - 1$ is a square and alternate periods of the continued fraction give solutions of $13a_0^2 - 1 = A^2$.[46]

To return to a more general consideration, many techniques developed in polynomial theory have an obvious (but not always factual) simplicity and seemed to require no profound justification, and we find such concepts as: if for all x $\Phi(x) = \Psi(x)$, then $\Phi \equiv \Psi$ (where $x \in$ some $[a, b]$)*. A particular case is the coefficient comparison axiom: taking $\Phi(x) = \sum_{0\leqq i\leqq n}(a_i x^i)$, $\Psi(x) = \sum_{0\leqq i\leqq n}(b_i x^i)$ then, if $\Phi(x) \equiv \Psi(x)$, for all i $a_i = b_i$—a technique widely used to effect such transforms as reversal of series or series reciprocation**, and, as we have seen, the (unique) factoring of a n-degree polynomial $f(x) \equiv \sum_{0\leqq i\leqq n}(a_i x^i)$ into its n factors $\prod_{1\leqq j\leqq n}(x - \alpha_j)$ which can be used

* Though, a case in point, some modern axiomatisations of set theory would not allow it in an unqualified form—which shows its essential arbitrariness.

** Given $\Phi(x) \equiv \sum_{0\leqq i\leqq n}(a_i x^i)$, to derive $\frac{1}{\Phi(x)} = \sum_{0\leqq j\leqq n}(b_j x_j)$ (with convergence more or less assumed as $m \to \infty$).

[46] No integers a_0, A can, of course, satisfy $n \cdot a_0^2 - 1 = A^2$ where $n \equiv 3 \pmod 4$. The equivalents of other continued fraction expansions are given by BROUNCKER, as

$$\sqrt{13} = (4 -: \overline{2+, 2-, 8-}),$$
$$\sqrt{109} = (10 +: \overline{2+, 4-, 3+, 5-, 7+, 7-, 5+, 3-, 4+, 2+, 20+}),$$
and
$$\sqrt{21} = (5 -: \overline{2+, 2+, 2-, 10-})$$
$$\sqrt{433} = (21 -: \overline{5+, 4+, 2+, 3-, 4+, 14-, 3-, 2+, 13+, 4-, 3+, 2+, 4+, 5-, 42-}),$$

where alternative periods solve $na_0^2 \pm 1 = A^2$. Really, all there remained to do was the not so difficult task of proving that for all non-square integers n the "BROUNCKER" periods are finite (and of even length). Indeed WALLIS, in ch. 99 of his *Algebra* (1685), tries to show existence of a solution by considering $na^2 + 1 = \lambda^2$. Since $a\sqrt{n} < \sqrt{na^2+1} < a\sqrt{n} + \frac{1}{2a\sqrt{n}}$, $< a\sqrt{n} + 1$, he easily proves that $\sqrt{na^2+1}$ is the integer next greater than $a\sqrt{n}$; and so reduces the existence condition to $\frac{x}{a} < p < \frac{z + (z^2 + 4pr)^{\frac{1}{2}}}{2a}$, where $z = [ap]$, $r = \frac{1}{2\sqrt{n}}$ $p = [\sqrt{n}] + 1 - \sqrt{n}$. WALLIS' further argument is circular, in effect stating that "obviously" this condition can be satisfied for all n. Indeed, J.L. LAGRANGE's first existence proof (in *Solution d'un probleme d'arithmétique*, Miscellania Taurinensis **4** (Turin, 1766): 41 ff.) uses a not unsimilar reduction; but his second proof (*Sur la détermination des problèmes indéterminés du second degré*, Histoire de l'ac. sc. de Berlin **23** (1767): 272ff.) uses the easy proof that the continued fraction expansion of $\sqrt{n}$ is finite-periodic.

to derive the relations between the a_i, α_j by equating $\sum_{0\leq i\leq n} (a_i x^i) \equiv a_n \prod_{1\leq j\leq n} (x-\alpha_j)$. More generally, we find a widespread use of standard factorisations in the period—elementary forms of which existed in classical Greek mathematics defined on a geometrical model of rectangle area[47] but of which a compact free variable notation allowed a much greater conciseness of expression and generality of treatment. A choice example is to be found in the sum-series $\frac{\pi}{2\sqrt{2}} = 1 + \frac{1}{3} - \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \frac{1}{11} - \frac{1}{13} - \frac{1}{15} + \cdots$ which NEWTON derived from the factorisation $x^4+1 \equiv (x^2+\sqrt{2}x+1)(x^2-\sqrt{2}x+1)$★ in a neat counterblast to the sum-series $\frac{1}{4}\pi = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$ communicated by LEIBNIZ.[48]

But what perhaps reveals most fully the incipient power of the free-variabled polynomial form are the subtle and widely varied structural delimitations to which it can be successfully applied. Without an adequate concept of and notation for free variable this would be, in all but the simplest cases, a supremely difficult if not impossible task. Usually the structural delimitation involves one or more conditions of the type $(x)\ [K(\Phi(x))]$, where $K(\Phi(x))$ is some delimiting condition on the function $\Phi(x)$. Understandably clear cases of such a reasoning pattern are rare in the 17th-century[49], but there occurs a fine example in the form which

★
$$\begin{aligned}\int_0^1 \frac{1+x^2}{1+x^4}\cdot dx &= \int_0^1 \left(\frac{1}{1+(\sqrt{2}x+1)^2} + \frac{1}{1+(\sqrt{2}x-1)^2}\right) dx \\ &= \left[\frac{1}{\sqrt{2}}\left(\tan^{-1}(\sqrt{2}x+1) + \tan^{-1}(\sqrt{2}x-1)\right)\right]_0^1 \\ &= \left[\frac{1}{\sqrt{2}}\tan^{-1}\frac{x\sqrt{2}}{1-x^2}\right]_0^1 = \frac{\pi}{2\sqrt{2}}, \\ &= \int_0^1 (1+x^2) \times \lim_{n\to\infty}\Big(\sum_{1\leq i\leq n}((-1)^{i-1}\cdot x^{4i})\Big)\cdot dx \\ &= \lim_{n\to\infty}\left(\sum_{1\leq i\leq n}\left((-1)^{i-1}\left(\frac{1}{4i-3}+\frac{1}{4i-1}\right)\right)\right).\end{aligned}$$

[47] In particular, the results $(x \pm y)^2 = x^2 \pm 2xy + y^2$ and $(x+y)(x-y) = x^2 - y^2$, to be found in EUCLID's *Elements* but probably PYTHAGOREAN.

[48] NEWTON communicated the series to LEIBNIZ through OLDENBURG (see his letter to OLDENBURG, 24 October 1676, GERHARDT (*B*) **1**: 203—225, especially 214). LEIBNIZ' series is of course, the very well known expansion of $\tan^{-1} 1$ (but given first in the 15th century by the Hindu mathematician NILAKANTHA, *cf.* ch. 5), and was communicated in the letter to OLDENBURG of 27 August 1676 (GERHARDT (*B*) **1**: 193—200, especially 193—196) but later to be published with faintly plausible but ill-founded number-mysticism of odd and even, and positive and negative in *AE* (1682): 41—46: *de vera proportione circuli ad quadratum inscriptum in numeris rationalibus*. (His derivation of the series is examined by J.E. HOFMANN in *Entwicklungsgeschichte der Leibnizschen Mathematik* ...; 32—35 on the basis of manuscript sources in the Royal Library, Hanover.)

[49] Perhaps the first such example of delimiting a function by an (implicit) quantified condition is that given by ARCHIMEDES in his treatise *On the equilibrium of planes* Bk 1, where he derives from the conjunction of the two quantified conditions $(h)\ (f(\lambda+h)+f(\lambda-h)=2f(\lambda))$ and $(\lambda)\ (f(\lambda)=-f(-\lambda))$ the result $\lambda\cdot f(\mu)=\mu\cdot f(\lambda)$, where λ, μ are any real numbers (see DIJKSTERHUIS: *Archimedes (op. cit.)*: 286—305). It is significant, however, that the largely verbal argument would become increasingly difficult to control under more complex delimiting conditions on the quantified function.

BROUNCKER derives to satisfy (x) $[\Phi(x-1)\cdot\Phi(x+1)=x^2]$ (where the variable x shall range at least over the positive integers) and from which he deduces his continued fraction expansion for $4/\pi$.* We do not have BROUNCKER's proof of this derivation but from hints given in WALLIS' first published statement[50] of it, it is possible to restore his train of ideas with some assurance of historical authenticity.

BROUNCKER's general result, stated explicitly for a large number of integral values of x,[51] expands $\Phi(x)$ as the continued fraction $\lim_{n\to\infty}\left(x+\frac{1^2}{2x+}\frac{3^2}{2x+}\cdots\frac{(2n-1)^2}{2x}\right)$, which is itself based on the tighter result, given for the two cases $\lambda=0,\infty$,[52]

$$(x,\lambda)\quad\left(\begin{array}{c}\left(x-1+\frac{1^2}{2(x-1)+}\cdots\frac{(2i-3)^2}{2(x-1)+}\frac{(2i-1)^2}{(1+\lambda)x+(2i-1)}\right)\times\\ \times\left(x+1+\frac{1^2}{2(x+1)+}\cdots\frac{(2i-3)^2}{2(x+1)+}\frac{(2i-1)^2}{\left(1+\frac{1}{\lambda}\right)x-(2i-1)}\right)=x^2\end{array}\right).$$

Some attempts to justify this (which may be due to WALLIS rather than BROUNCKER) are sketched in WALLIS' treatment[53] which gives the particular cases $i=1,2,3$ of the theorem that, where $D(x)_i$ is denominator of the i^{th} convergent of $\Phi(x)_n=x+\frac{1}{2x+}\frac{3^2}{2x+}\cdots\frac{(2n-1)^2}{2x}$,

$$\Phi(x)_i\times\Phi(x+2)_i=(x+1)^2-(-1)^i\left(\frac{1^2\cdot 3^2\cdots(2i-1)^2}{D(x)_i\,D(x+2)_i}\right)$$

—a theorem which can only have been derived by induction over particular cases worked out physically, but which makes plausible the limit-form $[\Phi(x)_\infty=]$ $\Phi(x)\times\Phi(x+2)=(x+1)^2$.

This does not, however, throw light on the derivation of the form of $\Phi(x)$, a process which is restorable from other hints given in the following way[54]: Clearly, since $(x-1)(x+1)=x^2-1<x^2$, $\Phi(x)>x$, and we may assume that $\Phi(x)=x+\frac{\alpha_1}{\Phi_1(x)}$, α_1 some constant to be particularised at will.** Substituting,

* The application is made in chapter four, where BROUNCKER shows $\Phi(1)=\square$ $(=4/\pi)$.

** A step more naturally taken in the 17th century, when a continued-fraction method of numerical approximation was widely used.

[50] In *AI* 1656: prop. 191: *propositum sit inquirere quantus sit terminus* $\square$ $[=4/\pi]$ ... *in numeris absolutis quam proxime*: *idem aliter* and *scholium*.

[51] *AI*: 182.

[52] The case $\lambda=1$ was considered by GUSTAV BAUER: *Von einem Kettenbruch Eulers und einem Theorem von Wallis*, Abhandlungen der kgl. bayr. Akademie der Wissenschaften zu München. **11.** 2 (1872): 92ff., but the general theorem is given here for the first time.

[53] *AI*: 183.

[54] The restoration given was made, in the first instance, solely on the basis of the text, but was confirmed later on reading various articles by EULER, who concerned himself with the problem intermittently over much of his life. Compare EULER's varying attempts in, for example, *de fractionibus continuis observationes*, Comm. ac. sc. Petrop. **11** (1739) 1750: 32–81 ·≡· *opera omnia* **15** 1 (1925): 291–345; *de seriebus in quibus producta ex binis terminis contiguis datam constituunt progressionem* ·≡· *opuscula analytica* **1** (St. Petersburg, 1783): 3–47; and *de fractionibus continuis Wallisii*, Méms. de l'ac. des sci. de St. Pétersburg. **5** (1812) 1815: 24–44 ·≡· *opera omnia* **16** 2 (1925): 178–199.

we have on multiplying and cancelling

$$-\Phi_1(x-1)\,\Phi_1(x+1)+\alpha_1(x+1)\,\Phi_1(x+1)+\alpha_1(x-1)\,\Phi\,(x-1)+\alpha_1^2=0,$$

and an obvious reduction is $\alpha_1=1$, or, on rearranging, $\big(\Phi_1(x-1)-(x+1)\big)\times\big(\Phi_1(x+1)-(x-1)\big)=x^2$. On testing we find a simple way to keep symmetry is the substitution $\Phi_1(x)=2x+\frac{\alpha_2}{\Phi_2(x)}$, so that $\left(x+3+\frac{\alpha_2}{\Phi_2(x+1)}\right)\left(x-3+\frac{\alpha_2}{\Phi_2(x-1)}\right)=x^2$, and we have the beginning of a periodic cycle. At some stage we have, say, $\left(x+(2n-1)+\frac{\alpha_n}{\Phi_n(x+1)}\right)\left(x-(2n-1)+\frac{\alpha_n}{\Phi_n(x-1)}\right)=x^2$. Multiplying, cancelling and rearranging, an obvious reduction is to take $\alpha_n=(2n-1)^2$, after which we can arrange to form $\big(\Phi_n(x+1)-(x-(2n-1))\big)\big(\Phi_n(x+1)-(x+(2n-1))\big)=x^2$, and the substitution $\Phi_n(x)=2x+\frac{\alpha_{n+1}}{\Phi_{n+1}(x)}$ gives the beginning of the next cycle $\left(x+\big(2(n+1)-1\big)+\frac{\alpha_{n+1}}{\Phi_{n+1}(x+1)}\right)\times\left(x-\big(2(n+1)-1\big)+\frac{\alpha_{n+1}}{\Phi_{n+1}(x-1)}\right)=x^2$. Working back from the i^{th} stage, we can "unwrap" the cycle, finding $\Phi(x)=x+\frac{1^2}{2x+}\,\frac{3^2}{2x+}\cdots\frac{(2i-3)^2}{2x+}\,\frac{(2i-1)^2}{\Phi_i(x)}$, and the BROUNCKER expansion is the limit-form as $i\to\infty$. Further, the extended BROUNCKER results follow by choosing special forms $\Phi_n(x+1)$, $\Phi_n(x-1)$ which make the condition $\left(x+(2n-1)+\frac{(2n-1)^2}{\Phi_n(x+1)}\right)\left(x-(2n-1)+\frac{(2n-1)^2}{\Phi_n(x-1)}\right)=x^2$ an identity.[55]

Unless the letters in which BROUNCKER revealed his ideas to WALLIS still exist—they appear irretrievably lost[56]—such restoration must remain merely plausible, and perhaps, after all, they were merely abstracted by induction from particular instances.[57] Yet the development remains a fascinating example of

[55] In fact a general form (worked out with a little trouble) is

$$\left(x+(2n-1)+\frac{(2n-1)^2}{\left(1+\frac{1}{\lambda}\right)x-(2n-1)}\right)\left(x-(2n-1)+\frac{(2n-1)^2}{(1+\lambda)\,x+(2n-1)}\right)\equiv x^2,$$

but the particular cases

$$(\lambda=0)\quad (x+(2n-1))\left(x-(2n-1)+\frac{(2n-1)^2}{x+(2n-1)}\right)\equiv x^2,$$

and

$$(\lambda=\infty)\quad \left(x+(2n-1)+\frac{(2n-1)^2}{x-(2n-1)}\right)(x-(2n-1))\equiv x^2$$

given by BROUNCKER are more immediate.

[56] These letters according to remarks in *AI:* prop. 191; *idem aliter* seem to have been communicated some time in 1654—1655, while the earliest extent correspondence between WALLIS and BROUNCKER (that printed in *CE*) dates from 1657. PAUL TANNERY, however cites unpublished letters of WALLIS to BROUNCKER of 16 and 20 October 1656 which he found in Vienna in 1899 in a collection then in the Hofbibliothek (manuscript 7050: 424—425) (*see Mémoires scientifiques*, **6**: 373). Perhaps some light will be shed if these are traced.

[57] A strong argument against accepting such an induction as plausible is that the BROUNCKER continued fraction expansion of $\Phi(x)$ is in no sense unique. So (for general λ) an alternative form is

$$\Phi(x)=\lambda-2+\frac{\alpha}{\lambda+}\,\frac{(x+1)^2(\alpha+2\beta)}{4\gamma+}\,\frac{(x+3)^2\alpha(\alpha+4\beta)}{4(\gamma+\beta)+}\,\frac{(x+5)^2(\alpha+2\beta)(\alpha+6\beta)}{4(\gamma+2\beta)+}\cdots,$$

where

$$\alpha=(x+1)^2-\lambda(\lambda-2),\quad \beta=2(x+2-\lambda),$$

a quantified delimitation, and the result of a subtlety not to be surpassed in the 17th century.

Finally, other than in free variable analysis little was done in 17th century algebra, though much now formulated in an abstract algebraic form—various concepts of transform, for example—was developed elsewhere, as part of pure geometry or in an unrelated technique. So, it is an historical curiosity that the theory of permutations remained a mere numerical study (tied up with a "LAPLACIAN" probability theory) which did little more than enumerate possible varieties under varying conditions without examining their nature (and so developing such concepts as group, invariance and identity transform). In general, transition to a developed concept of algebraic structure had to wait till ever-broadening ideas and techniques had been systematised and operational methods developed, and above all till analytical techniques in geometry were given an algebraical, and not as in the 17th century a classically intuitive and largely extra-logical, basis. Meanwhile the study of algebra had to remain inevitably an unsystematic, piecemeal collection of methods and results.

---

and

$$\gamma = \alpha - \left(\frac{\lambda}{2} - 1\right)\beta;$$

and, in particular, when $\lambda = x + 2$ ($\alpha = \gamma = 1$, $\beta = 0$),

$$\Phi(x) = x + \frac{1}{x+2+}\,\frac{(x+1)^2}{4+}\,\frac{(x+3)^2}{4+}\cdots.$$

More generally we note that the functional equation $x^2 = \Phi(x-1) \times \Phi(x+1)$ is satisfied by

$$\Phi(x) = (x+1)\,\frac{B\left(\frac{x+3}{4}, \frac{1}{2}\right)}{B\left(\frac{x+1}{4}, \frac{1}{2}\right)}, = \left(2 \times \frac{\Gamma\left(\frac{x+3}{4}\right)}{\Gamma\left(\frac{x+1}{4}\right)}\right)^2,$$

which we prove easily by defining

$$X(x) = \frac{\Phi(x)}{x+1} \times \frac{B\left(\frac{x+1}{4}, \frac{1}{2}\right)}{B\left(\frac{x+3}{4}, \frac{1}{2}\right)}$$

(compare ch. 4) and using the reduction

$$\frac{B\left(\frac{x+5}{4}, \frac{1}{2}\right)}{B\left(\frac{x+1}{4}, \frac{1}{2}\right)} = \frac{x+1}{x+3};$$

in fact,

$$\Phi(x)\;\Phi(x+2) \equiv (x+1)^2 = (x+1)\,(x+3)\,\frac{B\left(\frac{x+5}{4}, \frac{1}{2}\right)}{B\left(\frac{x+1}{4}, \frac{1}{2}\right)}$$

can be rearranged as $X(x) \cdot X(x+2) \equiv 1$, true for all x in some interval, or $X(x) \equiv 1$. Wherefore, any continued fraction expansion which takes on the values of $(x+1) \times \frac{B\left(\frac{x+3}{2}, \frac{1}{2}\right)}{B\left(\frac{x+1}{2}, \frac{1}{2}\right)}$ over some interval, $x \in [a, b]$ say, will satisfy the functional equation $\Phi(x)\;\Phi(x+2) = (x+1)^2$, $x \in [a, b]$. (The BROUNCKER expansion satisfies it for $0 \leqq x \leqq \infty$, with $\Phi(x) = -\Phi(-x)$.)

III. Concept of function

1. The logarithm as a type-function

The general idea of a function arose gradually over many years and through many increasingly abstract stages. Defined generally as a mapping $f(x, y)$: $x \rightarrow y$ of one variable, x, into a second, y, it is a product of the early 19th century effort to place the concepts of analysis on a rigorous basis: a stage which could be reached only after long familiarity with particular functions in the attempt to synthesize generally applicable methods and techniques. A previous stage, when a mass of particular functions but few standard methods were known, had been reached in the late 18th century (through the diligence of such mathematicians as EULER, LAGRANGE, the BERNOULLIS and JACOBI), but in the 17th century even particular functions known were few, and general methods were largely restricted to what was obvious treatment of the geometrical models in which they were widely used—notably areas and arc-lengths of the various species of conics—: an approach which, with all its advantages of immediacy and tangibility, was hardly conducive to the development of abstract treatment

For that reason a comprehensive general account, while possible, seems not very worthwhile, and it seems preferable to sketch in the complexities of 17th century functional treatments with regard to a particular function, seeing the difficulties faced and overcome by the evolving concept as in many ways typical. Such an approach, while bringing a considerable amount of cohesion to what must, in historical fact, inevitably be a collection of scattered aspects, however firmly linked, must depend for its value on the particular function chosen for study. Fortunately, in the later 17th century the logarithm is an almost automatic choice: an important and basic analytical function given a wide variety of treatment and interpretation in the period, both abstractly as a correspondence and geometrically as hyperbola-area (in which form it ties in closely with the trigonometrical functions themselves defined on the geometrical model of the circle or general ellipse*). As such, an understanding of its ramifications and varieties of form are essential to a full comprehension of 17th century mathematics and its limitations, and those aspects—notably, general series-expansions[1]—which are not treated in detail elsewhere will be discussed here approximately in chronological sequence.

Historically, the logarithmic function developed[2] as the attempt to render precise and to evaluate numerically the correspondence which exists between two sets of numbers, one increasing (or decreasing) in an arithmetical ratio, $\lambda + k\mu$ while the other increases in a geometrical ratio $L \times M^k$, where k varies

---

* Such a dual definition, analytical and geometrical, was typical of the 17th century, and it is important to notice that each aspect reinforced the other both conceptually and as a matter of practical technique. While such things as series-expansions (in the case of the logarithm and trigonometrical functions) and periodicity (restricted at first to the trigonometrical functions) are better dealt with analytically, others —especially the interrelationship of logarithm and trigonometrical function—are more naturally treated on the geometrical model.

[1] To be developed at length in the following chapters.

[2] Around the beginning of the 16th century. A detailed modern account with full references—which I will not try to duplicate—is given in TROPFKE$_3$ 2 (1933): Section E (204—262): *Die Logarithmen*, especially 207 ff.

among some integer set, $-r, -(r-1), \ldots, -1, 0, 1, 2, \ldots, (s-1), s$ say, in the first instance, and later (by natural extension) as a full real variable in the continuum $[-\infty, +\infty]$. Clearly the functional mapping $f(\alpha_k, A_k)$: $a_k \rightarrow A_k$, where $a_k = \lambda + k\mu$, $A_k = L \times M^k$, is given by $a_k = \log(A_k)$ if $\lambda = \log(L)$ and $\mu = \log(M)$.* When the concept of ratio had become widely understood in the late medieval period, such correspondences were used in the attempt to interpret natural phenomena on a mathematical basis[3] and especially to formulate a satisfactory law of resisted motion[4]. Typically the medieval approach was to

* The particular case $\lambda + \mu = \log(L \times M)$ seems to have been the overriding reason for the late 16th and early 17th century attempts at extensive tabulation of the logarithm.

[3] For example, historically one of the oldest such correspondences is the concept of speed, whose origins go back beyond exact record. Specifically, this is the correspondence between the two linear continua of space traversed by a moving body and time taken to traverse that space conventionally given by the numerical ratio: $\frac{\text{distance}}{\text{time}}$, where the time and distance are measured in suitable units—a ratio which, if taken in the inverse form of "inverse speed": $\frac{\text{time}}{\text{distance}}$ would have removed some of the difficulties which clogged medieval attempts to formulate the speed of a moving body as varying with time taken or, again, with distance traversed (but which, in the form, speed $= \frac{\text{distance}}{\text{time}}$ was seen as a "natural" definition to be upheld at all cost). (Compare note 4 below.) The derived motion of instantaneous speed, the limit-form where the distance—and time—intervals shrink to zero (a contribution, apparently, of the 14th century Merton School at Oxford) became increasingly mathematically valuable and is, indeed, the model on which NAPIER develops his theory of the logarithm. It is an unanswered (if answerable) question as to how far, if at all, the use made by the medieval philosophers of the idea of instantaneous speed in developing theoretical problems on motion which use a law of motion which is logarithmical in form—though defined by them only as a correspondence between two number sets varying in a simple way—influenced the early modern theories of the logarithm, notably NAPIER's (see note 4 and compare J.E. HOFMANN: *Geschichte der Mathematik* **1** (1953): 135).

[4] Especially in the critical studies of the early 14th century Merton School at Oxford whose influence was to be passed on through the late 14th century Paris school (of such scholastics as BURIDAN and ORESME) and early 15th century Spain and Italy to the Renaissance. Dissatisfied with the inconsistent (if at all exactly formulable) ARISTOTELIAN law of resisted motion—which we may take perhaps as speed $(V) \propto \frac{\text{motive power } (M)}{\text{resistance } (R)}$ —THOMAS BRADWARDINE had proposed a variant form which seemed better to correspond with physical fact: "The proportion of the speeds in motion follows the proportion of the proportions of the motive power to resistance", or, setting up a table of correspondences, $\lambda = 1, 2, 3, \ldots$, if $V_1 \leftrightarrow \frac{M_1}{R_1}$, then $\lambda V_1 \leftrightarrow \left(\frac{M_1}{R_1}\right)^\lambda$ [*cf.* H.L. CROSBY: *Thomas of Bradwardine: his 'tractatus de proportionibus ...'*, Madison (Wisconsin), 1955, 12ff.]. Here, we are not concerned with the efficacy of this as a law of nature—for that see ANNELIESE MAIER: *Die Vorläufer Galileis im 14. Jahrhundert*, Rome 1949, and its excellent review by KOYRÉ [Archives Internationales de l'Hist. des Sc. **4** (1951): 769ff.]—but many recent treatments uncritically state the law in its modern form $V \propto \log(M/R)$. Such an exact functional correspondence is found in no text before the 17th century, and completely distorts a function-form which was seen as exponential only in the vaguest way. So it is with all known scholastic and scholastic-influenced treatments: RICHARD SWINESHEAD [*cf.* the 14th century *liber calculationum* (printed) Venice, 1520: especially tract 11: *de loco*

tabulate instances on some numerical basis of two covarying phenomena, and it is natural to suppose—on a basic principle of the simplicity of nature—that the connection will manifest itself in an obvious way when we compare corresponding instances. Likewise in the 16th century we find the same method of tabulation of instances used in exploring the relations between numbers and power-indices[5]—a case which corresponds to $L=1$ $(\log(L)=\lambda=0)$ above. Using the decimal number base, it is natural that M be taken as 10 and $\mu=1$ (which defines the logarithmic base also to be 10), and we then have the type-example of the correspondence studied:

x	$\cdots -3$	-2	-1	0	1	2	$3\cdots$
10^x	$\cdots 10^{-3}$	10^{-2}	10^{-1}	1	10	10^2	$10^3\cdots$

where x is restricted to being integral. Clearly, the importance of this at a practical level is that, for $x\leftrightarrow 10^x$, $y\leftrightarrow 10^y$, then $x+y\leftrightarrow 10^{x+y}=10^x\times 10^y$, and that the correspondence allows us to replace the operation of multiplication by that of addition—a cherished ideal when there were no automatic computing techniques at more than the most elementary level. If then the (integral) values of x could so be interpolated that two values x_1, x_2 could be assigned with reasonable accuracy corresponding to any numbers X_1, X_2 which have to be multiplied, and such that a third number X be found corresponding to (x_1+x_2), then $X=X_1\cdot X_2$, and the problem is solved. The obvious way was to set up a "logarithmic canon" of corresponding values of x and 10^x—it is immediate that only values of x in the interval $[1, 10]$ need be tabulated, since $10^{k+x}=10^k\times 10^x$—but it was far from clear how this was to be done systematically, and indeed no general approach appeared till independent[6] methods were created by NAPIER and BÜRGI at the close of the century.

BÜRGI's development is by far the simpler[7], merely giving an extremely large number of values of $(1.0001)^k$, $k=1, 2, 3, \ldots$.* NAPIER's ideas[8] show the signs

---

elementi: 36vb—38ra], WILLIAM HEYTESBURY [*cf.* CURTIS WILSON: *William Heytesbury. Medieval logic and the rise of mathematical physics*, Madison (Wisconsin), 1956: *passim*] and DESCARTES (with regard to the law of motion communicated to MERSENNE in the letter of 13 November 1629 ·≡· *OE* ADAM & MILHAUD **1**: 83—88). The blunt fact is that, with the possible exception of DESCARTES, none had sufficient mathematical technique at his disposal further to define the correspondence.

* In the general scheme $\lambda=0$, $L=1$; $\mu=1$, $M=1.0001$. An obvious disadvantage is that BÜRGI's development discards the decimal logarithmic base.

[5] Beginning with isolated examples in CHUQUET's *Triparty* (1484) and PACIOLI's *summa* (1494) it quickly became obligatory to consider the correspondence in algebra texts (and remained so till the close of the 16th century), but an outstanding treatment was given by STIFEL in his *arithmetica integra* (1544): Bk. 3: 250*L*: 102. Compare TROPFKE, *op. cit:* **2**; 206ff., and an article by D. E. SMITH, *The law of exponents in in the works of the 16th century*. NAPIER *TV:* 81—91.

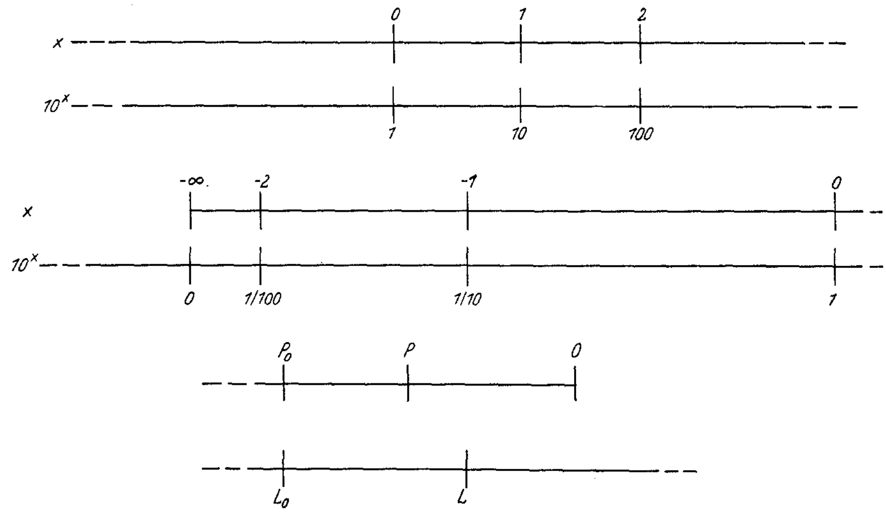
[6] Both NAPIER and BÜRGI seem to have begun their calculations about 1590 (see NAPIER *TV:* especially Lord MOULTON: *The invention of logarithms, its genesis and growth:* 1—32; E. W. HOBSON: *John Napier and the invention of logarithms*, London, 1914; and E. VOELLMY: *Jost Bürgi und die Logarithmen*, Basel, 1948, who suggests on manuscript evidence that priority is to be given to BÜRGI).

[7] Published in his *Arithmetische und geometrische Progress-tabulen*, Prag 1620.

[8] Given in the *constructio* (1617) when his logarithmic canon was already in print (in the *descriptio*, 1614).

of deeper and more imaginative thought, though left tantalisingly vague in his description of his ideas—a state of affairs which has led to several reconstructions of his process of thought[9], more or less inadequate.

It is indisputable, however, that NAPIER bases his development on a geometrical model in which he conceives two correlated points moving on separate line-segments such that one traverses segments in arithmetical progression while the other traverses corresponding segments which are in geometrical progression. Returning again to the popular 16th century correspondence between the integers and index powers, let us set it up on a model. Two simple forms are possible, the first of which—one naturally suggested by the physical layout of the cor-



respondence on the printed page in the typical 16th century treatment—maps the function x onto a line-length simply calibrated, and the function 10^x on to a second line to correspond; while the second has the function 10^x mapped onto the simply calibrated line, and the function x onto a second line to correspond. Taken together the two forms become powerfully suggestive[10], and would seem the root source of NAPIER's basic ideas. In fact, NAPIER's treatment is defined on a geometrical model slightly adapted from the second—probably for computational convenience[11]—and introducing an independent continuum of time.

Consider two points P and L moving one on each of two lines. The point P moves towards a point 0 on its line at a speed which varies directly as its distance away from it[12], while the corresponding point L moves uniformly along its line with the same speed as that which P has instantaneously at P_0. Then if point P

[9] Particularly in Lord MOULTON's essay (*op. cit.* note 6), whose derivation seems much too artificial and far-fetched in comparison with the reconstruction given.

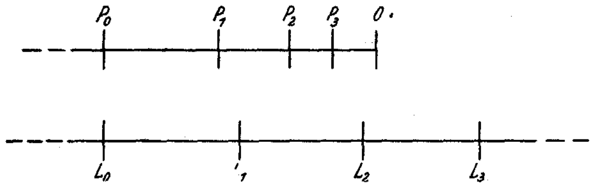
[10] In the first case the point correspondence is $x \leftrightarrow 10^n$: in the second the equivalent $\log_{10}(x) \leftrightarrow x$.

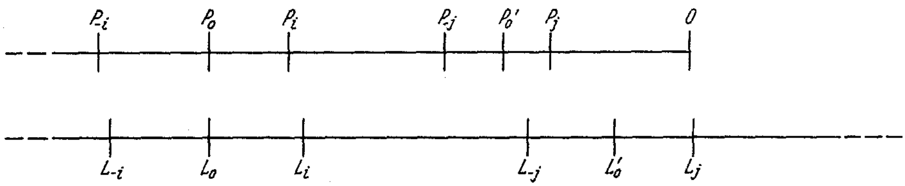
[11] MOULTON suggestively argues that the peculiar form arises from NAPIER's aim that his canon shall ease trigonometrical computation.

[12] Curiously—but only coincidentally—SWINESHEAD's law of motion in his *liber calculationum* (see note 4). It is interesting to notice that SWINESHEAD argued that the point P could not reach O in finite time if the starting speed was finite.

is at P_i and point L at L_i in the same moment of time, NAPIER defines the segment L_0L_i to be the *logarithmus* (ratio-number) of the segment P_iO.[13] Clearly, apart from introducing a time-continuum (whose function mainly is to add plausibility and emphasise certain obvious but non-trivial details, especially that P and L are at unique points P_i, L_i at the same moment, or that the correspondence $P \leftrightarrow L$ is 1, 1), the major modification made by NAPIER on the second model above has been to reverse the sense of the upper line.

It is easy to show the proposition, necessary and sufficient for the correspondence to be logarithmic, that the segments P_iO increase in (negative) geometrical proportion—specifically that, where the corresponding "logarithm" segments





are equal, $P_0P_1 : P_1P_2 : \ldots : P_iP_{i+1} : \ldots = P_0O : P_1O ; \ldots : P_iO ; \ldots$. It is immediate that $P_iO : P_{i+j}O = P_iP_{i+1} : P_{i+j}P_{i+j+1} = P_0P_1 : P_jP_{j+1} = P_0O : P_jO$, or $P_iO \times P_jO = P_{i+j}O \times P_0O$; or, where $L_0L_i \leftrightarrow P_iO$, $L_0L_j \leftrightarrow P_jO$, then $L_0L_i + L_0L_j\, [= L_iL_{i+j}] = L_0L_{i+j} \leftrightarrow P_{i+j}O$; so that, where similarly $L_0L_k \leftrightarrow P_kO$ and $L_0L_l \leftrightarrow P_lO$, $L_0L_i + L_0L_j = L_0L_k + L_0L_l$ (or equivalently $i+j = k+l$) implies $P_{i+j}O = P_{k+l}O$, or $P_iO \times P_jO = P_kO \times P_lO$ (which mirrors the fundamental logarithmic mapping of multiplication onto addition).

The problem remains of applying this structure, and NAPIER bases his numerical treatment on a general inequality derived verbally but which is clarified by being given symbolically.[14] Let us suppose that P traverses each of the intervals $P_{-i}P_0$, P_0P_i, $P_{-j}P_0'$, $P_0'P_j$ in equal intervals of time (measured by $L_iL_0 = L_0L_i = L_{-j}L_0' = L_0'L_j$): then $P_{-i}P_0 : P_0P_i = P_{-j}P_0' : P_0'P_j = P_0O : P_iO = P_0'O : P_jO$.

Clearly, since the speed of the point P continuously decreases as it moves towards O, its speeds at P_{-j}, P_i will be greater and less respectively than that at P_0 (where $P_{-i}O > P_0O > P_iO$), so that

$$\frac{P_0'P_j}{P_jO}\left[= \frac{P_{-i}P_0}{P_0O}\right] > \left[\frac{L_{-i}L_0}{P_0O} = \right] \frac{L_0L_j - L_0L_0'}{P_0O} \left[= \frac{L_0L_i}{P_0O}\right] > \left[\frac{P_0P_i}{P_0O} = \right] \frac{P_0'P_j}{P_0'O};$$

[13] *descriptio:* def. 6 (and compare *constructio:* 5ff.). From what is shown below it follows that, where $L_\alpha L_\beta = L_\gamma L_\delta$ (or $\alpha - \beta = \gamma - \delta$), $P_\alpha O : P_\beta O = P_\gamma O : P_\delta O$; so that $L_\alpha L_\beta$ is a "measure of the ratio" $P_\alpha O : P_\beta O$. This concept of a *mensura rationis* is fundamental in many 17th century analytical treatments of the logarithm (and very possibly underlies NAPIER's choice of the word *logarithmus*).

[14] *constructio:* 8ff.

or, taking $P_j O = x_j$, $L_0 L_j = L_N(x_j)$,

$$\left.\begin{array}{l} x_j < x_0' \\ L_N(x_j) > L_N(x_0') \end{array}\right\} \text{ defines the inequality } \frac{x_0' - x_j}{x_j} > \frac{L_N(x_j) - L_N(x_0')}{x_0} > \frac{x_0' - x_j}{x_0'}.$$

In particular, when $x_0 = x_0'$, $L_N(x_0') = L_N(x_0) = 0$; or $\frac{x_0 - x_j}{x_j} > \frac{L_N(x_j)}{x_0} > \frac{x_0 - x_j}{x_0}$, $(x_j < x_0)$.*

The way is now clear to construction of the numerical canon. NAPIER takes $x_0 = P_0 O = 10^7$ and in a series of tables calculates, first, $10^7\left(1 - \frac{1}{10^7}\right)^r$, $r = 0, 1, 2, \ldots, 100$[15] and so finds $10^7\left(1 - \frac{1}{10^7}\right)^{100} = 9999900.0004950$; next, the 51 numbers $10^7\left(1 - \frac{1}{10^5}\right)^s$, $s = 0, 1, 2, \ldots, 50$; and, finally, the 21×69 numbers $10^7 \times \left(1 - \frac{5}{10^4}\right)^p \cdot \left(1 - \frac{1}{10^2}\right)^q$, $p = 0, 1, 2, \ldots, 20$; $q = 0, 1, 2, \ldots, 68$, finding that $10^7 \times \left(1 - \frac{5}{10^4}\right)^{20} \cdot \left(1 - \frac{1}{10^2}\right)^{68}$ is a little less than $\frac{1}{2} \times 10^7$.[16] Using his inequality NAPIER derives bounds for all this dense set of numbers—or at least of an adequate number, according as the circumstances justified[17]—and finds that, by taking the arithmetic mean of the two bounds an accuracy of 7 significant figures is to be had. So he completes his logarithmic canon for $x \in [\frac{1}{2} \times 10^7, 10^7]$ and by straightforward extension to the remaining interval $x \in [0, \frac{1}{2} \times 10^7]$, and the whole canon is adapted to trigonometrical computation by changing the argument from natural instances to tabulated instances of $10^7 \cdot \sin\vartheta$, ϑ taken at 1′ intervals, $0 \leqq \vartheta \leqq 90°$.[18]

While the numerical aspect of logarithmic computation is not devoid of theoretical interest[19], it is the structure on which such numerical calculations are made which is significant in the concept of a logarithmic function.

---

* These inequalities correspond to the more familiar ones of natural logarithms: $a > b$ implies $\frac{a-b}{b} > \frac{\log(a) - \log(b)}{1} > \frac{a-b}{a}$. We cannot, of course—since $L_N(1)$ is not zero—suppose $L_N(\alpha) - L_N(\beta)\left[= L_N\left(\frac{\alpha}{\beta}\right) - L_N(1)\right]$ the same as $L_N\left(\frac{\alpha}{\beta}\right)$.

15 A computation which can be made simply by successive subtraction:

$$10^7\left(1 - \frac{1}{10^7}\right)^{r+1} = 10^7\left[\left(1 - \frac{1}{10^7}\right)^r \quad \frac{1}{10^7}\left(1 - \frac{1}{10^7}\right)^r\right].$$

16 The object, clearly, is to find a large number of approximately geometrical means in the interval $[10^7, \frac{1}{2} \times 10^7]$ and so have a fairly dense point-set scattered over it: $10^7\left(1 - \frac{1}{10^7}\right)^{100} \approx 10^7\left(1 - \frac{1}{10^5}\right)$, for example.

17 Mostly he seems merely to have calculated bounds for the 21×69 numbers $10^7 \cdot \left(1 - \frac{5}{10^4}\right)^p \left(1 - \frac{1}{10^2}\right)^q$ which form his "radical table", and to have filled in the remaining numbers by linear interpolation.

18 The canon was a gigantic labour of love which took twenty years to compute and check. It is a tribute to the accuracy of NAPIER's work (and to that of BRIGGS, who carried through an even more stupendous programme of calculation for his AL) that, even with the improved techniques available, no essentially new recalculation was made for a century. BRIGGS' adaptation to a decimal base ("common logarithms") involved merely the subtraction and division of constants and a change of sign.

19 As will be seen in the next chapter, numerical approximation is important in the early stages of interpolation theories.

If we take up NAPIER's basic idea, the concept can, in fact, be made to yield more than was ever taken from it in the 17th century. Consider once more the upper line in the NAPIERIAN definition, and as before suppose that the point P moves so that its speed varies as its distance from O, where $P_0O = 10^7$. Now consider the two-dimensional space in which a Cartesian coordinate system is defined by $P_iO = x$ and where $P_iQ_i = y$, normal to P_iO, measures the "inverse" instantaneous speed ($=1$/speed) of P at P_i. By loose limit considerations, the law of motion of P demands that "distance" $P_iO \times 1$/"speed" $P_iQ_i =$ "time" $=$ constant,* or $xy = P_iO \times P_iQ_i = P_0O \times P_0Q_0 = 10^7$, since NAPIER defines the instantaneous speed of P at P_0 to be unit speed—which shows the point-set of the Q_i to be a rectangular hyperbola of centre O and one asymptote P_0O. Further, the hyperbola-area $\text{Hyp}(P_0P_iQ_iQ_0)$ gives the total time taken by P to traverse the segment P_0P_i**; and so, in the above notation,

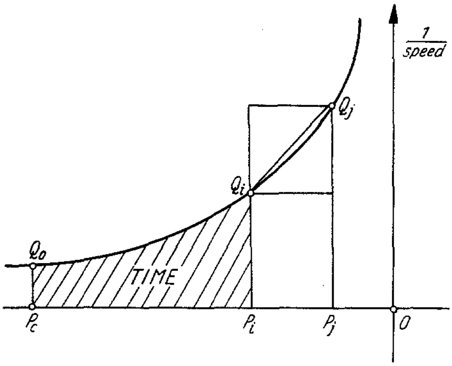


Fig. 8

$$L_N(x_i = P_iO) = \text{Hyp}(P_0P_iQ_iQ_0) = \int_{x_i=P_iO}^{x_0=10^7} y\left(=\frac{10^7}{x}\right)\cdot dx = 10^7 \log\left(\frac{10^7}{x_i}\right),$$

the known relation connecting the NAPIERIAN logarithm $L_N(x_i)$ and the natural logarithm $\log(x_i)$. Finally, if we consider two general points P_i, P_j and their corresponding Q_i, Q_j, the areal inequality (where $P_iO > P_jO$)

$$P_iP_j \times P_jQ_j = \text{rectangle } P_iQ_j > \text{Hyp}(P_iP_jQ_jQ_i) > \text{rectangle } P_jQ_i = P_iP_j \times P_iQ_i$$

proves

$$\frac{P_jQ_j \times P_iP_j}{P_0O} > \left[\frac{\text{Hyp}(P_iP_jQ_jQ_i)}{P_0O} = \right]\frac{\text{Hyp}(P_0P_jQ_jQ_0) - \text{Hyp}(P_0P_iQ_iQ_0)}{P_0O} > \frac{P_iQ_i \times P_iP_j}{P_0O},$$

or, where $x_i > x_j$, $\frac{x_i - x_j}{x_j} > \frac{L_N(x_j) - L_N(x_i)}{x_0} > \frac{x_i - x_j}{x}$, which is NAPIER's inequality. (We see, incidentally, how accurate is NAPIER'S final inspiration of taking the middle term as the arithmetic mean of the two bounds—on the model this is equivalent to equating the trapezium $P_iP_jQ_jQ_i$ with $\text{Hyp}(P_iP_jQ_jQ_i)$, slightly the smaller in fact.)[20]

Clearly, the use of hyperbola-area as a model of the logarithmic function is a richly suggestive idea, and one which, using an exhaustion proof, could fully

* The law has a *constant* time increment $dt = \left(\frac{dt}{ds}\right) \times ds$, where ds, dt are increments of distance and time respectively.

** Where t_i is the time taken by P over P_0P_i, $t_i = \int_{t=0}^{t=t_i} \left(\frac{dt}{ds}\right)\cdot ds = \int_0^{t_i} dt$.

[20] The whole argument, deliberately kept loose in keeping with NAPIER's own distance-speed model treatment, would have been understood by a 14th century scholastic, and indeed is medieval rather than modern, however attractive its rigorous treatment by calculus concepts.

be justified on mathematical techniques existing in the early 17^{th} century. Historically, however, we find a curious time-lag. Apparently the connection between the logarithm (and its basic property, $\log(\alpha)+\log(\beta)=\log(\alpha\times\beta)+\log(1)$) and the hyperbola-area was first to be noticed only half a century after NAPIER'S work by the relatively obscure Belgian Jesuit A.A. DE SARASA[21] reading through the *opus geometricum*[22] of his friend GREGORY ST. VINCENT (in whom a general viewpoint seemed to have been obscured by his love of detail). In fact, we can find in the *opus geometricum* everything except a statement of the logarithmic nature of hyperbola area: specifically, GREGORY proves that, where the points D, P, H, K are on a rectangular hyperbola of asymptotes AF, AC, then—if, say $DE:PQ=(\lambda:\mu)^m$ and $HI:KC=(\lambda:\mu)^n$ — $\mathrm{Hyp}(EQPD):\mathrm{Hyp}(ICKH)=m:n$.*

GREGORY'S proof[23] reduces the problem to the case $m, n=1$ by dividing EQ, IC in m, n segments respectively in geometrical progression of ratio $(\lambda:\mu)$ (decreasing from A): specifically, if $DE:PQ\,(=AQ:AE)=HI:KC\,(=AC:AI)$ defines the hyperbola $DPHK$ such that, for any ordinates $\lambda_i\mu_i$, L_iM_i, $A\lambda_i\times\lambda_i\mu_i=AL_i\times L_iM_i\,(=K^2$ constant), then $\mathrm{Hyp}(EQPD)=\mathrm{Hyp}(ICKH)$. The demonstration is carried through by an exhaustion method**. The general proof, however, is incomplete in that no freedom

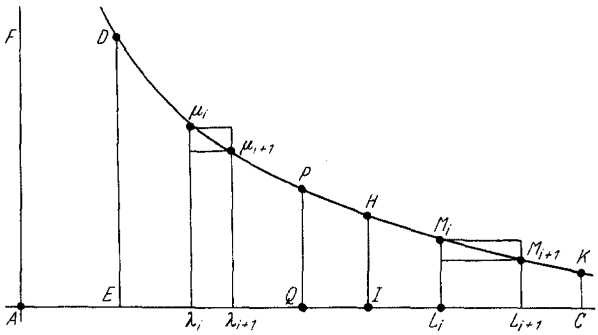


Fig. 9

* "... superficiem $DEQP$ toties continere superficiem $HICK$ quoties ratio lineae DE and PQ multiplicat rationem HI and KC."

** Where $AE<A\lambda_i<A\lambda_{i+1}<AQ$ orders the points of the segment EQ, we can set up a corresponding ordering $AI<AL_i<AL_{i+1}<AC$ of IC by $A\lambda_i:\lambda_i\lambda_{i+1}=AL_i:L_iL_{i+1}$, and show that

$$\lambda_i\lambda_{i+1}\times\begin{cases}\lambda_i\mu_i\\ \lambda_{i+1}\mu_{i+1}\end{cases}=L_iL_{i+1}\times\begin{cases}L_iM_i\\ L_{i+1}M_{i+1}\end{cases}$$

Finally, using the inequalities $\begin{cases}A\lambda_i<A\lambda_{i+1}\\ AL_i<AL_{i+1}\end{cases}$ or the equivalent $\begin{cases}\lambda_i\mu_i>\lambda_{i+1}\mu_{i+1}\\ L_iM_i>L_{i+1}M_{i+1}\end{cases}$, we have the Archimedean exhaustion scheme

$$(i)\begin{pmatrix}A_i=\lambda_i\lambda_{i+1}\times\lambda_i\mu_i>\mathrm{Hyp}(\lambda_i\lambda_{i+1}\mu_{i+1}\mu_i)>\lambda_i\lambda_{i+1}\times\lambda_{i+1}\mu_{i+1}=a_i\\ B_i=L_iL_{i+1}\times L_iM_i>\mathrm{Hyp}(L_iL_{i+1}M_{i+1}M_i)>L_iL_{i+1}\times L_{i+1}M_{i+1}=b_i\end{pmatrix},$$

where $A_i=B_i$, $a_i=b_i$—which proves $\mathrm{Hyp}(\lambda_i\lambda_{i+1}\mu_{i+1}\mu_i)=\mathrm{Hyp}(L_iL_{i+1}M_{i+1}M_i)$, and this is true for each pair of segments $\lambda_i\lambda_{i+1}$, L_iL_{i+1}.

[21] In an "appendix" to the *opus geometricum* (see note [22]) published shortly afterwards, *solutio problematis a ... Mersenno propositi: datis tribus quibuscunque magnitudinibus, rationalibus vel irrationalibus, datisque duarum ex illis logarithmis, tertiae logarithmum geometrice invenire.*

[22] GREGORY ST. VINCENT *OG:* Antwerp, 1647. See J.E. HOFMANN: *Das Opus Geometricum des Gregorius a S. Vincentio und seine Einwirkung auf Leibniz,* Abh. der Preuß. Akad. der Wiss., 1941, No. 13. Berlin 1942.

[23] *OG:* Bk. 6: *de hyperbola,* pt. 4: *de segmentis hyperbolicis convexis et concavis:* 583—603; especially prop. 125: 594.

is allowed for extension beyond rational values of the ratio $m:n$ (m, n are restricted to being integers, and GREGORY's GP-section divides the hyp-areas into respectively m, n equal hyp-areas), though Eudoxian schemes are easy to apply.

While GREGORY's *opus geometricum* attracted wide notice for another reason[24], mathematicians were at first slow to extend the hyperbola-area model of the logarithm[25]. Perhaps the correspondence seemed merely to convert a difficult analytical concept into an equally difficult one of hyperbola-area. In particular, how could the hyperbola-area be calculated for a suitable range of values of the asymptote—the very basis for setting up an improved logarithmic canon? It is this question, defined and solved with increasing precision from the early 1650's, which finally provoked the elementary infinite sum-series developments for the logarithm in the late 1660's.[26]

Perhaps the first attempt to calculate hyperbola-areas systematically was formulated by BROUNCKER in the mid-1650's. JOHN WALLIS, having had some success in finding approximations to circle-area using the (CARTESIAN) representing equation $y=\sqrt{R^2-x^2}$, had tried (in his *AI*) to apply the same techniques to the hyperbola, $y=\sqrt{R^2+x^2}$, and, failing, suggested the problem to BROUNCKER.[27] BROUNCKER succeeded in dissecting hyperbola-areas systematically, apparently in mid-1655, but did not publish his method for a decade.[28]

The BROUNCKERIAN approach typifies the solid, common-sense attitude to mathematical difficulties which so often—contrary to myth—yields a workable solution*. When confronted by some area whose numerical measure in terms of unit-area we wish to find, we naturally narrow approximation error by suitably splitting the area. So BROUNCKER, faced with the hyperbola-area $ABCE$, where $O\lambda$ is an asymptote and general point μ on the (rectangular) hyperbola λEC is defined by $O\lambda\times\lambda\mu=K^2$, begins by repeated bisection of the base-line AB such that at some λ^{th} stage the points $a', b', c', \ldots$ dissect the interval AB into 2^λ equal intervals $Aa'=a'b'=b'c'=\cdots=g'B$; and then considers two distinct ways of approximation. First, we can see hyp-area $(ABCE)$ as the limit of the sum sequence of inscribed rectangles (denoted as in the figure): $\square ABCF+\square KFNd+\square MNPb+\square HKLf+\cdots$; or secondly, we can take it as the limit

* Though I do not deny that outstanding advance has taken place on the basis of a flash of insight or a clarifying redefinition of the problem.

24 His illusory proof that circle quadrature is impossible—*cf.* ch. 1, note 25.

25 Though both HUYGENS and NEWTON realized its full significance at an early point in their mathematical development and use the logarithmic function in full generality in geometrical schemes. *Cf.* HUYGENS *OE* **12** 1910): 234ff., in which with a "*ἐύρηκα*, 27 October 1657" he reduces the rectification of the parabola to a suitable hyperbola-area; and CUL. Add 4004: (to be dated early 1665) where NEWTON notes: "In y^e Hyperbola y^e area of it beares y^e same respect to its asymptote w^{ch} a logarithme doth [to its] number."

26 Compare the next chapters.

27 See WALLIS: *adversus M. Meibomii de proportionibus dialogum: dedicatio* ·≡· *operum mathematicorum pars prima,* Oxford 1657: *dedicatio,* iii, ·≡· *opera* **1** (1695): 231—232.

28 In *PT* **3** (1668): 645—649: *The squaring of the hyperbola by an infinite series of rational numbers ....*

of the (negative) sum sequence of inscribed triangles: $\square ABDE - [\triangle CED + \triangle CdE + (\triangle dbE + \triangle Cfd) + \cdots]$.* In BROUNCKER's example, $K^2 = 1$, $= \vartheta\lambda \times \lambda\mu$, and $OA = AE = AB$ (so that hyp-area $(ABCE) = \log 2$). Using the first approach,

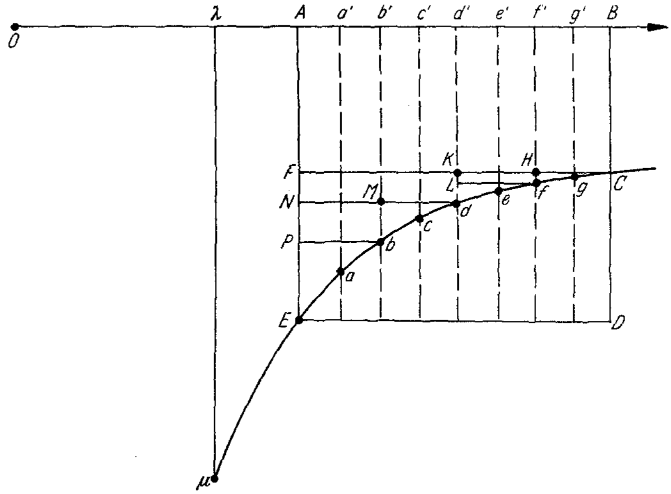


Fig. 10

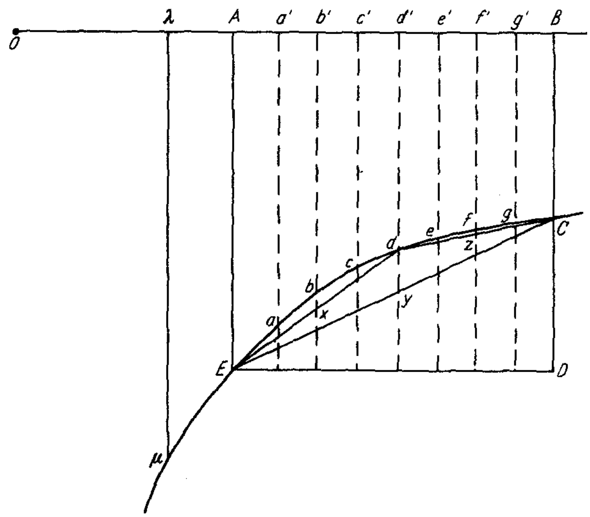


Fig. 11

we find, since $AE = 1$, $aa' = \frac{8}{9}$, $bb' = \frac{8}{10}, \ldots, gg' = \frac{8}{15}$, $BC = \frac{1}{2}$, that $\square ABCF = \frac{1}{1\cdot 2}$, $\square KFNd$ $(= \square d'ANd - \square d'AFk) = \frac{2}{3\cdot 4} \times \frac{1}{2} = \frac{1}{3\cdot 4}$, $MNPb = \frac{1}{5\cdot 6}$, $\square HKLf = \frac{1}{7\cdot 8}, \cdots$; and so the general law of formation is clear to the eye. That is,

$$\text{hyp-area}\,(ABCE)\ (= \log 2) = \frac{1}{1\cdot 2} + \frac{1}{3\cdot 4} + \frac{1}{5\cdot 6} + \frac{1}{7\cdot 8} + \cdots.^{29}$$

* A basic assumption made is, of course, that the hyperbola μEC is everywhere convex (except at points at infinity, but these do not trouble in the present case).

[29] This is, of course, the "MERCATOR" expansion of log 2.

By the second approach, we find $\triangle CDE = \frac{1}{2^2}$, $\triangle CdE = \frac{1}{2\cdot 3\cdot 4}$, $\triangle dbE = \frac{1}{4\cdot 5\cdot 6}$, $\triangle Cfd = \frac{1}{6\cdot 7\cdot 8}, \ldots$, and so, in this case,

$$\text{hyp-area}\,(ABCE) = 1 - \left(\frac{1}{2^2} + \frac{1}{2\cdot 3\cdot 4} + \frac{1}{4\cdot 5\cdot 6} + \frac{1}{6\cdot 7\cdot 8}\cdots\right).$$

Clearly, the method is general* but laborious—what makes the method appealing is that the complicated expressions reduce (as the particular case $x=1$ of the general series below) to more amenable shape. The same is true for a similar approach instituted by PIETRO MENGOLI[31], apparently some time in the mid-1650's also. MENGOLI'S method, in fact, yields the same series for log 2 as BROUNCKER'S approach by rectangles, but the interesting conceptual development arises by suitable definition and particularisation from a deliberate attempt to create an analytical theory of the logarithm, based on the model of hyperbola-area in inspiration but independent of it in form.**

MENGOLI begins with two basic (and complementary) concepts: the "hyperlogarithmus", $\overline{L}(m/n)_r$, and the "hypologarithmus" $\underline{L}(m/n)_r$, defined respectively by

$$\overline{L}\left(\frac{m}{n}\right)_r = \sum_{rn\leq\lambda\leq rm-1}\left(\frac{1}{\lambda}\right) \text{ and } \underline{L}\left(\frac{m}{n}\right)_r = \sum_{rn+1\leq\lambda\leq rm}\left(\frac{1}{\lambda}\right).$$

Clearly

$$\overline{L}\left(\frac{m}{n}\right)_r > \underline{L}\left(\frac{m}{n}\right)_r \text{ and } \lim_{r\to\infty}\left(\overline{L}\left(\frac{m}{n}\right)_r - \underline{L}\left(\frac{m}{n}\right)_r\right) = 0;$$

and we can show[32]

$$\underline{L}\left(\frac{m}{n}\right)_r < \underline{L}\left(\frac{m}{n}\right)_s, \quad \overline{L}\left(\frac{m}{n}\right)_r > \overline{L}\left(\frac{m}{n}\right)_s \qquad \text{for } r > s.$$

---

* BROUNCKER, indeed, sketches in the extension[30] where $OA = AE = 1$, $AB = x$; so that hyp-area $(ABCE) = \log(1+x)$ as a more complicated case of the above dual procedure. The two approaches, in fact, yield rather unwieldy series expansions for $\log(1+x)$, namely, where $\lambda_{r,s} = 2^r + 2sx$,

(by rectangles)

$$\frac{x}{1+x} + x^2 \times \lim_{n\to\infty}\left(\sum_{1\leq r\leq n}\ \sum_{1\leq s\leq 2^{r-1}}\left(\frac{1}{\lambda_{r,s}(\lambda_{r,s}-x)}\right)\right),$$

(by triangles)

$$x - \frac{1}{2}\,\frac{x^2}{1+x} - x^3 \times \lim_{n\to\infty}\left(\sum_{0\leq r\leq n}\ \sum_{1\leq s\leq 2^{r-1}}\frac{1}{\lambda_{r,s}(\lambda_{r,s}-x)(\lambda_{r,s}-2x)}\right).$$

** This abstraction of structure from geometrical form is MENGOLI'S professed ideal throughout *GS*. It is interesting to interpret the analytical discussion given here on the model of the hyperbola $xy = 1$.

30 BROUNCKER, *op. cit.* 349: "By any of which ... series it is not hard to calculate, as near as you please, these and the like hyperbolic spaces, whatever be the rational proportion of *AE* to *BC*."

31 In MENGOLI: *GS*; Bologna, 1659. The series expansion for the logarithm seems to have been introduced while the book was printing, in the lengthy introduction (*cf.* appendix: 73–75 "cum haec scriberem, mihi contigit rectum tramitem invenire ad persequendos omnium numerosarum rationum logarithmos") while the analytical theory of the logarithm is pursued at great length in Books 4, 5; compare A. AGOSTINI: *L'opera matematica di Pietro Mengoli*, Archives int. de l'hist. des sciences **3** (1950): 816–834.

32 A point proved not quite rigidly by MENGOLI.

Further

$$\underline{L}\left(\frac{m}{n}\right)_{rp} + \underline{L}\left(\frac{o}{p}\right)_{rm} = \underline{L}\left(\frac{m}{n}\right)_{ro} + \underline{L}\left(\frac{o}{p}\right)_{rn} = \underline{L}\left(\frac{m\cdot o}{n\cdot p}\right)_{r}.$$

and similarly

$$\overline{L}\left(\frac{m}{n}\right)_{rp} + \overline{L}\left(\frac{o}{p}\right)_{rm} = \overline{L}\left(\frac{m}{n}\right)_{ro} + \overline{L}\left(\frac{o}{p}\right)_{rn} = \overline{L}\left(\frac{m\cdot o}{n\cdot p}\right)_{r}.$$

Using this as his analytical basis MENGOLI defines $\log(m/n)$ as the limit of the two sum-sequences $\overline{L}$ and $\underline{L}$:

$\log\left(\frac{m}{n}\right)$ is the function which satisfies $\overline{L}\left(\frac{m}{n}\right)_r \geqq \log\left(\frac{m}{n}\right) \geqq \underline{L}\left(\frac{m}{n}\right)_r$, for all r.

By use of an analytical counterpart of the exhaustion-method (using, in fact, the same logical proof-form) the property which defines the logarithmic nature of $\log(m/n)$:

$$\log\left(\frac{m}{n}\right) + \log\left(\frac{o}{p}\right) = \log\left(\frac{m\cdot o}{n\cdot p}\right)$$

is easy to show.

Finally, define the "prologarithmus" $P(n)_r$ by

$$P(n)_r = \sum_{1\leqq s\leqq n}\left(\frac{1}{(r-1)\,n+s}\right),$$

and it follows immediately that

$$\sum_{1\leqq r\leqq R}(P(n)_r) = \sum_{1\leqq r\leqq R}\ \sum_{1\leqq s\leqq n}\left(\frac{1}{(r-1)\,n+s}\right) = \sum_{1\leqq\lambda\leqq Rn}\left(\frac{1}{\lambda}\right).$$

Then

$$\underline{L}\left(\frac{m}{n}\right)_R = \sum_{R-1\leqq\lambda\leqq Rm}\left(\frac{1}{\lambda}\right) = \underline{L}\left(\frac{m}{1}\right)_R - \underline{L}\left(\frac{n}{1}\right)_R = \sum_{1\leqq r\leqq R}(P(m)_r - P(n)_r)$$

$$= \sum_{1\leqq r\leqq R}\left(\sum_{1\leqq s\leqq m}\left(\frac{1}{(r-1)\,m+s}\right) - \sum_{1\leqq t\leqq n}\left(\frac{1}{(r-1)\,n+t}\right)\right).$$

Finally

$$\log\left(\frac{m}{n}\right) = \lim_{R\to\infty}\left(\underline{L}\left(\frac{m}{n}\right)_R\right) = \lim_{R\to\infty}\left(\sum_{1\leqq r\leqq R}(P(m)_r - P(n)_r)\right).\star$$

Both BROUNCKER'S and MENGOLI'S general expansions for the logarithmic function are, in practice, clumsy and unwieldy. No workable approximations, for example, to particular logarithms are forthcoming without a quite unjustified amount of work. Well into the 1660's it remained the ideal of many mathematicians to construct methods which, based on the model of hyperbola-area for their justification, would give a close approximation without undue computation. This problem was, of course, resolved with the aid of intègration techniques by

⋆ From which BROUNCKER'S "Mercator" series for log 2 follows by taking $m = 2$, $n = 1$:

$$\log\left(\frac{2}{1}\right) = \log 2 = \lim_{R\to\infty}\sum_{1\leqq r\leqq R}\left(\frac{1}{2(n-1)+1} + \frac{1}{2(r-1)+2} - \frac{1}{(r-1)+1}\right)$$

$$= \lim_{R\to\infty}\sum_{1\leqq r\leqq R}\left(\frac{1}{2r-1} - \frac{1}{2r}\right).$$

several sum-series expansions which were (or could be made) quickly converging[33], but a wonderfully ingenious and accurate approach had in the meantime been developed by JAMES GREGORY as a corollary to the well-known converging sequences which he abstracts from the geometrical model of a general sector of a central conic (ellipse or hyperbola).[34]

Let us take a general sector BPC of the conic whose centre is A, with the tangents at B, P meeting in F: it is immediate that AF, meeting BP in I, bisects BP, and that the tangent DIL is parallel to BP. With more difficulty we can show that the areas $(BAPI)=(GM)\,(BAPF, BAP)$, and $(ABDLP)=(HM)\,(BAPI, BAPF)$*; and now we see the beginning of two converging sequences (i_k), (I_k), in which $(ABP)=i_0$, $(ABFP)=I_0$; $(ABIP)=i_1$, $(ABDLP)=I_1$. In the case of the ellipse i_k is a (convex) area which has BA, BP for two sides, and the remaining (2^k) ones have their end-points in the ellipse arc BP; and I_k is a similar (convex) area of two sides BA, BP and whose remaining (2^k+1) sides are each tangent to the ellipse arc touching it in the end-points of sides of i_k. The case of the hyperbola is similar: we merely reverse the definitions of i_k, I_k.[35]

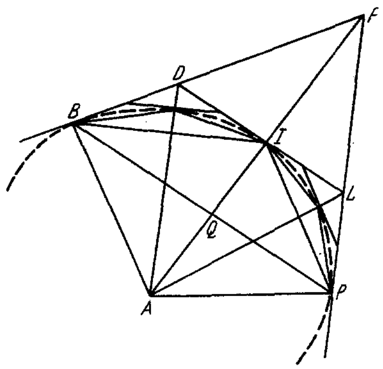


Fig. 12

We have, then, a "converging sequence" *(series convergens)* of (i_k), (I_k) which are generated by

$$i_{\lambda+1}=(GM)\,(i_\lambda, I_\lambda),$$
$$I_{\lambda+1}=(HM)\,(i_{\lambda+1}, I_\lambda),^{36}$$

and it is from this that GREGORY derives a subtle numerical technique. Thus, consider now the hyperbola ISL whose representing (CARTESIAN) equation is $xy=10^{25}$, and centre A, asymptotes AK, AO: The tangents at I, L meet in λ

* 1. By the pole-polar property $AI^2=AQ\cdot AF$, so that

$$\frac{(BAPI)}{(BAPF)}=\frac{AI}{AF}=\frac{AQ}{AI}=\frac{(BAP)}{(BAPI)}.$$

2. $$\frac{(BAPLD)}{(BAPF)}=\frac{(BAP)+\dfrac{QF^2-IF^2}{QF^2}(BPF)}{(BAPF)}=\frac{AQ\cdot QF+QF^2-IF^2}{AF\cdot QF}$$
$$=\frac{2AI(AR-AI)}{AF^2-AI^2}=\frac{(HM)\,(AI, AF)}{AF}=\frac{(HM)\,(BAPI, BAPF)}{(BAPF)}.$$

33 See chapter five. The first published account of the development was given by MERCATOR in *Logarithmotechnia:* especially prop 17: 31–33, though several people developed the method independently.

34 The method was developed apparently in postgraduate research at Padua in the mid-1660's, but first published in *VCHQ:* prop. 1 ff.

35 The i_k, I_k are, in GREGORY's terminology, "regularia inscripta", "... circumscripta". Clearly $\lim_{k\to\infty}(i_k)=\lim_{k\to\infty}(I_k)=$ conic sector ABP.

36 *Cf. VCHQ:* prop. 5: scholium, where GREGORY introduces parameters for this recursive procedure.

and $A\lambda N$ is drawn (bisecting the hyperbola chord IL). Taking $KA=LM=10^{13}$, $IK=AM=10^{12}$, and so $KM=OP=9\cdot 10^{12}$, so that if we can find Hyp-area $(KMLI)$, we shall have $10^{25}\times\log(10^{13}/10^{12})=10^{25}\times\log(10)$. However,

Hyp-area $(KMLI)$

$= \text{hyperbola} - \text{triangle } ALSI$

(since $\triangle AKI=\triangle ALM$ ★), and we can find this as the limit of the sequences i_k, I_k which begin with $i_0=\triangle AIL$, $I_0=$ area $(AL\lambda I)$. GREGORY evaluates $\triangle AIL$ by showing it to be equal to the trapezium $(LOPI)$ ★★, and proves area $(AL\lambda I)=(HM)\,((NOPI),(COPQ))$ ★★★, so that

$$i_0=\frac{99}{2}10^{24}\left(=OP\times\frac{1}{2}(LO+IP)\right),$$

$$I_0=(HM)\,(9\cdot 10^{24}, 9\cdot 10^{25}).$$ [37]

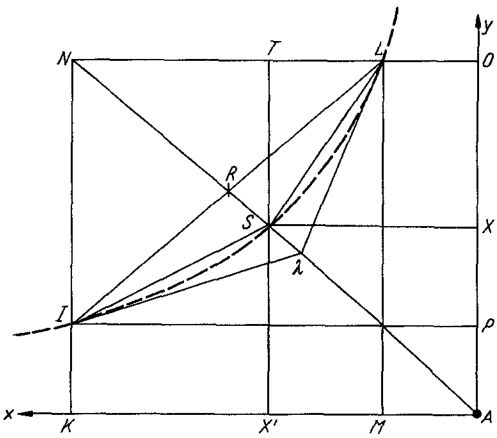


Fig. 13

To these GREGORY applies his formation rule 20 times, and has

$$i_{20}=23025\ \ 85092\ \ 99312\ \ 03593\ \ 18112\ \ 4$$
$$I_{20}=23025\ \ 85092\ \ 99589\ \ 61534\ \ 17386\ \ 4,$$

which he rounds off by a "triplicating" inequality[38], reaching finally Hyp-area $(KMLI)\,[=10^{25}\cdot\log(10)]=10^2\times 23025\ 85092\ 99404\ 56240\ 1787$, *proxime*. In further development he sketches in how the technique might be adapted to calculating $\mathrm{Log}(X)$ using the hyperbola $xy=10^{25}$ deriving by calculating $10^{25}\times\log(X)$ from given (close) values, $\log(X_1)$, $\log(X_2)$ where $\log(X_1)<\log(X)<\log(X_2)$.[39]

Such "brute-force" methods were rapidly superseded by simpler but—from a theoretical viewpoint—less subtle methods, and certainly with an increase in power there was a corresponding lack of rigour. However, the methods of geometrical approximation were, in effect, mere corollaries of the geometrical hyperbola-area model of the logarithmic function, and till an adequate analytical definition was developed—significantly, by abstracting from the geometrical

---

★ For it is the hyperbola property that $IK\times KA=LM\times MA$.

★★ Area $(AOLI)-\Delta AOL=$ area $(AOLI)-\Delta API$, with $AO\times OL=AP\times PI$.

★★★ 1. $(ALI):(ALSI)=(ALSI):(AL\lambda I)$

2. $(AM):(GM)=(GM):(HM)$

3. $\begin{cases}(ALI)=(AM)\,(NOPI, LOPQ)\ [=(LOPI)]\\(ALSI)=(GM)\,(NOPI, LOPQ), \text{ since } SX^2=SX\cdot SX'\\ \qquad =LO\times(OA=)\,ON.\end{cases}$

[37] *VCHQ*: props. 25–29.

[38] Probably that of *VCHQ*: prop. 24: *scholium*: sector $\approx\dfrac{8I_{k+1}+8i_{k+1}-i_k}{15}$. This apparently had been derived empirically by GREGORY at the time of writing *VCHQ*, but is stated in more exact form in *BG*: 11. Compare GREGORY-OLDENBURG, 25 December 1668 ·≡· HUYGENS *OE* 6; 309.

[39] Props. 30–32 (and conversely in props. 33, 34), *op. cit.*

model in the form $\log|x| = \int_1^x \frac{dx}{x}$ —all applications of the logarithmic function in mathematical analysis continued to be on a geometrical basis.

Thus we find that MERCATOR's publication of his sum-series treatment of $\log(1+x)$ in *logarithmotechnia* inspired WALLIS[40] to give an exact form of the equivalent of $\int_\beta^\alpha \log(x)\cdot dx$ (improving on MERCATOR's sum-series treatment[41]). Stated precisely, by a method equivalent to a change of order of integration in double integral his result is that $\int_b^1 \mathrm{Hyp}\,{}_b^x \cdot dx = \mathrm{Hyp}\,{}_b^1 - b^2(1-b)$, where $\mathrm{Hyp}\,{}_\beta^\alpha$ is the area under the hyperbola $xy=b^2$ between $x=\alpha$ and $x=\beta$ $\left(\text{or } \mathrm{Hyp}\,{}_\beta^\alpha = b^2\log\left(\frac{\alpha}{\beta}\right)^{\star}\right)$.

The method was, however, used most elegantly and powerfully in England by JAMES GREGORY[42] and ISAAC BARROW[43]. Outstanding in its beauty and ingenuity as well as its complexity is JAMES GREGORY's proof of the equivalent of $\int_0^\vartheta \sec x\cdot dx = -\log(\sec\vartheta - \tan\vartheta)$, important for its use in the theory of the MERCATOR projection (GREGORY's "nautical planisphere")[44]. Of this—in an equivalent form: $\int_0^\vartheta \sec x\cdot dx = \log\left(\frac{1+\sin\vartheta}{1-\sin\vartheta}\right)$, at least—BARROW gave a much shorter proof,[45] and its analysis will show the power of the geometrical model of the logarithm.

BARROW assumes a geometrical transform of the integral[46] which in effect, yields the equality $\int_0^\vartheta \sec x\cdot dx = \int_{x=0}^{x=\vartheta} \sec^2 x\cdot d(\sin x)$. This latter he sets up in geometrical form by considering a circle quadrant ABC (of centre C): taking any

---

⋆ Substituting this we have the modern form of WALLIS' result:

$$\int_b^1 \left(b^2\log\left(\frac{x}{b}\right)\right)\cdot dx = \left(b^2\log\left(\frac{1}{b}\right)\right) - b^2(1-b), \quad \text{or} \quad \int_b^1 \log\left(\frac{x}{b}\right)\cdot dx = \log\left(\frac{1}{b}\right) - (1-b).$$

40 In *PT* **3** (1668): no 38: 753—764, which reviews MERCATOR Log giving extracts from two letters of his to BROUNCKER of 8 July and 5 August 1668.

41 *Mercator* Log: prop. 19, where by simple integration of his sum-series MERCATOR gives $\int_0^x \log(x)\cdot dx = x\left(\frac{x}{2} - \frac{x^2}{2\cdot 3} + \frac{x^3}{3\cdot 4}\cdots\right)$, where the logarithmic function is defined on the hyperbola $xy=1$.

42 In *GPU* and *EG:* appendix especially.

43 In his *LG:* especially lectio 9ff.

44 *EG:* 14—21L *analogia inter lineam meridionalem planisphaerici nautici..., seu quod secantium naturalium additio efficiat tangentes artificiales*, especially props. 1, 2: 14—17.

45 BARROW, *LG:* lectio **12**, appendix: 5—6: 111. As will be seen in chapter 6 of this work, many years later HALLEY gave a further ingenious proof that the stereographic projection of a loxodrome on a sphere is a logarithmic spiral. *Cf.* *PT* **19** (1695): No. 215.

46 Elaborated in *EG: analogia ...*: prop. 1: 14—15.

parallel $T\lambda$ to AC (meeting as shown), we define ϱ on the curve XA by the meet of λT with $\varrho\sigma$, drawn parallel to BC through the meet σ of the tangent at τ and AC. Then, where $A\hat{C}T = x$,

$$2\lambda\varrho^2 (= 2BC^2 \sec^2 x) = 2BC^2 \left(\frac{\lambda\varrho}{BC}\right)^2$$
$$= 2TBC^2 \left(\frac{BC}{\lambda T}\right)^{2\,\star}$$

with

$$\lambda T^2 = CT^2 (= BC^2) - C\lambda^2$$
$$= (BC + C\lambda)(BC - C\lambda)$$
$$= B\mu \times B\lambda,$$

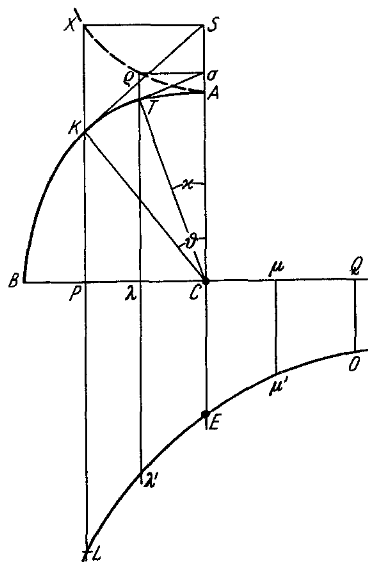


Fig. 14

if we take μ (on the further side of C) in BC such that $\lambda C = C\mu$. Further $2BC = B\lambda + B\mu$, so that

$$2 \cdot \lambda\varrho^2 = BC^3 \left(\frac{B\lambda + B\mu}{B\lambda \cdot B\mu}\right) = BC\left(\frac{BC^2}{B\lambda} + \frac{BC^2}{B\mu}\right),$$

and it is completely natural to introduce the rectangular hyperbola LEO (of centre B and asymptote BC) by: for all points λ' on it, $B\lambda \cdot \lambda\lambda' = BC^2$. We can then reduce further by: $2\lambda\varrho^2 = BC(\lambda\lambda' + \mu\mu')$, and finally, summing by the elements of BC** over $0 \leq x \leq \vartheta$, where $\vartheta = A\hat{C}K$ defines the maximum range of integration $PC = CQ = \sin\vartheta \times BC$, we have

$$\sum_{x=0 \leq x = x \leq x = \vartheta} 2BC^2 \sec^2 x \cdot d(BC \cdot \sin x) \cdot \left(= 2BC^3 \int_0^\vartheta \sec^2 x \cdot d(\sin x)\right)$$
$$= \sum_{x=0 \leq x = x \leq x = \vartheta} BC(\lambda\lambda' + \mu\mu') = BC \times \text{Hyp-area}\,(PQOL)$$
$$= BC^3 \log\left(\frac{BQ}{BP}\right) = BC^3 \log\left(\frac{1 + \sin\vartheta}{1 - \sin\vartheta}\right).$$

It remains, to complete discussion of 17th century attitudes towards the logarithm, to note that, in keeping with the increasing analytical tone of the late 17th century, attempts were made to give a fully analytical definition of the logarithm—specifically, it was required that this definition should lead naturally and immediately to the known sum-series expansions. In contrast with the flexibility of the modern definition $\log|x| = \int_1^x \frac{dx}{x}$—which still has the fossil-mark of the hyperbola on it—fluxional calculus, lacking a usable sign for the operation of integration, had to fall back on a definition which was largely verbal. Inevitably, too, such verbal definition was in some sense a return to the loosely expressed kinematical approach of NAPIER.

* Since $\lambda\varrho (= C\sigma) : BC (= CT) = \sec x : 1 = 1 : \cos x = CT (= BC) : \lambda T$.
** $d(BC \sin x) = BC\, d(\sin x)$.

So we find it with HALLEY's attempt at an analytical definition[47]. Following a strictly Napierian approach, HALLEY takes as his (verbal) definition of the logarithm of a number the fact that logarithms are "numbers which are the exponents of ratios" (numeri rationum exponentes), and considers some very small "ratiuncula" which shall be a unit-measure for logarithms. Then to measure the ratio of the logarithms of two line segments, α and β, he sets up in each a scale of continued proportionals of which this unit-ratiuncula is the first segment, so that, as the unit-ratiuncula is indefinitely decreased in magnitude, the ratio of the number of geometrical proportionals in each line will approximate ever more closely to the ratio of their logarithms. Thus, if $(1+\lambda)^a = A$ and $(1+\lambda)^b = B$, $\log_k A : \log_k B = \underset{\lambda \to \infty}{\text{limit}}\, (a:b)$.[48]

HALLEY now has an ingenious idea*: "... if, instead of supposing the logarithms composed of a number of equal ratiunculae proportional to each ratio, we shall take the ratio of unity to any number to consist always of the same infinite number of ratiunculae, their magnitude in this case will be as their number in the former; wherefore, if between unity and any number proposed there be taken any infinity of mean proportionals, the infinitely little segment or decrement of the first of those means from unity will be a ratiuncula; that is, the momentum or fluxion of the ratio of unity to the said number. And seeing that in these continued proportionals all the ratiunculae are equal, their sum, or the whole ratio, will be as the said momentum directly; that is, the logarithm of each ratio will be as the fluxion thereof. Wherefore, if the root of any infinite power be extracted out of any number, the differentiola of the said root from unity shall be as the logarithm of that number."

The verbal treatment obscures the basic concept—and the whole passage was not understood widely at the time because of such obscurity of what was at its clearest a difficult concept—but a symbolic sketch will point his meaning. Let the ratiunculae of the two line-segments $(1+\alpha)$, $(1+\beta)$ be, respectively, $(1+\alpha)^{1/m}-1$, $(1+\beta)^{1/m}-1$, where m, indefinitely large, is the number of mean proportionals in each line. By his verbal argument HALLEY shows that the magnitudes of these ratiunculae are, in the limit, as the numbers of the original ones (the ratio of which numbers is as that of the logarithms). Symbolically:

$$\log_k(1+\alpha):\log_k(1+\beta) = \lim_{m\to\infty}\left[\left((1+\alpha)^{1/m}-1\right):\left((1+\beta)^{1/m}-1\right)\right]$$
$$= \lim_{m\to\infty}\left[\frac{(1+\alpha)^{1/m}-1}{1/m} : \frac{(1+\beta)^{1/m}-1}{1/m}\right],$$

* Significantly, if the above restoration of NAPIER's thought-process is correct, HALLEY is unconsciously repeating NAPIER.

[47] In *A most compendious ... method of constructing the logarithms, exemplified and demonstrated from the nature of numbers, without any regard to the hyperbola*, *PT* **19** (1695) No. 215. Interestingly, HALLEY gives as his explicit reason for writing the article: "... I find very few of those who make constant use of logarithms to have attained an adequate notion of them, to know how to make or examine them, or to understand the extent of the use of them; contenting themselves with the tables of them as they find them, without daring to question them, or caring to know how to rectify them."

[48] This is, of course, a variation of BÜRGI's approach, and, in particular, had been developed into a practical technique by MERCATOR in *logarithmotechnia*, props. 1, 2: 1—10.

or $\log_k(1+\alpha)$ is proportional to $\lim_{m\to\infty}\left[\frac{(1+\alpha)^{1/m}-1}{1/m}\right]$. More generally, $\log_k(1\pm\alpha)$ is proportional to $\lim_{m\to\infty}\left[\frac{(1\pm\alpha)^{1/m}-1}{1/m}\right]$* and, as HALLEY points out, taking a suitable factor of proportionality—$\frac{1}{\log k}$, in fact, —gives us logarithms to a particular base (k). In particular, natural—"Lord NAPIER'S"—logarithms arise when the proportion factor is unity, or $k=e$.

This result, $\log(1\pm\alpha)=\pm\lim_{m\to\infty}\left[\frac{(1\pm\alpha)^{1/m}-1}{1/m}\right]$, is HALLEY'S analytical definition of the natural logarithm. Using it he finds the sum-series expansions of $\log(1\pm x)$ **, and of the exponential function $e^{\pm L}$*** very neatly.

By the end of the 17^{th} century we can say that, much more than being a calculating device suitably well tabulated, the logarithmic function—very largely on the geometrical model of hyperbola-area—had been accepted into mathematics. When, in the 18^{th} century, this geometrical basis was discarded in favour of a fully analytical one, no extension or reformulation was necessary—the concept of "hyperbola-area" was transformed painlessly into that of "natural logarithm". What remained to be done at the end of the 17^{th} century was, above all, to make precise its relationship with that of the circular functions, the narrowness of which seemed clear from several correspondences already verified—especially the dual nature of GREGORY'S analytical sequences in *VCHQ*—but whose nature was to be pin-pointed by such relations as COTES' $e^{\pm i\vartheta}=\cos\vartheta\pm i\cdot\sin\vartheta$ $\left(\cdot\equiv\cdot\, i\vartheta=\log\left(\frac{\cot\vartheta+i}{\cot\vartheta-i}\right)\right)$. Otherwise the (real function) logarithm had been tolerably well discussed.

---

* Indeed

$$\lim_{m\to\infty}\left[\frac{(1\pm\alpha)^{1/m}-1}{1/m}\right]=\lim_{n\to 0}\left[\frac{(1\pm\alpha)^{n}-(1\pm\alpha)^{0}}{n}\right],$$

the differential ("fluxion") of

$$\lim_{n\to 0}(1\pm\alpha)^{n}=\lim_{m\to\infty}(1\pm\alpha)^{1/m}:$$

then

$$\lim_{m\to\infty}\left(\frac{d}{dm}(1\pm\alpha)^{1/m}\right)=\lim_{m\to\infty}(\log(1\pm\alpha)\times(1\pm\alpha)^{1/m})=\log(1\pm\alpha),$$

$$=\log k\times\log_k(1\pm\alpha).$$

** Expanding by the binomial theorem,

$$\log(1\pm\alpha)=\lim_{m\to\infty}\left[\left(1\pm\frac{1}{m}\alpha+\frac{1(1-m)}{2!\,m^2}\alpha^2\pm\frac{1(1-m)(2-m)}{3!\,m^3}\alpha^3+\cdots\right)-1\right]m$$

$$=\pm\alpha+\tfrac{1}{2}\alpha^2\pm\tfrac{1}{3}\alpha^3+\tfrac{1}{4}\alpha^4\pm\cdots$$

*** Take $\log(1\pm\alpha)=\pm L$, or $\pm L=\lim_{m\to\infty}\left[\frac{(1\pm\alpha)^{1/m}-1}{1/m}\right]$. Unwrapping, $e^{\pm L}=1\pm\alpha$ $=\lim_{m\to\infty}\left(1\pm\frac{L}{m}\right)^m$; and expanding this by the binomial theorem, $e^{\pm L}=1\pm L+\frac{1}{2}L^2\pm\frac{1}{3!}L^3+\cdots$.

IV. Concept of function

2. *Interpolation*

By the first decades of the 17th century, the elementary mathematical functions (trigonometrical and logarithmic) had been tabulated to the accuracy of roughly, six or seven decimal figures for a large number of particular values densely packed in some adequate interval.[1] As with all tabulated functions it was a natural desire to seek ways of deriving intermediate values of the function from neighbouring (known) tabulated instances without the wearisome toil necessary in calculating each value of the function afresh from first principles.[2]

Fortunately, these elementary functions are well-behaved, having singularities only at a few exceptional points. More important—from the viewpoint of 17th century mathematics at least—a small variation in the argument provokes in such functions an equally small increase (or decrease), and such increases for uniform increase in the argument occur likewise very nearly at a uniform rate. On this fact is justified (usually only implicitly) the widespread use of linear interpolation to interpolate values of a function between given tabulated ones "not too widely differing". Briefly, for $h \in [O, H]$ we interpolate the value $f(x+h)$ between given values $f(x)$ and $f(x+H)$ by assuming $f(x+h)-f(x)=\frac{h}{H}\times[f(x+H)-f(x)]$, or equivalently by $f(x+H)-f(x+h)=\frac{H-h}{H}\times[f(x+H)-f(x)]$.[3]

Such a linear interpolation is, however, only accurate to an assignable error, and with the accuracy required of 17th century mathematical tables the method did not yield accurate enough tabulations except where the values $f(x)$, $f(x+H)$ differ only very slightly. Where and how, then, were improved methods to be found?

In historical fact, the refined methods were introduced by taking into consideration the differences of the differences $\Delta' f(x+\lambda H)=f(x+(\lambda+1)H)-f(x+\lambda H)$, and in general, the general n^{th} differences $\Delta^n f(x+\lambda H)$ defined recursively by $\Delta^n f(x+\lambda H)=\Delta^{n-1} f(x+(\lambda+1)H)-\Delta^{n-1} f(x+\lambda H)$. Indeed, the very form of the number-system accepted—where a general number N is denoted with respect to some number-base B by the unique sum-series $N=\sum_{0\leqq i\leqq I}(y_i B^i)$,

---

[1] Beginning with the HIPPARCHUS-PTOLEMY table of chords (which forms part of PTOLEMY's *Almagest*), the common trigonometrical functions—tabulated at first in sexagesimal fractions for suitable division of the interval $0°\leqq\vartheta°\leqq 90°$, but in Renaissance times more commonly in decimal form—had been calculated to an accuracy of several figures and roughly at 1′ intervals of angle. Of these outstanding were RHETICUS' 16th century tabulations. And with NAPIER's table of logarithms (strictly of logarithmic sines) and BRIGGS' adaptation to base 10 usable tables of the logarithmic function existed from the period 1614—1625 onwards.

[2] Years of work must have gone into the comparatively meagre chord tables of PTOLEMY, and we know that lifetimes were spent in the 16th century in improving the accuracy of existing trigonometrical tables.

[3] This is clearly a rounding-off of the general BRIGGS-NEWTON interpolation formula elaborated below: *viz*

$$f(x+h)\left[=f\left(x+\frac{h}{H}\times H\right)\right]=f(x)+\frac{h}{H}\times\Delta^1 f(x)+\cdots$$

[where $\Delta^1 f(x)=f(x+H)-f(x)$]. Such linear interpolation, in particular, was widely used by NAPIER in constructing his canon of logarithms (see previous chapter).

where each a_i is positive integral and less than B—made it natural that such an approach should arise.* So we find that BRIGGS, in the introduction to his *arithmetica logarithmica*[4] implicitly gives the basis for interpolation with regard to functions (tabulated at equal intervals of the argument) whose n^{th} differences are constant—functions, that is (as NEWTON at least was to see), whose representing polynomial is of finite degree n.

A clear insight into BRIGGS' process of thought is given if we consider in detail the section of AL [5] where he derives, apparently for the first time, the case $n=\frac{1}{2}$ of the binomial expansion of $(1+\alpha)^n$, $(|\alpha|<1)$.

In an earlier chapter[6] BRIGGS had sketched out an improvement on NAPIER's method for constructing the logarithms of numbers: briefly, he constructs a large number of geometric means between 1 and X (whose logarithm we seek) by repeatedly extracting square roots. In particular, to find[7] $\lambda=\log(1.0077696)$ he extracts 46 successive square roots and finds

$$\frac{1}{2^{46}}\times\lambda=\log(1.000000000000000109985934),$$

which is very nearly 10^{-24} (109985934).** While historians have largely seen this method as impractical, BRIGGS makes it a workable construction by considering n^{th} differences[8]. Symbolising his largely verbal method, we consider $\lambda=1.0077696$. By physical root-extraction BRIGGS tabulates $\lambda^{2^{-i}}$, $i=1, 2, 3, \ldots, 11$ in a form which, taking $\lambda^{2^{-\gamma}}=K$ $\left(\text{or } K^{2^i}=\lambda^{2^{i-\gamma}}\right)$ we can set out as:

$$\begin{aligned}
K-1 &= 0.00003\ 02331\ 60505\ 65775\ldots = e_0,\\
K^2-1 &= 0.00006\ 04672\ 35055\ 30968\ldots = e_1,\\
K^4-1 &= 0.00012\ 09381\ 26397\ 13459\ldots = e_2,\\
\ldots &= \ldots\ldots\ldots = \ldots\\
K^{128}-1 &= 0.00387\ 72833\ 36962\ 45663\ldots = e_7,\\
K^{256}-1 &= 0.00776\ 96 \qquad\qquad = e_8=\lambda,
\end{aligned}$$

where $e_s=K^{2^s}-1$.

Looking at this table BRIGGS saw that e_i is, for each i, very nearly equal to $\frac{1}{2}e_{i+1}$; and so, taking $\Delta_i^1=\frac{1}{2}e_{i+1}-e_i$, he sets up a second table of the Δ_i^1, $i=0, 1, 2, \ldots$. Looking at this BRIGGS finds that, even more nearly, Δ_i^1 is approximately equal to $\frac{1}{4}\Delta_{i+1}^1$, and so considers "modified" second differences $\Delta_i^2=\frac{1}{4}\Delta_{i+1}^1-\Delta_i^1$. In general, he finds that the "modified" k^{th} difference Δ_i^k

* When we consider general types of operation which can be performed on the two numbers $\lambda=\sum_{0\leq i\leq I}(a_iB^i)$, $\mu=\sum_{0\leq j\leq J}(b_jB^j)$, the sum or difference $\lambda\pm\mu$, seems easiest.

** Using the approximation (true for μ small) $\log(1+\mu)\approx\mu$.

[4] Especially in chapters 8—13.

[5] *AL*: Chapter 8.

[6] *AL*: Chapter 6.

[7] The example arises in finding log(6), since $\lambda=9\times\log(6)-7\times\log(10)$.

[8] *AL*: ch. 8: 17: "atque ad hunc modum cujuscunque numeri propositi logarithmum per continue medios invenire poterimus: quos nobis lateris quadrati inventis suppenditat satis laboriose. *hujus autem tanti laboris molestia minuetur plurimum per differentias*" (my italics).

is very nearly equal to $\frac{1}{2^k}\Delta^k_{i+1}$, and so defines a "modified" $k+1^{\text{th}}$ difference recursively by $\Delta^{k+1}_i = \frac{1}{2^k}\Delta^k_{i+1} - \Delta^k_i$. Reformulating BRIGGS' empirical observations, the kernel of his insight is that, for Δ^k_i so defined, $\lim_{k\to\infty}(\Delta^k_i) = 0$.

BRIGGS now unwraps the "modified" differences, beginning with some stage λ and taking all higher differences to be zero. We have then, since $\Delta^\lambda_i = \frac{1}{2^{\lambda-1}} \times \Delta^{\lambda-1}_{i+1} - \Delta^{\lambda-1}_i$ $\left(\text{or equivalently } \Delta^{\lambda-1}_i = \frac{1}{2^{\lambda-1}} \times \Delta^{\lambda-1}_{i+1} - \Delta^\lambda_i\right)$, that

$$e_i = \tfrac{1}{2} e_{i+1} - \Delta^1_i,$$
$$\Delta^1_i = \tfrac{1}{4}\Delta'_{i+1} - \Delta^2_i,$$
$$\cdots\cdots$$

and

$$\Delta^\lambda_i = \tfrac{1}{2}\Delta^\lambda_{i+1}$$

or

$$e_i = \frac{1}{2} e_{i+1} - \frac{1}{4}\Delta^1_{i+1} + \frac{1}{8}\Delta^2_{i+1} - \cdots (-1)^{\lambda-1}\frac{1}{2^\lambda}\Delta^\lambda_{i+1},$$

and in particular

$$e_{-1} = \tfrac{1}{2} e_0 - \tfrac{1}{4}\Delta^1_0 + \tfrac{1}{8}\Delta^2_0 - \cdots.$$

It only remains to evaluate these Δ^λ_0—specifically, taking $K = 1 + c$, BRIGGS tabulates the Δ^λ_0 in terms of powers of α [9], expanding $e_i = (1+\alpha)^{2^i} - 1$, $i = 1, 2, 3, \ldots$. Thus

$$\Delta^1_0 = \tfrac{1}{2} e_1 - e_0 = \tfrac{1}{2}\alpha^2,$$
$$\Delta^2_0 = \tfrac{1}{4}\Delta^1_1 - \Delta^1_0 = \tfrac{1}{2}\alpha^3 + \tfrac{1}{8}\alpha^4,$$
$$\Delta^3_0 = \tfrac{7}{8}\alpha^4 + \tfrac{7}{8}\alpha^5 + \tfrac{7}{16}\alpha^6 + \tfrac{1}{8}\alpha^7 + \tfrac{1}{64}\alpha^8,$$
$$\cdots\cdots$$
$$\Delta^9_0 = 2805\,527\,\alpha^{10} + \cdots.$$

The general pattern now becomes obvious: Δ^{m-1}_0 has no powers of α less than α^m—the difficult proof of which BRIGGS does not attempt. After so much that is dull the final stage becomes enormously exciting. Substituting these expansions of e_i in the expansion of $e_{-1} = (1+\alpha)^{\frac{1}{2}} - 1$ we have, on collecting powers of α, the binomial expansion

$$(1+\alpha)^{\frac{1}{2}} - 1 = \tfrac{1}{2}\alpha - \tfrac{1}{8}\alpha^2 + \tfrac{1}{16}\alpha^3 - \tfrac{5}{128}\alpha^4 + \tfrac{7}{256}\alpha^5 \ldots,$$

and since BRIGGS specifically notes that he used (an equivalent of) this expansion in improving NAPIER'S canon, there emerges the interesting fact that the first construction of logarithms by series-approximations used a binomial expansion rather than a direct logarithmic function expansion.*

---

* A similar "BRIGGSIAN" process with respect to $f_{ik} = (1+\alpha)^{p^i}$ using $\Delta^1_i = \frac{1}{p} f_{i+1} - f_i$, $\Delta^{k+i}_i = \frac{1}{p^{k+1}}\Delta^k_{i+1} - \Delta^k_i$, $k = 1, 2, \ldots$, yields the (unit-fractional) binomial expansion of $(1+\alpha)^{1/p}$.

[9] *AL:* 16, where the BOMBELLI ring-notation for powers of the variable makes the text extremely difficult to follow.

BRIGGS seems to have looked on this method only as a computing convenience, missing its general significance[10], but we know so little about the development of BRIGGS' mathematical thought that it is difficult to begin to guess how highly he thought of his square-root method. It is clear, however, that he had made a profound study of the n^{th}-order finite differences. In later chapters of his *AL*[11] he gives, without prior investigation or justification, rules which contain implicitly the general "NEWTON-GAUSS" interpolation formula,

$$f(x+h) = f(x+\lambda H) = f(x) + \binom{\lambda}{1}\Delta^1 f(x) + \binom{\lambda}{2}\Delta^2 f(x) + \cdots,$$

where the function instances $f(x+L\cdot H)$, $L=0, \pm 1, \pm 2, \ldots$ are given at H-intervals of the argument,

$$\lambda = \frac{h}{H}, \quad \text{and} \quad \Delta^1 f(x+L\cdot H) = f\big(x+(L+1)\cdot H\big) - f(x+L\cdot H)$$

$$\Delta^{k+1} f(x+L\cdot H) = \Delta^k f\big(x+(L+1)\cdot H\big) - \Delta^k f(x+L\cdot H).$$

The formula is used, for example, in *AL*[12] specifically tied to the logarithmic function tabulated at unit intervals and rounded off at the second difference, in the form

$$\log(a+h) \approx \log(a) + h\,\Delta^1 \log(a) + \frac{h(h-1)}{2}\Delta^2 \log(a), \qquad 10h = 1, 2, \ldots, 9$$

to derive easy rules for finding $\log(a+h)$ from the instances $h=-1, 0, 1, 2$*— that is, a rule for subtabulating by $1/10^{th}$ in the interval $[a, a+1]$. More generally, he seems to have used it in deriving the general rules for treating mean differences in subtabulation which he lists in *AL*[13].

This unwillingness to commit his methods to print contributed without doubt to the general lack of recognition of BRIGGS' mathematical worth in the 17^{th} cen-

---

* The instances $f(a-1)$, $f(a)$, $f(a+1)$, $f(a+2)$ are obviously sufficient to yield the necessary second differences.

[10] Partly that may be due to the inadequate representation afforded by his ring-notation for powers, but it is certain that no others in the 17^{th} century, if they understood the equivalence of BRIGGS' approach with the general binomial expansion —which is highly doubtful—, considered it as anything but an abstruse computing technique for logarithmic tabulations. Curiously BRIGGS in his (posthumous) *trigonometria britannica*. Gouda, 1633 (apparently deriving his inspiration from VIETA) had given in his ABACUS *ΠΑΓΧΡΗΣΤΟΣ* the construction of a table of figurate numbers—in effect a "Pascal"-triangle modified into a rectangular array such that the number in his i^{th} column and j^{th} row is $\binom{i+j-1}{j-1} = \binom{i+j-1}{i}$. Nowhere, however, does he hint that these numbers have anything to do with the coefficients of powers of α in his expansion, and the application had to wait till NEWTON.

[11] *AL:* chapters 12, 13. Chapter 13: 27—32, omitted from VLACQ's continental edition, is reprinted with the slight changes necessitated by the substitution of sin and tangent functions for the logarithm in *trigonometria britannica:* 38ff.

[12] *AL:* ch. 12.

[13] *AL:* 29ff. = *trigonometria britannica:* 38. Specifically BRIGGS gives rules for correcting the mean differences (as far as the 20^{th} difference) in quinquisecting the interval to be subtabulated, and it is significant that they agree exactly with the rule given by ROGER COTES using the NEWTON-BESSEL and NEWTON-STIRLING formulas in his *canonotechnica, sive constructio tabellarum per differentias:* prop. 6: 48—50 (printed at the end of his *harmonia mensurarum*, London, 1722).

tury[14] and with his death interest in the theory of tabular interpolation lapsed till the 1660's, when it was revived in an elementary way by MERCATOR, perhaps inspired by a reading of BRIGGS' *AL*, but more especially by NEWTON and JAMES GREGORY, who clearly saw the equivalence of tabular interpolation with the problem of fitting an n-degree polynomial to a set of points (the end-points of CARTESIAN coordinate lengths which represent the known tabulated instances) on the basis of successive differences (up to those of the n^{th} order).

Meanwhile in the 1650's JOHN WALLIS had developed a variant type of interpolation method which he used virtually to interpolate between integral functions tabulated for certain regularly separated values of the arguments.[15] As WALLIS gives it the method is very loosely founded on what is basically only a strong feeling for pattern; yet, though—as will be clear from a detailed analysis—this laxity could at times introduce more complexities than he could control (or even be aware of), when the method is put on a rigorous basis and made precise it proves very fertile.[16]

---

[14] Few 17th century mathematicians seem to have read BRIGGS' lengthy and apparently obscure introductions to his tables—certainly not WALLIS, who is usually only too ready to overestimate English mathematical achievement. JAMES GREGORY is, however, the exception—compare his answer to a query of COLLINS about BRIGGS' subtabulation methods (GREGORY *TV:* 118—122, especially 120). It is tempting to conjecture (with D.C. FRASER: *Newton's interpolation formulas:* 57—58) that NEWTON studied BRIGGS' work at an early stage in his life, but there is nothing in any of the Portsmouth Collection of Newton manuscripts which corroborates this, and it would seem likely that if he had done so he would have realized the significance of BRIGGS' square-root procedure and given him due credit as a formative influence on his own ideas along with WALLIS (*cf. CUL Add.* 4000: 14V).

Appreciative accounts of BRIGGS' work and its influence are given by CHARLES HUTTON in his historical preface to his revised (5th) edition of SHERWIN'S *Mathematical tables* London 1785 (=MASERES: *scriptores logarithmici*, **1** London 1791: i-cxi, especially lxiii—lxxxiii); and in H.W. TURNBULL *James Gregory: a study in the early history of interpolation.* Proc. Edinburgh Math. Soc.$_2$ **3** (1932—33): 151—178, especially 164—168.

[15] *AI: passim—cf.* a faithful but uninspired account in J.F. SCOTT: *The mathematical work of John Wallis:* ch. 4; 26—64. It is interesting to see how WALLIS' methods may be related to his work on codes during the English Civil War. In particular, the whole pattern of his layout on the printed page corresponds closely with the natural way of setting out a coded message for decoding. Moreover, the two problems are akin on a logical level. Essentially WALLIS in his interpolation approach sets up the pattern of tabulated instances in a two-dimensional array, and then compares individual instances with surrounding ones in a search for general aspects of the pattern—much as the decoder uses context checks in trying to abstract a meaning from the pattern of symbols before him. Codes in use in the Civil War were suggestively numerical, with easily recognizable frequency patterns occurring among the various number-sets used—typically, such a pattern as la, le, li, lo, lu (a consonant together with the five vowels in order) is represented by the number pattern $\alpha + \lambda\beta$, $\lambda = 0, 1, 2, 3, 4$, α, β suitable integers (very often multiples of 5 in the codes I have checked). JOHN DAVIS in *An essay on the art of decyphering, in which is inserted a discourse of Dr. Wallis ...*, London, 1737: 26 gives a numerically coded letter from the Civil War period; while two further letters, dated 1689, deciphered by WALLIS are given in his *opera* **3** (1699): 660—672 together with keys and transcriptions.

[16] It was EULER who, above all others, established more rigorous treatments of WALLIS' suggestive ideas in many papers (too numerous to enumerate here) and intermittently over most of his life.

Wallis, after a false start* seeks virtually[17] in the latter part of his *arithmetica infinitorum*[18] to interpolate $f(\frac{1}{2}, \frac{1}{2})$ in tabulated instances, λ, μ positive integral, of $f(\lambda, \mu) = \dfrac{1}{\int_0^1 (1-x^{1/\lambda})^\mu \cdot dx}$**. These "brute-force" tabulations are made on the basis of elaborate and diffuse techniques developed in the early part of AI[19]. Briefly, he gives (with strict proof only for a few particular cases) the equivalent of $\int_0^1 x^k \cdot dx = \frac{1}{k+1}$, where k is of the form p or $\frac{1}{p}$, p a positive integer, and then assumes the rule true on a mere instinctive basis of analogy for all k rational. It then becomes possible to evaluate any particular $f(\lambda, \mu)$ by physically expanding $(1 - x^{1/\lambda})^\mu = (1 - k)^\mu$ as a binomial in powers of $k = x^{1/\lambda}$, and then integrating the resulting sequence term by term. So, for example

$$\int_0^1 (1 - x^{\frac{1}{3}})^2 \cdot dx = \int_0^1 (1 - 2x^{\frac{1}{3}} + x^{\frac{2}{3}}) \cdot dx = \frac{1}{1+0} - 2 \cdot \frac{1}{1+\frac{1}{3}} + \frac{1}{1+\frac{2}{3}} = \frac{1}{10},$$

or

$$f(3, 2) = 10 \left[= \frac{\Gamma(3+2+1)}{\Gamma(3+1) \cdot \Gamma(2+1)} = \frac{(3+2)!}{3!\,2!} = \frac{5!}{3!\,2!} \right].$$

Next Wallis sets up a square table of $f(\lambda, \mu)$, $\lambda, \mu = 1, 2, \ldots, 10$, which he extends by analogy to include the cases where either (or both) $\lambda, \mu = 0$.

---

* His aim, to find an approximate circle quadrature by interpolation of $\int_0^1 (1-x^2)^{\frac{1}{2}} \cdot dx$ between suitable integrals, naturally led him first to treat $\Phi(\lambda, \mu) = \dfrac{1}{\int_0^1 (1-x^\lambda)^\mu \cdot dx}$, trying to interpolate $\Phi(2, \frac{1}{2})$ between $\Phi(2, 0)$ and $\Phi(2, 1)$; but with his techniques he could see no pattern coming through and abandoned it.

** In fact, $\dfrac{1}{f(\lambda, \mu)} = \int_0^1 (1-x^{1/\lambda})^\mu \cdot dx$, which for $x \geqq 0$, is transformable by $x^{1/\lambda} \to y$, into

$$\int_0^1 (1-y)^\mu \cdot d(y^\lambda), \; = \int_0^1 (1-y)^\lambda \cdot d(y^\mu) = \frac{1}{f(\mu, \lambda)}, \; = \lambda \cdot B(\lambda, \mu+1) = \frac{\Gamma(\lambda+1) \cdot \Gamma(\mu+1)}{\Gamma(\lambda+\mu+1)}:$$

so that $x \geqq 0$, $f(\lambda, \mu) = \dfrac{\Gamma'(\lambda+\mu+1)}{\Gamma(\lambda+1) \cdot \Gamma(\mu+1)}$. For $x < 0$, the integral bounds are changed to $\int_0^1 (1-y)^\mu \cdot d(y^\lambda)$, which takes on *no* real values, unlike $\dfrac{\Gamma(\lambda+\mu+1)}{\Gamma(\lambda+1) \cdot \Gamma(\mu+1)}$ which is defined as a real function for $\left.\begin{matrix}\lambda \\ \mu \\ \lambda+\mu\end{matrix}\right] \geqq -1$ (a fact which is taken up below). We should not despise this too much. Euler, following up many of Wallis' root ideas, frequently appeals to the extramathematical concept "ex lege continuitatis".

[17] He has no symbolism for integration but defines the integral loosely as a limit sum sequence—see chapter 8.

[18] AI: props. 128—191, with omissions.

[19] A fuller consideration will be given when we treat of general indivisible theories (see chapter 8).

Thus:

λ	μ 0	1	2	3	...	10
0	1	1	1	1	...	1
1	1	2	3	4	...	11
2	1	3	6	10	...	66
3	1	4	10	20	...	286
⋮	⋮	⋮	⋮	⋮	⋮	⋮
10	1	11	66	286	...	184756

The symmetry of the table, $f(\lambda, \mu) = f(\mu, \lambda)$, stands out, but WALLIS also notes that the number-sequences are figurate (forming a "PASCAL" triangle), or $f(\lambda+1, \mu+1) = f(\lambda, \mu+1) + f(\lambda+1, \mu)$[20] *. WALLIS, however, has set himself the problem of finding $f\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{\int_0^1 (1-x^2)^{\frac{1}{2}} \cdot dx} = \frac{4}{\pi}\left(= \frac{\Gamma(2)}{(\Gamma(\frac{3}{2}))^2} = \frac{4}{(\Gamma(\frac{1}{2}))^2}\right)$, which he denotes by '□'. This, in preparation for interpolating intermediate values, he

λ	μ $-\frac{1}{2}$	0	$\frac{1}{2}$	1	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	...
$-\frac{1}{2}$	·	·	·	·	·	·	·	·	...
0	·	*1*	·	*1*	·	*1*	·	*1*	...
$\frac{1}{2}$	·	·	□	·	·	·	·	·	...
1	·	*1*	·	*2*	·	*3*	·	*4*	...
$1\frac{1}{2}$	·	·	·	·	·	·	·	·	...
2	·	*1*	·	*3*	·	*6*	·	*10*	...
$2\frac{1}{2}$	·	·	·	·	·	·	·	·	...
3	·	*1*	·	*4*	·	*10*	·	*20*	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

inserts an expanded version of his table[21], and tries to abstract a general pattern on which he can introduce interpolated values.

To simplify further discussion we take $\lambda = \frac{1}{2}l$, $\mu = \frac{1}{2}m$, $f(l, m) = l_m = m_l$ (the tabulated instance to be found on the l^{th} row/column and m^{th} column/row in the revised table below). **

WALLIS, stressing that the tabulated instances are figurate and considering only the rows—the diagonal symmetry of the table clearly implies that there is an equivalent treatment by columns—shows numerically that this property

* $$f(\lambda, \mu+1) + f(\lambda+1, \mu) = \frac{\Gamma(\lambda+\mu+2)}{\Gamma(\lambda+1)\cdot\Gamma(\mu+2)} + \frac{\Gamma(\lambda+\mu+2)}{\Gamma(\lambda+2)\cdot\Gamma(\mu+1)} = \frac{\Gamma(\lambda+\mu+3)}{\Gamma(\lambda+2)\cdot\Gamma(\mu+2)} = f(\lambda+1, \mu+1).$$

** $$l_m, \text{ for } l, m \geqq 0, = \frac{\Gamma\left(\frac{l}{2}+\frac{m}{2}+1\right)}{\Gamma\left(\frac{l}{2}+1\right)\cdot\Gamma\left(\frac{m}{2}+1\right)} = \frac{\Gamma\left(\frac{l+m+2}{2}\right)}{\Gamma\left(\frac{l+2}{2}\right)\cdot\Gamma\left(\frac{m+2}{2}\right)}.$$

[20] *AI:* prop. 131.
[21] *AI:* prop. 169.

implies, for l, m both even,

$$l_m \left[= \binom{\frac{l+m}{2}}{\frac{m}{2}} = \frac{\frac{l+m}{2}}{\frac{m}{2}} \times \binom{\frac{l+m}{2}-1}{\frac{m}{2}-1} \right] = \frac{l+m}{m} \times l_{m-2},$$

and supposes by "analogy" (and it is immediately provable*) that this holds for all l, m (at least, in WALLIS' table, $l, m \geqq -1$). From there the interpolation goes fairly neatly, yielding the table[22] (where $l_1 = \square$ furnishes the basis for setting up the l_m, l, m both odd). [-1_{-1} is tabulated as infinity—"∞" the first use of the symbol—for consistency, since $-1_{-1} = \frac{1}{1-1} \times 1_{-1} \to \infty$; specifically -1_{-1} is the limit of $\frac{\Gamma(\varepsilon)}{\Gamma(\frac{1}{2}) \cdot \Gamma(\frac{1}{2}+\varepsilon)}$ where $\lim_{\varepsilon \to 0} (\Gamma(\varepsilon)) \to \infty$.]

l \ m	-1	0	1	2	3	4	...	8	...
-1	"∞"	1	$\frac{1}{2} \cdot \square$	$\frac{1}{2}$	$\frac{1}{3} \cdot \square$	$\frac{5}{8}$	...	$\frac{105}{384}$	...
0	1	1	1	1	1	1	...	1	..
1	$\frac{1}{2} \cdot \square$	1	$\square$	$\frac{3}{2}$	$\frac{4}{3} \cdot \square$	$\frac{15}{8}$	...	$\frac{945}{384}$	...
2	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	...	5	...
3	$\frac{1}{3} \cdot \square$	1	$\frac{4}{3} \cdot \square$	$\frac{5}{2}$	$\frac{8}{3} \cdot \square$	$\frac{35}{8}$	...	$\frac{3465}{384}$	...
4	$\frac{3}{8}$	1	$\frac{15}{8}$	3	$\frac{35}{8}$	6	...	15	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
8	$\frac{105}{384}$	1	$\frac{945}{384}$	5	$\frac{3465}{384}$	15	...	...	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

WALLIS finally achieves his interpolation by noting that $l_m < l_{m+2}$ for all l, m in his table (excluding $l, m < 0$) and he assumes true for all l, m (positive) by "analogy" the "interpolated" law, $l_m < l_{m+1} < l_{m+2}$ (or $l_m < (l+1)_m < (l+2)_m$).

* In fact,

$$l_m = \frac{\Gamma\left(\frac{l+m+2}{2}\right)}{\Gamma\left(\frac{l+2}{2}\right) \cdot \Gamma\left(\frac{m+2}{2}\right)} = \frac{l+m}{m} \times \frac{\Gamma\left(\frac{l+m}{2}\right)}{\Gamma\left(\frac{l+2}{2}\right) \cdot \Gamma\left(\frac{m}{2}\right)} = \frac{l+m}{m} \times l_{m-2}.$$

[22] *AI:* prop. 189. This table is a fine example of WALLIS' lack of control over his interpolation—it represents $f\left(\frac{l}{2}, \frac{m}{2}\right) = \frac{1}{\int_0^1 (1-x-2/l)^{m/2}} \cdot dx$ *only* for $l, m \geqq 0$. In fact, WALLIS' interpolation, based on the recursive formation rule $l_m = \frac{l+m}{m} \times l_{m-2} = m_l$, interpolates $g\left(\frac{l}{2}, \frac{m}{2}\right) = \frac{\Gamma\left(\frac{l+m}{2}+1\right)}{\Gamma\left(\frac{l}{2}+1\right) \cdot \Gamma\left(\frac{m}{2}+1\right)}$ which takes on the values of $f\left(\frac{l}{2}, \frac{m}{2}\right)$ for all positive l, m, but which—as the table and unlike $f\left(\frac{l}{2}, \frac{m}{2}\right)$—is defined also for $l, m \geqq -1$ (and indeed for $l, m \geqq -2$ with $l+m \geqq -2$).

Using the particular case of this, $1_m < 1_{(m+1)} < 1_{(m+2)}$, WALLIS' product arises as the limit form $m \to \infty$: a simplification introduced historically by NEWTON[23]. In fact, isolating the row 1_m we can tabulate it as

m	0	1	2	3	4	5	6	...
1_m	1	$\square\left(=\frac{4}{\pi}\right)$	$\frac{3}{2}$	$\frac{4}{3}\times\square$	$\frac{3}{2}\times\frac{5}{4}$	$\frac{4}{3}\times\frac{6}{5}\times\square$	$\frac{3}{2}\times\frac{5}{4}\times\frac{7}{6}$	...

where

$$1_{2n} = \prod_{1\leq i\leq n}\left(\frac{2i+1}{2i}\right)\times 1_0(=1) = 1\times\frac{3}{2}\times\frac{5}{4}\times\frac{7}{6}\times\cdots\times\frac{2n+1}{2n}$$

and

$$1_{2n+1} = \prod_{1\leq i\leq n}\left(\frac{2i+2}{2i+1}\right)\times 1_1(=\square) = \square\times\frac{4}{3}\times\frac{6}{5}\times\frac{8}{7}\times\cdots\times\frac{2n+2}{2n+1}$$

and therefore $1_{(2n-1)} < 1_{2n} < 1_{(2n+1)}$ implies

$$\square\times\prod_{1\leq i\leq n-1}\left(\frac{2i+2}{2i+1}\right) < \prod_{1\leq i\leq n}\left(\frac{2i+1}{2i}\right) < \square\times\prod_{1\leq i\leq n}\left(\frac{2i+2}{2i+1}\right),$$

or

$$\prod_{1\leq i\leq n}\left(\frac{(2i+1)^2}{2i\,(2i+2)}\right) < \square\left(=\frac{4}{\pi}\right) < \prod_{1\leq i\leq n}\left(\frac{(2i+1)^2}{2i\,(2i+2)}\right)\times\frac{2n+2}{2n+1},$$

which, on slight rearrangement, yields the infinite sequence ("WALLIS' theorem") for $\frac{1}{2}\pi$,

$$\frac{\pi}{2} = \lim_{n\to\infty}\left(\prod_{1\leq i\leq n}\frac{(2i)^2}{(2i-1)\,(2i+1)}\right).^{\star}$$

WALLIS, however, in his *AI*, states this in a stronger form, using (the equivalent of a SCHWARZ inequality) $(1_m)^2 > 1_{(m-1)}\times 1_{(m+1)}$**—a procedure which yields the more powerful result[25]

$$\prod_{1\leq i\leq n}\left(\frac{(2i+1)^2}{2i\,(2i+2)}\right)\times\left(\frac{2n+3}{2n+2}\right)^{\frac{1}{2}} < \square < \prod_{1\leq i\leq n}\left(\frac{(2i+1)^2}{2i\,(2i+2)}\right)\times\left(\frac{2n+2}{2n+1}\right)^{\frac{1}{2}}.$$

To return to a general viewpoint, this reasoning by analogy—or perhaps more correctly from a feeling for a general pattern which seems to run through a set of particular results—exemplifies a process which must be fundamental to any system of interpolation: since there are an infinite number of ways of filling in a pattern, we choose that way which seems best suited (in a sense wider than the strictly mathematical), best conforms, to the instances known. WALLIS' assumptions in his derivation are quite audacious, and in a rigorous treatment must be carefully justified—yet in following through an intuition that he was

* Found independently by PIETRO MENGOLI in much the same way as WALLIS about 1659, but published only in his *circolo*.[24]

** Stated "by analogy" (per analogiam), *sc.* by induction from a few numerical instances.

[23] In his manuscript annotations from WALLIS' *AI* (to be dated 1665) in CUL Add. 4000: 16V–17V.

[24] *circolo: ... il problema della quadratura del circolo*. Bologna, 1672.

[25] *AI:* prop. 190. It is interesting to note that it is a particular case derivable from more general theorems given by MENGOLI in *circolo* which draw their analytical justification from a logarithmic inequality established by MENGOLI at the end of Book 5 of his *geometrica speciosa* (1659).

thereby achieving a result which is both true and important WALLIS was doing something practised by every creative mathematician, however lucky in that he did not seriously have to consider the boundary-cases where such general reasoning by pattern must break down.

One aspect, however, of his interpolation scheme did not satisfy WALLIS' instinct for consistent pattern and harmony. He had gained his continued-product sequence by assuming a numerical ordering of particular values of l_m (and in particular 1_m), but he had not been able to give a unified treatment which harmonized the two independent product-sequences,

$$1_{2n} = 1 \times \frac{3}{2} \times \frac{5}{4} \times \cdots \times \frac{2n+1}{2n}, \quad \text{and} \quad 1_{2n+1} = \square \times \frac{4}{3} \times \frac{6}{5} \times \cdots \times \frac{2n+2}{2n+1}.$$

WALLIS' instinct for symmetry resented this essential lack of formal similarity between odd and even values of 1_m, but, despite his trying many ways of modification, he could find no general pattern which would generate both as particular instances. Sometime in 1654, therefore, he seems to have asked BROUNCKER for a solution which should preserve the essential unity of 1_m, independently of m being odd or even, present in its definition, $1_m = \dfrac{1}{\int\limits_0 (1-x^2)^{m/2} \cdot dx}$.[26] The solution which BROUNCKER returned is sketched by WALLIS[27]—specifically

$$1_m = \frac{\square}{2} \times \frac{2}{\Phi(1)} \times \frac{4}{\Phi(3)} \times \cdots \times \frac{2m+2}{\Phi(2m+1)} = \frac{\square}{2} \times \prod_{0 \leq i \leq n} \left(\frac{2i+2}{\Phi(2i+1)}\right),$$

where $\Phi(x)$ is that function which satisfies $\Phi(x-1) \times \Phi(x+1) \equiv x^2$.* It is an immediate consequence that $1_0 (=1) = \frac{\square}{2} \times \frac{2}{\Phi(1)}$, or $\Phi(1) = \square$, and, if the nature of $\Phi(x)$ can be precisely delimited, we have a calculable value for $\square\,(=4/\pi)$. In fact[28] BROUNCKER found that $\Phi(x)$ can be given (for $x > 0$ implicitly) by the infinite continued fraction $\Phi(x) = \lim\limits_{i \to \infty} \left(x + \frac{1^2}{2x+} \frac{3^2}{2x+} \cdots \frac{(2i-1)^2}{2x}\right)$—from which the "BROUNCKER" continued fraction $\frac{1}{\square} = \frac{1}{\Phi(1)} \left(= \frac{\pi}{4}\right) = \frac{1}{1+} \frac{1^2}{2+} \frac{3^2}{2+} \cdots$ is an immediate deduction.[29]

---

* He gives also the immediate extension to general l_m.

[26] *AI:* prop. 191: scholium, where he tries to express this concept in a verbal statement—compare *AI* (1656): 181–182. "When I had proposed to Brouncker some of my propositions and had indicated by what law they proceeded, I asked him to show in what form that quantity □ could most conveniently be designated."

[27] *AL:* prop. 191: *scholium* and *idem aliter.*

[28] See the previous chapter.

[29] I find it curious that the process should give this form, whose convergents (as EULER showed on many occasions) are the successive convergents to the NILAKANTHA-LEIBNIZ sum sequence, $\frac{\pi}{4} = \lim\limits_{n \to \infty} \sum\limits_{0 \leq \lambda \leq n} \left[(-1)^\lambda \times \frac{1}{2\lambda+1}\right]$, rather than $\frac{\pi}{2} = \lim\limits_{n \to \infty} \left(1 + \frac{1}{1+} \frac{1}{\frac{1}{2}+} \cdots \frac{1}{1/n} + \cdots\right)$, whose convergents yield WALLIS' continued product. (The latter is given by EULER in *de fractionibus continuis observationes*, Comm. ac. sc. Petrop. **11** (1739) [1750]: 39–81, especially § 36: 51 ·≡· *opera* **14**a 1 (1925): 316; but the identification of its convergents with successive approximations to WALLIS' product-sequence was first made by J. J. SYLVESTER in *Note on a new continued fraction applicable to the quadrature of the circle*, Phil. Mag. **37** (1869): 373–375, especially 375 ·≡· *Collected mathematical papers*, **2**: 692.)

WALLIS did not restrict his method of interpolation by "analogy" to this example, but a few years later showed the power of the method[30] in evaluating the area under a cissoid.[31] Taking his definition of the cissoid as the point set of B such that $BL:AL = AL:KL$, where BLK is drawn perpendicular to the diameter AD of circle $\overparen{ACDC'}$ and F is the point at infinity on the tangent DH, he shows that area $(\overparen{ABFD}) = 3 \times$ area of the semicircle (ACD). In fact, where AD is unit-length, $AL = x$, $LK^2 = x(1-x)$ so that $BL = \frac{LA^2}{LK} = \left(\frac{x^3}{1-x}\right)^{\frac{1}{2}}$ and area $(\overparen{ABFD}) = \lim \cdot \sum (BL \cdot \Delta(AL)) = \int_0^1 \left(\frac{x^3}{1-x}\right)^{\frac{1}{2}} \cdot dx$. Similarly, the area of the semi-(circle ACD) is

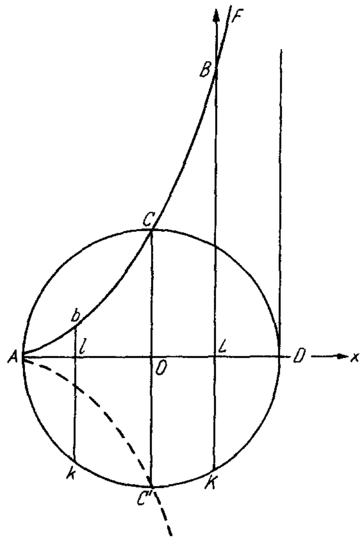


Fig. 15

$$\int_0^1 (x(1-x))^{\frac{1}{2}} \cdot dx = \frac{\pi}{8} = (\text{WALLIS'}) \frac{1}{2\square},$$

and WALLIS sets up a sequence of integrals, "justified" by appeal to analogy,

$$f_\lambda = \int_0^1 x^{\frac{1}{2}} (1-x)^{\lambda/2} \cdot dx$$
$$= \int_0^1 (1-x)^{\frac{1}{2}} x^{\lambda/2} \cdot dx \left(= B\left(\frac{3}{2}, \frac{\lambda}{2} + 1\right)\right),$$

and

$$g_\mu = \int_0^1 x^{\frac{3}{2}} (1-x)^{\mu/2} \cdot dx$$
$$= \int_0^1 (1-x)^{\frac{3}{2}} x^{\mu/2} \cdot dx \left(= B\left(\frac{5}{2}, \frac{\mu}{2} + 1\right)\right),$$

with which to compare these two integrals.

In detail his approach is very much as that developed in AI. Thus, by straight multiplication and integration, WALLIS tabulates particular values of f_λ, g_μ; λ, μ positive even: for example,

$$f_4 = \int_0^1 x^{\frac{1}{2}} (1-x)^2 \cdot dx = \int_0^1 (x^{\frac{1}{2}} - 2x^{\frac{3}{2}} + x^{\frac{5}{2}}) \cdot dx = \frac{1}{1+\frac{1}{2}} - \frac{2}{1+\frac{3}{2}} + \frac{1}{1+\frac{5}{2}}$$
$$= \frac{2 \cdot 2 \cdot 4}{3 \cdot 5 \cdot 7} \left(= B\left(\frac{3}{2}, 3\right) = \frac{\Gamma(\frac{3}{2}) \cdot \Gamma(3)}{\Gamma(\frac{9}{2})}\right).$$

Using these tabulated instances, he is able to set up the table:

λ	0	2	4	6	...
f_λ	$\frac{2}{3}$	$\frac{2}{3} \times \frac{2}{5}$	$\frac{2}{3} \times \frac{2}{5} \times \frac{4}{7}$	$\frac{2}{3} \times \frac{2}{5} \times \frac{4}{7} \times \frac{6}{9}$	...

[30] HUYGENS wrote in a manuscript draft of a letter in 1658: *in cissoide apparet vis methodi*. HUYGENS *OE*. 3: 58.

[31] In his *tractatus duo de cycloide* ..., Oxford 1659: especially 81–90, which is a part transcription of a 1658 letter to HUYGENS—compare JOSEPHA & J. E. HOFMANN: *Erste Quadratur der Kissoide*, Deutsche Mathematik 5 (1940–1941): 571–584, especially § 2: WALLIS.

from which he derives the formation-rule (for even λ) $f_\lambda = \frac{\lambda}{\lambda+3} \times f_{\lambda-2}$. This recursion rule is assumed by "analogy" to hold generally* and, tabulating f_1 by its value of $\frac{1}{2\square}$, he derives the expanded table:

λ	-1	0	1	2	3	4	…
f_λ	$\frac{\frac{1}{2\square}}{\frac{1}{4}}$	$\frac{2}{3}$	$\frac{1}{2\square}$	$\frac{2}{3}\times\frac{2}{5}$	$\frac{1}{2\square}\times\frac{3}{6}$	$\frac{2}{3}\times\frac{2}{5}\times\frac{4}{7}$	…

Using the recursion $g_\mu = \frac{\mu}{\mu+5} \times g_{\mu-2}$**, he tabulates even values of μ in a second table

μ	-1	0	1	2	3	4	…
g_μ	$\frac{\frac{1}{4\square}}{\frac{1}{6}}$	$\frac{2}{5}$	$\frac{1}{4\square}$	$\frac{2}{5}\times\frac{2}{7}$	$\frac{1}{4\square}\times\frac{3}{8}$	$\frac{2}{5}\times\frac{2}{7}\times\frac{4}{9}$	…

where the odd values of μ are tabulated analogously from the known

$$g_1 = \int_0^1 x^{\frac{3}{2}}(1-x)^{\frac{1}{2}} \cdot dx = \int_0^1 x^{\frac{1}{2}}(1-x)^{\frac{3}{2}} \cdot dx = f_3 = \frac{1}{2\square} \times \frac{3}{6} = \frac{1}{4\square}.$$

Immediately $g_{-1} = \int_0^1 x^{\frac{3}{2}}(1-x)^{-\frac{1}{2}} = \frac{\frac{1}{4\square}}{\frac{1}{6}} = 3 \times \frac{1}{2\square} = 3 \times f_1$, and the result follows.

It has been shrewdly conjectured[32] that such a WALLISIAN principle of induction by analogy from a set of instances played an important part in the formation of NEWTON's mathematical thought, influencing in particular the growth of his views on infinite series, curve quadrature and above all his statement of the binomial expansion[33]. Indeed, NEWTON himself, in his letter to OLDENBURG of October 24, 1676, stresses his debt to WALLIS for the inspiration which led to his formulation of the general binomial expansion, and "by similar reasoning there also came forth the ... area of the hyperbola ..."[34]. The debt becomes, however, very obvious when we consult the manuscripts on which NEWTON based his letter[35].

* In fact, $B\left(\frac{3}{2}, \frac{\lambda}{2}+1\right) = \frac{\lambda}{\lambda+3} B\left(\frac{3}{2}, \frac{\lambda}{2}\right)$.

** Or, in modern style, $B\left(\frac{5}{2}, \frac{\mu}{2}+1\right) = \frac{\mu}{\mu+5} B\left(\frac{5}{2}, \frac{\mu}{2}\right)$.

[32] By J.M. CHILD in *Newton and the art of discovery, Isaac Newton, 1642—1727*, London 1927: 117—129 especially 117—122.

[33] CHILD, *op. cit:* 117—118 "NEWTON ... was inspired to consider WALLIS' finite series as capable of bearing an intelligible meaning if they were indefinitely continued and the rest was perfectly simple and a natural consequence of what WALLIS had proved."

[34] Compare GERHARDT (*B*): **1**: 203—225, especially 203ff.

[35] Especially the undergraduate notebook *CUL Add.* 4000: 15R—22V: *Annotations out of Dr. Wallis, his arithmetica infinitorum,* with an alternative draft in *Add.* 3958: 70R—73V.

NEWTON finds[36] the hyperbola-area $(apqd)$, where the rectangular hyperbola $(1+x)\,y=1$ is defined with regard to centre c and asymptote cq, and where $cp=ap=1$, and general ordinate $dq=y$ corresponds to abscissa $cq=1+x$, as the limit-sum equivalent to $\int\limits_0^x \frac{1}{1+x}\cdot dx$, and it is to this integral that a WALLIS-type induction is applied. So far NEWTON follows WALLIS' attempt in his *AI*[37] to apply such an induction to $f(\lambda,\mu)=\int\limits_0^1 x^\lambda(1+x)^\mu\cdot dx$*, but with a flash of insight NEWTON solves the knot by generalizing the integral bounds, leaving the upper one, X, freely variable and tabulating $\Phi(\lambda)=\int\limits_0^X (1+x)^\lambda\cdot dx$ for ascending positive integral powers of λ in terms of the coefficients of $X, \frac{X^2}{2}, \frac{X^3}{3}, \frac{X^4}{4}, \ldots$ in the ensuing sequence. Thus $\Phi(2)=\int\limits_0^X (1+x)^2\cdot dx=\int\limits_0^X (1+2x+x^2)\cdot dx=1\cdot X+2\cdot\frac{X^2}{2}+1\cdot\frac{X^3}{3}$ and more generally the coefficient of X^μ/μ in $\Phi(\lambda)$ will be that of $X^{\mu-1}$ or the table of coefficients will be a "PASCAL" triangle**. By "analogy" NEWTON assumes that the pattern holds also for negative values of λ, and in particular for $\lambda=-1$ so that the general binomial coefficient $\binom{\lambda}{i}=\frac{\lambda}{1}\times\frac{\lambda-1}{2}\times\cdots\times\frac{-i+1}{i}$ becomes $\left[\binom{-1}{i}\right]=\frac{-1}{1}\times\frac{-1-1}{2}\times\cdots\times\frac{-1-i+1}{i}=(-1)^i$. Sub-

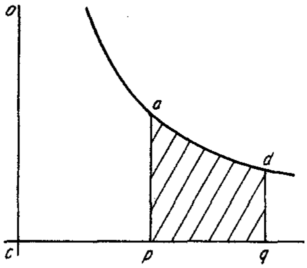


Fig. 16

	λ						
	-1	0	1	2	3	4	...
$X\times$	1	1	1	1	1	1	...
$\frac{X^2}{2}\times$	-1	0	1	2	3	4	...
$\frac{X^3}{3}\times$	1	0	0	1	3	6	...
$\frac{X^4}{4}\times$	-1	0	0	0	1	4	...
$\frac{X^5}{5}\times$	1	0	0	0	0	1	...
$\frac{X^6}{6}\times$	-1	0	0	0	0	0	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

* $f(\lambda,\mu)$ is easily calculable by multiplication and integration for positive integral λ,μ but WALLIS needed to interpolate $f(\frac{1}{2},\frac{1}{2})$—which yields the area under the hyperbola $y^2=x(1+x)$ between $x=1$ and $x=0$—and that he could not do.

** $\Phi(\lambda)=\sum\limits_{0\leq i\leq\lambda}\left[\binom{\lambda}{i}\times\frac{X^{i+1}}{i+1}\right]$ in general for λ positive integral.

36 *Add.* 4000: 20R–20V.

37 *AI:* props. 165ff.

stituting, he has immediately that hyp-area $(apqd) = \Phi(-1)$

$$= \lim_{n\to\infty} \sum_{0\leq i\leq n} \left[(-1)^i \times \frac{x^{i+1}}{i+1}\right].^{38}$$

Similarly, by interpolating $\psi(\lambda) = \int_0^x (1-x^2)^\lambda \cdot dx$ (readily calculable for λ positive integral) NEWTON finds[39] with respect to the geometrical model of a general circle segment

$$\psi(\tfrac{1}{2}) = \int_0^X (1-x^2)^{\frac{1}{2}} \cdot dx \left[= \tfrac{1}{2}(\sin^{-1}X + X(1-X^2)^{\frac{1}{2}})\right]$$
$$= \lim_{n\to\infty} \sum_{0\leq i\leq n} \left((-1)^i \binom{\frac{1}{2}}{i} \frac{X^{2i+1}}{2^{i+1}}\right).$$

Clearly, here we have two binomial expansions in integral form:

$$\int_0^X (1+x)^{-1} \cdot dx = \int_0^X \lim_{n\to\infty} \sum_{0\leq i\leq n} \left[\binom{-1}{i} x^i\right] \cdot dx,$$

and

$$\int_0^X (1-[x^2])^{\frac{1}{2}} \cdot dx = \int_0^X \lim_{n\to\infty} \sum_{0\leq i\leq n} \left[\binom{\frac{1}{2}}{i} [x^2]^i\right] \cdot dx,$$

and the way is open to abstract the binomial expansion pattern:

$$(1+\alpha)^r = \lim_{n\to\infty} \sum_{0\leq i\leq n} \left[\binom{r}{i} \alpha^i\right],$$

particularly since it agrees with the known form of the coefficients where r is positive integral.

The advance NEWTON has made on WALLIS' inductive approach to integrals —taking the upper bound of the integral variable—is that, in allowing a free variable (and its powers) into the pattern, he has been able to use the ordering of coefficients given by powers of the variable to point a more general aspect of the pattern lost in WALLIS' tabulated numerical instances. By chance, the form of these coefficients show them to be the same figurate numbers of WALLIS' function $\frac{1}{\int_0^1 (1-x^{1/\lambda})^\mu \cdot dx}$ and—as CHILD pointed out[40]—it only remained for NEWTON to rearrange WALLIS' table slightly, and make the same generalization that in the general expansion the coefficients are likewise figurate.

[38] The "MERCATOR" series for $\log(1+x)$, used, in fact, by NEWTON in the "plague" year 1665 to calculate particular logarithms to impractically large numbers of decimal places—compare *Add.* 4000: 14V: "... in summer 1665, being forced from Cambridge by the plague, I computed y^e^ area of y^e^ hyperbola at Boothby in Lincolnshire (to) two and fifty figures ..." Such detailed calculations for $x = \pm 0.1$, ± 0.2, ± 0.001, ± 0.002 to differing numbers of places are found variously in *CUL Add.* 4004: 81 R to 81 V; *Add.* 3958: Section 4, and *Add.* 4000: 20R—20V.

[39] *Add.* 4000: 18R—18V: "Having y^e^ signe of any angle to find y^e^ angle, or to find y^e^ content of any segment of a circle." See next chapter.

[40] CHILD, *op. cit.* 118ff.

With NEWTON this general scheme of interpolation by induction of a general pattern from inspection of tabulated instances shades into the general theory of infinite sequences, gradually to be replaced there by a less suggestive but tighter and more reliable basis in the theory of the integral as the limit of a sum-sequence. Indeed, uncontrolled use of induction by pattern is valuable only at a certain stage of discovery, after which its very suggestiveness and vagueness may hinder the precision of concept needed for further advance. In the mid-17th century it was not important that WALLIS should, in fact, tabulate a function more general than the one he defined, but a little later it had become supremely important that such a confusion should not be made.

The question of precision remained relatively unimportant in the theory of finite differences which evolved in the 17th century as, and has remained, an eminently practical study. It is important, however, to emphasise that the practical techniques developed were dependent on a pattern of ideas which were akin to those on which a WALLIS-type induction was based. Further, it is instructive to see how the patterning produced by the concept of n^{th}-order functional difference played an essential part in that development.

NICOLAUS MERCATOR sparked off new interest in the subject with his *logarithmotechnia* of 1668[41], showing himself familiar with the formula (derivable in an immediate way by unwrapping the differences) $e_i = e_0 + \sum_{0 \leq j \leq i} \left[\binom{i}{j} \Delta_0^j\right]$, where the differences Δ_k^i are defined in the BRIGGSIAN manner by the recursion scheme

$$\begin{cases} \Delta_\lambda^1 = e_{\lambda+1} - e_\lambda \\ \Delta_\lambda^{i+1} = \Delta_{\lambda+1}^i - \Delta_\lambda^i. \end{cases}$$

More important is how such a codification could lead to apparently unrelated mathematical results. Thus MERCATOR himself, stating an equivalent of

$$\sum_{0 \leq i \leq \lambda} \left[(-1)^i \binom{\lambda}{i} \log\left(\frac{a+ib}{a+(i+1)b}\right)\right] < 0,^\star$$

uses the formula to derive a q^{th} root approximation.[42] In particular[43] he shows

$$\sum_{1 \leq l \leq q} \sum_{0 \leq i \leq l-1} \left[\binom{\lambda-1}{i} \cdot a_i\right] = \sum_{0 \leq i \leq q-1} \left[\binom{q}{i+1} \cdot a_i\right]$$

$$\left(\text{since } \binom{q}{i+1} = \sum_{1 \leq l \leq q} \left[\binom{l-1}{i}\right]\right),$$

$$= q \times \left(a_0 + k\, a_1 + \frac{k(2k-1)}{3} a_2 + \frac{k(2k-1)(k-1)}{6} a_3 + \cdots\right),$$

★ Which has an easy proof, accessible to MERCATOR, by taking $\log\left(\frac{a+ib}{a+(i+1)b}\right) = \int_b^1 \frac{1}{a-ib+x} \cdot dx$, and reducing to a problem in hyperbola-area.

[41] MERCATOR: *logarithmotechnia*; prop. 3. Compare J. E. HOFMANN: *Nicolaus Mercator's logarithmotechnia (1668)*, Deutsche Mathematik **3** (1938): 446–466, especially 449–451.

[42] *logarithmotechnia:* props. 5–11: 15–23—compare HOFMANN, *op. cit.* 451–456.

[43] *logarithmotechnia:* prop. 7.

where $k = \frac{q-1}{2}$,

$$= q \times \left(\sum_{0 \leq i \leq k} \left[\binom{k}{i} a_i \right] + \frac{k(k+1)}{6} \left(a_2 + (k-1)\, a_3 + \frac{(k-1)(k-18)}{20}\, a_4 + \cdots \right) \right)$$

$$\approx q \times \left(\sum_{0 \leq i \leq k} \left[\binom{k}{i} a_i \right] + \frac{q^2-1}{24} \times \Delta_{k-1}^2 \right)$$

$$> q \times \sum_{0 \leq i \leq k} \left[\binom{k}{i} a_i \right].$$

Substituting $e_j = \sum_{0 \leq i \leq j} \left[\binom{k}{i} a_i \right]$, we have $\sum_{0 \leq l \leq q-1} [e_l] > q \times e_k$. Finally, taking $e_l = \log \left(\frac{a + \left(1 - \frac{2}{q} \times l\right) x}{a + \left(1 - \frac{2}{q} \times (l+1)\right) x} \right)$, we can show $\sum_{0 \leq l \leq q-1} [e_l] = \log \left(\frac{a+x}{a-x} \right) > q \times e_k = q \times e_{\frac{q-1}{2}} = q \times \log \left(\frac{aq+x}{aq-x} \right)$ or that $\left(\frac{a+x}{a-x} \right)^{\frac{1}{q}} > \left(\frac{aq+x}{aq-x} \right)$, $0 \leq x \leq a$ (with a similar proof when q is even).*

On a more practical level—an aspect which leads into GREGORY'S and NEWTON'S extensions of the finite-difference formulas—MERCATOR[44] uses the easily provable fact that, where $e_x = x^n$, Δ_λ^n is constant (and so $\Delta_\lambda^{n+1} = 0$) to build up integral powers of the integers by setting up a suitable difference table.

Conversely, NEWTON could use the convenience of logical form implicit in the difference table to tackle a problem in any way untypical of the age—the strengthening of sum-series convergence[45]. Specifically NEWTON takes his start from the (known) limit-sum $\frac{1}{1-x}$ of the geometrical progression $\lim_{n \to \infty} \sum_{0 \leq i \leq n} [x^i]$. Then, given some $\Phi(x) = \lim_{n \to \infty} \sum_{0 \leq i \leq n} [a_i \cdot x^i]$, we can transform successively by:

$$\Phi(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$$

$$= a_0 + a_1 \times \frac{x}{1-x} + (a_2 - a_1) \times \frac{x^2}{1-x} + (a_3 - a_2) \times \frac{x^3}{1-x} + \cdots$$

$$= a_0 + \left(\frac{x}{1-x} \right) (a_1 + (a_2 - a_1) \times x + (a_3 - a_2) \times x^2 + \cdots)$$

$$= a_0 + \left(\frac{x}{1-x} \right) \Big(a_1 + \left(\frac{x}{1-x} \right) \times$$
$$\times [(a_2 - a_1) + (a_3 - 2a_2 + a_1)\, x^2 + (a_4 - 2a_3 + a_2)\, x^3 + \cdots] \Big)$$

$$= \lim_{n \to \infty} \left(\sum_{0 \leq i \leq n} [\Delta_0^i] \times z^i \right), \qquad \text{where } z = \frac{x}{1-x}, \text{ and, as before,}$$

* Further approximation is possible (and given by MERCATOR) using the term $\Delta_{k-1}^2 = \Delta_{\frac{q-3}{2}}^2$.

[44] *logarithmotechnia:* prop. 12: 23—24.

[45] In various drafts of a tract *de serierum proprietatibus* (tentatively to be dated 1684) now in *CUL Add.* 3964: Section 3: 7R—20V. The method is not unlike some presented by JAMES STIRLING in his *methodus differentialis*, London, 1730: part 1: 1—83: *de summiatone serierum*, and STIRLING, indeed, explicitly attributes many of his ideas to NEWTON'S inspiration.

$$\Delta_0^i = \Delta_1^{i-1} - \Delta_0^{i-1} = \sum_{0 \leq \lambda \leq i} \left[(-1)^\lambda \binom{i}{\lambda} a_\lambda\right].^\star$$

NEWTON'S development was never published, but MERCATOR'S finite-difference technique is interesting in that it reflects how far such techniques had become accepted into conventional mathematics by the 1660's.[46] MERCATOR, however, had confined his difference-formula to interpolation at unit intervals of the function. What remained was to assume that the pattern held good universally.

This step had been taken at least by 1670 by JAMES GREGORY, developing the concept of approximating a continuous function by a power-polynomial.[47] Already in 1668 in his *exercitationes geometricae*[48] he had given some examples.

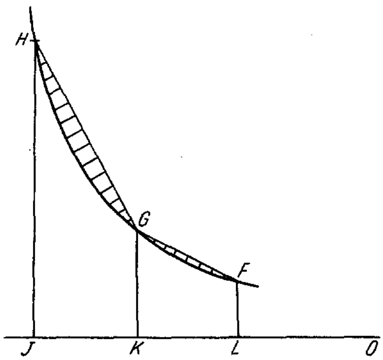


Fig. 17

For approximately equal second differences (and where the argument proceeds by unit-intervals) GREGORY assumes Δ_λ^2 constant, which restricts the form of the function to be interpolated to $y = ax^2 + bx + c$. Thus, where the ordinates $y_0(=FL)$, $y_1(=GK)$, $y_2(=HJ)$ are applied to corresponding abscissas $x_0(=OL)$, $x_1 = x_0 + h(=OK)$, $x_2 = x_0 + 2h(=OJ)$, GREGORY derives an equivalent of the SIMPSON rule: area $(FGHJL) \approx \frac{h}{3}(y_0 + 4y_1 + y_2)$. Though he gives little but a sketch proof, we easily restore his ideas.

Analytically, we have some function $y_i = f(x_0 + ih) = f(x_i) \approx a x_i^2 + b x_i + c$, and so $\Delta_\lambda^2 = \Delta_{\lambda+1}^1 - \Delta_\lambda^1 = y_{\lambda+2} - 2y_{\lambda+1} + y_\lambda = 2ah^2$. Then in the geometrical model shaded area (GH) = trapezium $(GHJK)$ — area $(GHJK)$

$$= \frac{h}{2}(y_2 + y_1) - \int_{x_1}^{x_2} (a x^2 + b x + c) \cdot dx$$

$$= \frac{a}{6}(x_2 - x_1)^3 \left(= \frac{a h^3}{6}\right) = \frac{h}{12} \Delta_\lambda^2;$$

and similarly for the shaded area (FG).

★ NEWTON'S favourite example transforms

$$\tan^{-1} y = y - \frac{1}{3} y^3 + \frac{1}{5} y^5 \cdots$$

into

$$(\tan^{-1} y =) \frac{1}{y} \times \left(\frac{y^2}{1+y^2} - \frac{2}{3}\left(\frac{y^2}{1+y^2}\right)^2 + \frac{2 \cdot 4}{3 \cdot 5}\left(\frac{y^2}{1+y^2}\right)^3 \cdots\right).$$

[46] MERCATOR, who wrote or introduced several elementary works on mathematics, in no way claims the concept of tabulated differences as his own. In fact, his discussion of the logarithmic concept in *logarithmotechnia* shows distinct traces of BRIGGS' influence.

[47] An "obvious" idea, but not to be put on a rigorous footing till WEIERSTRASS created adequate concepts of continuity.

[48] GREGORY *EG*: 25—26: *methodi componendi tabulas tangentium et secantinum artificialium ex tabulis tangentium et secantium naturalium* ...— compare GEORG HEINRICH: *Notiz zur Geschichte der Simpsonschen Regel*, Bibliotheca mathematica$_3$ **1** (Leipzig, 1900): 90—92.

Then area $(\overset{\frown}{FGHJL})$ = area $(FGHJL)$ — (shaded area (FG) + shaded area (GH))

$$= \frac{h}{2}(y_2 + 2y_1 + y_0) - \frac{2}{12} h \Delta_\lambda^2$$

and the result follows by substituting for Δ_λ^2 its value $(y_2 - 2y_1 + y_0) = \Delta_0^2$.*

By the late 1668 GREGORY seems to have become familiar with BRIGGS' work on "interpositions"[49] and certainly by November 1670 he had a completely general finite-difference interpolation formula (and apparently also its limit-form of the "TAYLOR" expansion by which he seems to have derived the general binomial theorem independently of NEWTON[50]) giving[51] the equivalent of

$$f(x_0 + x) = f(x_0) + \binom{x}{1} \Delta f^1(x_0) + \binom{x}{2} \Delta f^2(x_0) + \cdots$$

where the argument is given at unit intervals, and the variable x is left completely free. Further H.W. TURNBULL has argued[52] that GREGORY knew also the easily derivable form

$$f(x_0 + x) \left[= f\left(x_0 + \frac{x}{H} \times H\right)\right]$$

$$= f(x_0) + \binom{\frac{x}{H}}{1} \times \Delta^1_{f(x_0)} + \binom{\frac{x}{H}}{2} \times \Delta^2_{f(x_0 - H)} + \binom{\frac{x}{H}+1}{3} \times \Delta^3_{f(x_0 - H)} +$$

$$+ \binom{\frac{x}{H}+1}{4} \times \Delta^4_{f(x_0 - 2H)} + \cdots,$$

$$= \sum_{0 \leq i \leq \cdots} \left[\binom{\frac{x}{H}+i-1}{2i} \times \Delta^{2i}_{f(x_0 - iH)} + \binom{\frac{x}{H}+i}{2i+1} \times \Delta^{2i+1}_{f(x_0 - iH)} \right]. ^{**}$$

---

* GREGORY gives a tighter rule also by assuming an approximating cubic $y = a \cdot x^3 + b$.

** Where the argument is tabulated at (equal) H-length intervals.

[49] COLLINS in his letter of 30 December 1668 (see note 14 above) asks GREGORY to send him his ideas on the subject, and especially a proof of BRIGGS' results for corrected n^{th} differences in his *AL*.

[50] See next chapter.

[51] In his letter to COLLINS of 23 November 1670—*cf.* GREGORY *TV:* 117: "I remember you did once desire of me my method of proportional parts in tables, which is this . ." and states the expansion verbally with examples:

1. He takes $e_i = f(x_0 + i) = b\left(1 + \frac{d}{b}\right)^i$, so that $\Delta^k_{e_0} = b\left(\frac{d}{b}\right)^{k-1}$, and so finds the binomial expansion, where $b = 100$, $d = 6$, $e_{\frac{1}{365}} = 100 \times \left(\frac{106}{100}\right)^{\frac{1}{365}}$.

2. In an interesting generalization of MERCATOR's work in *logarithmotechnia*, GREGORY wishes to interpolate cubes of integers in the sequence $(5i)^3$, $i = 0, 1, 2, \ldots$ — that is, he takes $e_i = (5i)^3$, or $\Delta^1_{e_2} = 2375$, $\Delta^2_{e_2} = 2250$, $\Delta^3_{e_2} = 740$, $\Delta^k_{e_2} = 0$, $k > 3$, and subtabulates 23^3 by $(23)^3 = (5(2 + \frac{13}{5}))^3 = (5 \times 2)^3 + \binom{\frac{13}{5}}{1} \Delta^1_{e_2} + \binom{\frac{13}{5}}{2} \Delta^2_{e_2} + \binom{\frac{13}{5}}{3} \Delta^3_{e_2}$.

[52] In *James Gregory: a study in the early history of interpolation*, Proc. Edin. Math. Soc.$_2$ **3** (1932—1933): 151—178, arguing from examples given in an enclosure to GREGORY's letter of 23 November 1670—compare G.A. GIBSON: *James Gregory's*

However, the full working out of the theory of finite-difference interpolation is due to NEWTON, probably during the middle 1670's, and has been exhaustively described by D.C. FRASER.[53] NEWTON introduced both divided differences[54] defined on the recursive pattern of

$$\bar{\Delta}^{n+1}(x_0, x_1, \ldots, x_{n+1}) = \frac{\bar{\Delta}^n(x_0, \ldots, x_n) - \bar{\Delta}^n(x_1, \ldots, x_{n+1})}{x_0 - x_{n+1}}$$

and adjusted differences,[55] defined by the pattern

$$\Delta^{n+1}(x_0, x_1, \ldots, x_{n+1}) = \frac{\Delta^n(x_0, x_1, \ldots, x_n) - \Delta^n(x_1, \ldots, x_{n+1})}{\frac{1}{n}(x_0 - x_{n+1})},$$

and his work, especially that part of it printed in his *methodus differentialis*, formed the basis of all later elaborations.[56]

The details are too rich to summarise, but from a general viewpoint it is important to emphasise two points. First, as with GREGORY his methods all derive from taking a power-polynomial $f(x) = \sum_{0 \leq i \leq n} (a_i x^i)$ as a close approximation (for suitable choice of the a_i) to the (continuous) function to be interpolated. On that basis it is easily shown that the n^{th} divided difference $\bar{\Delta}^n$ (and so the n^{th} adjusted difference) is constant,[57] and merely by successively unwrapping the differences it is immediate that

$$\begin{aligned} f(x) &= f(x_0) + (x - x_0)\,\bar{\Delta}^1(x_0, x_1) + (x - x_0)(x - x_1)\,\bar{\Delta}^2(x_0, x_1, x_2) + \\ &\quad \cdots + (x - x_0)(x - x_1)\cdots(x - x_n)\,\bar{\Delta}^n(x_0, \ldots, x_n), \\ &= f(x_0) + \frac{(x - x_0)}{1!}\,\Delta^1(x_0, x_1) + \frac{(x - x_0)(x - x_1)}{2!}\,\Delta^2(x_0, x_1, x_2) + \\ &\quad \cdots + \frac{(x - x_0)(x - x_1)\cdots(x - x_n)}{n!}\,\Delta^n(x_0, x_1, \ldots, x_n) \end{aligned}$$

(which is a general form of the "NEWTON-GAUSS" formula where the argument intervals are unequal).

---

mathematical work ..., Proc. Edin. Math. Soc. **41** (1922—1923): 2—25. J.E. HOFMANN: *Über Gregorys systematische Näherungen für den Sektor eines Mittelpunktkegelschnittes*, Centaurus **1** (1950—1951): 24—36 has sketched how these formulas may have served to derive the approximations to central-conic sectors which GREGORY gives at great length in his *EG*: 6—8.

[53] In *Newton's interpolation formulas*, London, 1928; and his article *Newton and interpolation* in *Isaac Newton, 1642—1727*, London, 1927: 45—69.

[54] In his *methodus differentialis*, London 1711—there is a part draft (arranged in a different sequence) in *CUL Add.* 4004: 82R—84R (*MD*: Props. 1—4).

[55] In *CUL Add.* 3964: section 6: *regula differentiarum*—printed by FRASER (with translation) in his *Interpolation formulas*: 75—95.

[56] Especially those of COTES (in his *canonotechnia*, published with his *harmonia mensurarum*, London, 1722) and STIRLING (collected in his *methodus differentialis*, London 1730).

[57] Compare NEWTON's sketch proof in *MD*: prop. 1 ≡ *CUL Add.* 4004: 84R.

More important, perhaps, is NEWTON'S insistence that all the interpolation formulas should be subsumed under a single general rule,[58] and it is in this spirit of generalizing a pattern (which lies deep in the concept of interpolation) that he introduced adjusted differences as a variant on divided ones. In the scheme of adjusted differences the interpoland $f(x)$ is incorporated into the tabulated values $f(x_i)$ as a further "instance" (and what remains is to show that $\lim_{n\to\infty} \Delta^n(x_0, x_1, \ldots, x_n)$ is zero, and that the sum-sequence thus defined for $f(x)$ is convergent for a suitable range of values)*. In fact, —as NEWTON intended—all particular finite difference formulas are incorporated in the lozenge-scheme, and NEWTON must clearly have had some equivalent development in mind.[59]

In summary, the growth of the concept of interpolation is a typical aspect of the stage reached in mathematical development in the late 17th century—a stage where discovery was all-important, and where precision of the logical structures treated and justification of the methods of investigation both counted for little in comparison. It is a very practical viewpoint which sees an especial value in numerical computation—even NEWTON in his long logarithmic calculations was caught up in the tide—,and we find it equally influential in conditioning the development of the concept and technique of series expansions to which we now turn.

Appendix to IV: Fraser's lozenge diagram

(*cf.* D.C. FRASER: *Newton and interpolation, Isaac Newton, 1642—1727*: 45—69)

Taking up NEWTON'S concept of adjusted difference:

$$\Delta^{n+1}(x_0, x_1, \ldots, x_n) = \frac{\Delta^n(x_0, \ldots, x_n) - \Delta^n(x_1, \ldots, x_{n+1})}{\frac{1}{n} \times (x_0 - x_{n+1})}$$

we can show that

$$1.\ \frac{x_1 \cdot x_2, \ldots, x_n}{n!} \times \Delta^n(x_0, x_1, \ldots, x_n) + \frac{x_0 \cdot x_1, \ldots, x_n}{(n+1)!} \times \Delta^{n+1}(x_0, x_1, \ldots, x_{n+1})$$

$$= \frac{x_1 \cdot x_2, \ldots, x_n}{n!} \times \Delta^n(x_1, x_2, \ldots, x_{n+1}) + \frac{x_1 \cdot x_2, \ldots, x_{n+1}}{(n+1)!} \times \Delta^{n+1}(x_0, x_1, \ldots, x_{n+1}),$$

and

$$2.\ \frac{x_1 \cdot x_2, \ldots, x_n}{n!} \times \Delta^n(x_0, x_1, \ldots, x_n) + \frac{x_0 \cdot x_1, \ldots, x_n}{(n+1)!} \times \Delta^{n+1}(x_0, x_1, \ldots, x_n, x)$$

$$= \frac{x_1 \cdot x_2, \ldots, x_n}{n!} \times \Delta^n(x_1, x_2, \ldots, x_n, x),$$

and so set up the development given overleaf. Here all non-returning routes passing from left to right across the page yield particular interpolation formulas

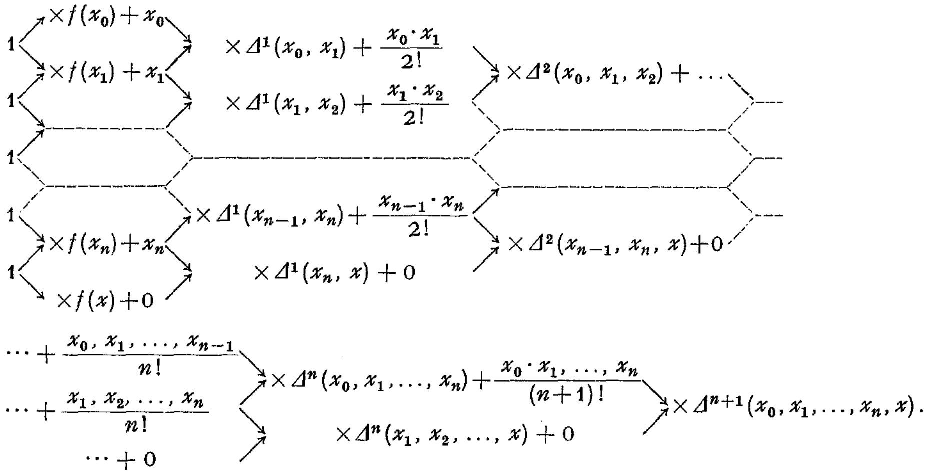
---

* A. FRASER'S "lozenge" diagram underlines the point visually. See appendix to this chapter.

[58] As he writes in his *regula differentiarum* (FRASER, *op. cit.* 82): "possunt aliae hujusmodi regulae tradi, sed mallem rem omnem una regula generali complecti, et ostendere, quomodo series quaevis in loco imperato intercalare possit".

[59] So in his *MD:* prop. 3 he derives "STIRLING'S" and "BESSEL'S" formulas as mere cases 1. and. 2. of his general divided difference expansion.

which all have the same value (specifically, that of the bottom-most route, $f(x)$).



V. Concept of function

3. Infinite series, limit-processes and convergence

The development of infinite convergent sequences is an accepted highlight of later 17th century English mathematics, and played a large role in formulating the need for a concept of limit-convergence which was not, however, adequately to be defined till the early 19th century.

This concept of a converging sequence must be very old in time in so far as the infinite sum-sequence is implicit in the theory of numerical approximation. The complex historical development of adequate notations for representing numbers, both integers and—especially in the model of a directed, calibrated line-segment—the general real, gave rise to such ideas as uniqueness and adequacy of representation.[1] In dealing with large numbers practical considerations led to the introduction of number bases, along with suitable rules for operating with such bases*—notably (and sufficiently) addition and multiplication, together with their inverses, subtraction and division. Essentially, in dealing with a large number λ, we use the property that, given some other number α, $\alpha<\lambda$, there is a unique k such that $k\alpha \leqq \lambda < (k+1)\alpha$; and we can then define a unique remainder l, given by $l=\lambda-k\alpha$ $(0\leqq l<\alpha)$. Extending the concept, for a given set of numbers α_i we can develop the sequence

$$\lambda = k_0 \times \alpha_0 + l, \qquad k_i = k_{i+1} \times \alpha_{i+1} + l_{n-i}, \qquad i = 0, 1, 2, \ldots, n-1,$$

and so, where $l_0 = k_n$, and the l_{n-i} are defined uniquely from k_i, k_{i+1}, α_{i+1} as l from λ, k_0 and α_0,

$$\begin{aligned}\lambda &= \sum_{0\leqq i\leqq n}\Big[l_i \times \prod_{0\leqq j\leqq i}(a_j)\Big] + l\\ &= l_0\cdot\langle n\rangle + l_i\cdot\langle n-1\rangle + \cdots + l_n\cdot\langle 0\rangle + l, \qquad \text{where } \langle i\rangle = \prod_{0\leqq j\leqq i}(a).\end{aligned}$$

* Commonly, in historical fact, decimal, sexagesimal and biquinary.

[1] Compare chapters 1, 2 *passim*.

The further advance implicit in the concept of place-value is that we can denote λ by the ordered set $[l_0, l_1, \ldots, l_n, l]$.*

When the required level of abstract thought is reached[2] it is natural to consider in what way meaning can be given to the unbounded sequence $l_0, l_1, \ldots, l_n, \ldots$ where n is taken unlimitedly large, and where we can define the l_i by some recursive pattern which is sufficient to generate them. In particular, the EUDOXIAN theory of number-ratio had been created to include such cases; but all such elaborations are dependent on a quantified definition—"for all i ...", or "there exists i such that ...", typically—which can have no proof in the general case though justifiable by an infinity of particular instances, and the essential arbitrariness of their introduction makes them "unnatural" and difficult to grasp. With, however, the further introduction of a concept of free variable[3] the generalization is immediate from the representation $\lambda = \sum_{0 \leq i \leq n} [l_i \times \langle i \rangle]$ to the general sum-sequence form $\lambda[x_0, x_1, \ldots, x_n] = \sum_{0 \leq i \leq n} [l_i \times \langle\!\langle i \rangle\!\rangle]$, where $\langle\!\langle i \rangle\!\rangle = \prod_{0 \leq j \leq i} [x_j]$ (with the variables x_j ranging over defined intervals); and in particular, where all the x_i are the same variable x, to the general power-polynomial form $\lambda(x)_n = \sum_{0 \leq i \leq n} [l_i \cdot x^i]$.

The first sequences so to be considered were the arithmetical and geometrical sum-series (codified in Greek times[4] as theorems in proportion theory and defined on a geometrical line-interval model), which we can represent analytically by

$$(AS)(\lambda, \mu : k)_n = \sum_{0 \leq k \leq n} (AP)(\lambda, \mu : k)$$

and

$$(GS)(L, M : k)_n = \sum_{0 \leq k \leq n} (GP)(L, M : k),$$

where (AP) and (GP) are the arithmetical and geometrical progressions $\lambda + k\mu$, $L \times M^k$ respectively; or, equivalently, in the recursive schemes $(AS)_0 = \lambda$, $(AS)_{i+1} - (AS)_i = \mu$, and $(GS)_0 = L$, $(GS)_{i+1}/(GS)_i = M$. As n increases indefinitely $(AS)_n$ is clearly unbounded, but ARCHIMEDES[5] had given a convergence criterion which showed $(GS)_n$ convergent to a limit for $|\mu| < \frac{1}{2}$ and, more generally, it was accepted that $(GS)_n$ is convergent for $|\mu| < 1$ by the medieval calculators[6] who generalised $(GS)_n$ into the form $\overline{(GS)}_n = \sum_{0 \leq k \leq n} [k \times (GP)(L, M : K)]$.

---

* Thus, in a decimal base, we take each $\alpha_j = 10$ (or $0 \leq l_k \leq 9$). The complication is, of course, that we have to introduce a zero-symbol for each $l_k = 0$.

[2] Historically, this was at least as early as the Greeks, HIPPASUS, EUDOXUS and others, who in the 5th century B.C. developed theories of such infinite number sequences to define real-number ratios. Such an advance led immediately to the distinction between actual and potential infinity and to the concept of irrational $\cdot\!\equiv\!\cdot$ "incapable of a (rational) ratio".

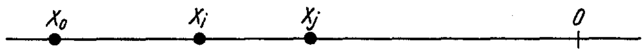
[3] The concept, while it existed verbally and defined on a convenient geometrical line-interval model from Greek times, had no adequate analytical representing notation ill the 16th century (through the invention of BOMBELLI, VIETA and others). Compare chapter 2.

[4] And so treated in EUCLID's *Elements* and by ARCHIMEDES.

[5] Compare DIJKSTERHUIS: *Archimedes*: 132—133. ARCHIMEDES applies it, of course, in his *Quadrature of the parabola* to the example $M = \frac{1}{4}$ and derives $(GS)(L, \frac{1}{4} : k)_n \to \frac{4}{3} L$, when $n \to \infty$.

[6] Especially SWINESHEAD, who seems to have introduced $(GS)_n$ in his *liber calculationum*, and the 16th century ALVARUS THOMAS (who gave a number of infinite

While it is not known how widely these medieval contributions were familiar to mathematicians of the 17th century[7], they give general credit to GREGORY ST. VINCENT for a definitive treatment of the geometrical sum-sequence[8]. Adopting GREGORY's terminology, we consider the positive (GP)-ratio $\frac{\lambda}{\mu}$ $(\lambda<\mu)$, and on the line-length X_0X_1 (as defined by fix-points X_0, X_1) we find the third fix-point O such that $X_1O:X_0O=\lambda:\mu$, and the unbounded sequence of points X_i, $i=2, 3, \ldots$, such that $X_iX_{i+1}:X_{i-1}X_i=\lambda:\mu$. Thus, for each i, we easily show $X_iX_{i+1}:X_0X_1=(\lambda:\mu)^i$, and so we can set up (GS) $(L, M:K)_n$ on the model by



$(GS)_n=X_0X_{n+1}$, where $L=X_0X_1$, $M=\lambda/\mu$, or equivalently $(GS)_n=\sum_{0\leq i\leq n}\left[X_0X_1\times\left(\frac{X_1X_2}{X_0X_1}\right)^i\right]=X_0X_1\times\sum_{0\leq i\leq n}[(\lambda/\mu)^i]$. Finally GREGORY states the equivalent of: limit-sum $(GS)_n=X_0O$.[9] In effect, since $X_0X_1:X_1X_2$ $(=\mu:\lambda)=X_0O:X_1O$, $X_0X_1:X_0O=X_1X_2:X_1O$, and we can show in general that $X_0X_1:X_0O=X_iX_{i+1}:X_iO$; therefore, since $X_0X_1<X_0O$, for all i $X_iX_{i+1}<X_iO$, with limit-equality only where X_iO can be made unlimitedly small (and this can clearly be done since the ratio $X_iO:X_0O=X_iX_{i+1}:X_0X_1=(\lambda:\mu)^i$ and $(\lambda/\mu)^i$, $\lambda<\mu$, can be made as near zero as we wish for large enough i).* A similar proof holds for the negative case where $\lambda/\mu\in[-1, 0]$,** and further, as GREGORY sketches, the whole argument is readily put into an exhaustion proof-form.

---

sum-sequences, based on $(GS)_n$ in inspiration, to some of which (by comparison with $(GS)_n$) he could give bounds only in the limit, his ingenuity not being matched by a corresponding mathematical maturity). Compare H. WIELEITNER: *Zur Geschichte der unendlichen Reihen im christlichen Mittelalter*, Bibliotheca mathematica$_3$ **14** (1913 to 1914): 150–168.

* The proof has a distinct flavour of NAPIER's derivation of his concept of logarithms by measuring on a calibrated scale the motion of a point traversing in equal times segments which are in decreasing geometrical progression.

** In a scholium[10] to his treatment GREGORY makes the first historical application of the limit geometrical progression to the "solution" of ZENO's paradox of Achilles and the Tortoise—however tempting the supposition there is no factual evidence to show that any such convergence consideration of the paradox was formulated in Greek times—and gives the now common argument that the corresponding points in the two line-length continua can be made to coincide in the limit, where the paths of Achilles and the Tortoise are traversed by points moving at proportional speeds in the same line-interval (but starting from different fix-points).

[7] LEIBNIZ had, however, studied SWINESHEAD's *liber calculationum* and possibly the (corrupt) 16th century printed edition of ORESME's tract on proportions.

[8] In his *opus geometricum*, Book 2: 51–177: *de progressionibus geometricis*. GREGORY himself admits only to classical influences—compare *OG*: 51: "Various places in ARCHIMEDES and EUCLID gave rise to this treatment ...; these particular cases tickled my imagination and led to my pondering over them seriously, and I now set out what came to me in thought ..."

[9] GREGORY expresses the concept of limit in Aristotelian terminology by *sine termino ... actu posse* ("taken unboundedly ... becomes able actually ...").

[10] GREGORY: *OG*: Book 2: prop. 78, scholium: 101–105, and compare 52. Apparently GREGORY thought out the application as a contribution to the recent revival in Belgian Jesuit circles of interest in the logical niceties of ZENO's arguments.

The straightforward analytical counterpart of this, using an algebraic free variable, was given by John Wallis a little later.[11] Wallis states that $\sum_{1\leq i\leq n}[a_0\times\lambda^i]=a_0\times\frac{1-\lambda^{n+1}}{1-\lambda}$ and proves it by a perfectly general method by "brute-force" division for a few small values of i and, though he does not explicitly give the limit form as $n\to\infty$, $|\lambda|<1$, he uses it several times in his *arithmetica infinitorum*, and indeed it is accepted by all mathematicians of the period as a standard result.

More generally, the limit-sum of the geometrical progression is a particular case of the binomial theorem:

$$(1\mp\lambda)^{-1}=\lim_{n\to\infty}\sum_{0\leq i\leq n}[(\pm\lambda)^i],\quad\text{since}\quad\binom{-1}{i}=(-1)^{12}$$

but the particular application received the name of "Mercator" division, deriving from Nicolaus Mercator's use of it to develop the "Mercator" expansion,[13] $\log(1+X)=\lim_{n\to\infty}\sum_{1\leq i\leq n}\left[(-1)^{i+1}\frac{X^i}{i}\right]$. Mercator's proof comes straightforwardly enough by defining $\log(1+X)$ as the area under the hyperbola $(1+x)\times y=1$ between $x=0$ and $x=X$, or by hyp-area $(LlmM)$ $[=\mathrm{Log}(1+X)]=\int_0^x\frac{1}{1+x}\cdot dx=\int_0^x\lim_{n\to\infty}\sum_{1\leq i\leq n}[(-1)^{i+1}\cdot x^{i-1}]\cdot dx.$[14]

[11] Wallis: *MU*: ch. 33: *progressio geometrica fusius traditur=operum mathematicorum pars prima* (1657): 38ff.

[12] As Newton pointed out in his letter to Oldenburg of 24 October 1676.

[13] Compare Mercator: *logarithmotechnia*: prop. 17: 31—33. In fact, as we now know, the Mercator technique of deriving an infinite series by straight division and then integrating term by term was used (in an equivalent form) by the 15[th] century Hindu Nilakantha in the "mandapam" constructions of the Yuktibhāsâ" (ed. Iyar & Tamparān. Trichur, 1948) which is a commentary c. 1639 on Nilakantha's *Tantrasāngraha* to derive the sum-sequence

$$\tan^{-1}z=\int_0^z\frac{1}{1+x^2}\cdot dx=z-\frac{1}{3}z^3+\frac{1}{5}z^5-\cdots=\lim_{n\to\infty}\sum_{1\leq i\leq n}\left[(-1)^{i-1}\frac{z^{2i-1}}{2i-1}\right],$$

a series found independently by Leibniz in 1673 (see J.E. Hofmann: *Entwicklungsgeschichte*...: 32—35)—together with its transform into a more rapidly converging form, and also the series expansions for sin ϑ, cos ϑ and $\sin^2\vartheta$. (Compare also various articles by C.T. Rajagopal and T.V.V. Aiyar in Scripta mathematica: **15** (1949): 201—209; **17** (1951): 65—74; **18** (1952): 25—30).

Moreover, Leibniz (in the Hanover manuscript quoted in Gerhardt (*B*) **1**: 228) gives prior discovery of the Mercator series to Johann Hudde: "Huddius mihi ostendit se jam anno 1662 habuisse quadraturam hyperbolae quam deprehendi esse illam ipsam quam Mercator quoque de suo invenit...", while Newton (see for example, *CUL Add.* 4000: 20L—20V) had the series by interpolation by 1665.

[14] The complementary $\log\left(\frac{1}{1-X}\right)=\int_x^0\frac{1}{1-X}\cdot dx=\lim_{n\to\infty}\sum_{1\leq i\leq n}\left[\frac{X^i}{i}\right]$ was found by Wallis immediately after publication of Mercator's *logarithmotechnia*. See his review of in *PT* **2** (1668): 753—759.

In fact, MERCATOR makes a very bald, loose use of indivisibles and even at the time, though the series was accepted immediately as an excellent calculating aid, there seems to have been a widespread desire for a more rigorous justification—indeed it is perhaps true to say that the MERCATOR series was accepted more because its value for $\log(1+1) = \log(2)$ was identical with that given by BROUNCKER using geometrical dissection than because of satisfaction with its logical proof-form. The loose passage to infinity in particular, introduced casually by MERCATOR, was felt to need further justification, and the remodelling of the MERCATOR proof in geometrical form by JAMES GREGORY[15] a few months later was accepted as and remained its standard derivation. GREGORY'S proof is a straightforward adaptation to the geometrical model of the hyperbola (defined by the usual asymptote property $((1+x)y=1)$ which is tightened up by being given an exhaustion-form—specifically in GREGORY'S preferred shape of "quatuor sunt igitur quantitates ..." to effect the necessary reversal of inequalities—and based explicitly[16] on GREGORY ST. VINCENT'S limit geometrical progression sum.*

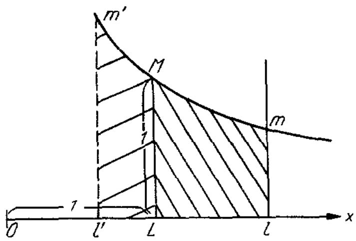


Fig. 18

In the late 1660's, however, there was a sudden proliferation of infinite sum-sequences (almost all particular logarithmic expansions) not immediately derived from a combination of limit-sums of geometrical progressions. For example, both BROUNCKER and MENGOLI had developed expansions of the logarithmic function which share with "MERCATOR'S" series the particular case, $\log(2) = \lim_{n\to\infty} \sum_{0 \leq i \leq n} \left[(-1)^{i-1} \cdot \frac{1}{i}\right]$. Above all the introduction of a whole general class of sum-sequences contained in the general binomial expansion brought with it an onrush of particular series, most approximating in the limit to particular geometrical forms—such as circle area, ellipse-length—which had proved virtually intractable (at least, on a numerical level) by previous methods. In view of the importance of the general binomial expansion in later analysis of the period, and because it typifies how a general result may bring together several independent aspects and methods, we will go into its development in some detail.

NEWTON, deservedly credited with its most general formulation[17], has tried to recapture the original train of thought which led him to the result in the

* But it is worth remarking that GREGORY in his later work never uses the geometrical hyperbola-model of the logarithmic function, preferring the analytical "logarithmus numeri" defined by the limit of a suitable sum-sequence.

[15] *EG*: part 2: 9—13: *N. Mercatoris quadratura hyperbolae geometrice demonstrata*; and compare J. E. HOFMANN: *Weiterbildung der logarithmischen Reihe Mercators in England*. Deutsche Mathematik **3** (1938): 598—605, especially 598—603.

[16] GREGORY'S prop. 1 ("si fuerint quantitates continue proportionales A, B, C, D, E, F etc. numero infinitae, quarum prima et maxima A, erit $A-B$ ad A ut A ad summam omnium") is referred for proof to GREGORY ST. VINCENT'S *Opus geometricum*.

It is striking that one of the two figures given for prop. 4 (11—12)—that for $\log(1+x)$—implicitly gives x a value *greater* than 1, which must have been very confusing to anyone trying to delimit convergence of the series expansion.

[17] Though, as we have seen in the previous chapter, BRIGGS had the particular expansion, $(1+\alpha)^{\frac{1}{2}}$, (in equivalent form) in the 1620's.

opening passages of his letter to OLDENBURG of 24 October 1676.[18] He begins by remarking on the diversity of methods by which infinite sum-series had been obtained in the past, and noting that the general binomial expansion incorporates "MERCATOR" division and "physical" root-extraction as particular cases, and then describes the birth of the ideas which led him to give the general formulation:

"At the beginning of my mathematical studies, when I had fallen upon the works of ... WALLIS, I came to consider the series through whose interpolation he develops the area of the circle and the hyperbola ...". He then sketches WALLIS' attempt to interpolate the sequence of integrals $f(\lambda)=\int\limits_0^x (1-x^2)^{\lambda/2}\cdot dx$,* $\lambda=0, 2, 4, \ldots$ by the odd values of $f(1)$, $f(3)$, ... Specifically, he multiplies out and integrates term by term to derive, for $\lambda=0, 2, 4, \ldots$,

$$f(\lambda)=\binom{\frac{1}{2}\lambda}{0}\times\int\limits_0^x 1\cdot dx+\binom{\frac{1}{2}\lambda}{1}\times\int\limits_0^x(-x^2)\cdot dx+\cdots=\sum_{0\leqq i\leqq\frac{1}{2}\lambda}\left[\binom{\frac{1}{2}\lambda}{i}\times\int\limits_0^x(-x^2)^i\cdot dx\right],$$
$$=\sum_{0\leqq i\leqq\frac{1}{2}\lambda}\left[\binom{\frac{1}{2}\lambda}{i}\times(-1)^{i+1}\frac{x^{2i+1}}{2i+1}\right],$$

where—as yet—the form of the binomial coefficients $\binom{\frac{1}{2}\lambda}{i}$ remains hidden, and they are listed only as numerical values. He begins to think out how to interpolate odd values of λ, $\lambda=1, 3, 5, \ldots$, and in particular obtain $f(1)=\int\limits_0^x(1-x^2)^{\frac{1}{2}}\cdot dx$, "which is the circle. I considered ... that the denominators $[2i+1]$ were in arithmetical progression, and so only the numeral [the binomial] coefficients remained to be investigated. But these [for even powers of λ] were the figures which represent powers of the number 1, 1 namely $(11)^0$, $(11)^1$, $(11)^2$ ..., that is, ..., 1; 1,1; 1,2,1; 1,2,3,1; 1,4,6,4,1.

"And so I sought how in these sequences, given the two first figures, the rest might be derived, and I found that, assuming the second figure to be m, the rest could be produced by continued multiplication of the terms of this sequence: $\frac{m-0}{1}\times\frac{m-1}{2}\times\frac{m-2}{3}\times\cdots$ *etc.* [and so he derives the general binomial coefficient $\binom{m}{i}=\frac{m\cdot(m-1)\cdot\cdots\cdot(m-i+1)}{1\cdot 2\cdot\cdots\cdot i}\cdots$]. So I applied this rule to interpolate the sequence ...". Thus, NEWTON supposes this binomial coefficient form to hold for intermediate values, and, in particular, uses the coefficient

$$\binom{\frac{1}{2}}{i}=\frac{\frac{1}{2}\cdot-\frac{1}{2}\cdot-\frac{3}{2}\cdots-(i-\frac{3}{2})}{1\cdot 2\cdot 3\cdot\cdots\cdot i}$$

to evaluate (on the geometrical model of the circle $y^2=1-x^2$) the area of the general circle segment, $\int\limits_0^x(1-x^2)^{\frac{1}{2}}\cdot dx$.

---

* We note that NEWTON, in being overfair to WALLIS, at the same time removes implicitly the block of thought WALLIS could not overcome, *viz:* WALLIS instinctively treated his integrals as having definite bounds, but NEWTON introduces without comment the free variable upper bound.

[18] First published in WALLIS: *opera* 3 (1699): 624ff., but I use the annotated version (based on the Hanover copy) of GERHARDT (B) 1: 203—225, especially 203 to 206.

The original manuscript on which NEWTON based this account exists in the Portsmouth Collection[19], and gives a fuller, more immediate account than NEWTON'S own statement of 1676 (which, written over ten years after, tends to touch up the crudities of the original discovery[20]). There NEWTON carries through the interpolation very elaborately, tabulating known (calculable) instances of $f(\lambda)$ very much in the style of his "model", WALLIS' methods of *arithmetica infinitorum*—specifically he justifies his generalization of the binomial coefficient $\binom{m}{i}$ to values of m other than positive integers by an argument from the logical shape of the tabulated coefficients: since the terms $(1-x^2)^{\lambda/2}$, $\lambda=0, 1, 2, \ldots$, are in geometrical proportion, their "areas $\left[\int_0^x (1-x^2)^{\lambda/2}\cdot dx\right]\ldots$ will observe some proportion amongst one another." By considering the geometrical model of the circle quadrant, he deduces that $\int_0^X (1-x^2)^{\frac{1}{2}}\cdot dx$ is the area under the circle $y^2=1-x^2$ between radius $Oc=1$ and the parallel half-chord $ba=(1-X^2)^{\frac{1}{2}}$, where $Oa=X$. Therefore area $(Oabc)$ $\left[=\frac{1}{2}\cdot\sin^{-1}X+\frac{1}{2}\cdot X\cdot(1-X)^{\frac{1}{2}}\right]$ $=\lim_{n\to\infty}\sum_{0\leq i\leq n}\left[(-1)^i\cdot\binom{\frac{1}{2}}{i}\cdot\frac{X^{2i+1}}{2i+1}\right]$.*

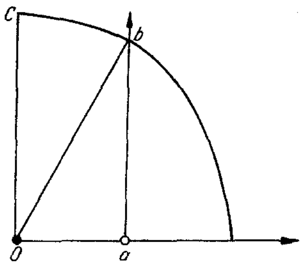


Fig. 19

The corresponding interpolation for the hyperbola-area $\int_0^x (1+x^2)^{\frac{1}{2}}\cdot dx$ is derived in a similar way[21]: "By the same method, again, the interpolated areas of the other curves $[f(\lambda), \lambda$ odd$]$ are forthcoming, as also the area of the hyperbola and other alternate terms in this series ... $\Big[$whose general term he takes by $g(\mu)=\int_0^x (1+x^2)^{\mu/2}\cdot dx, \mu=-1, 0, 1, 2, \ldots\Big]$... This was my first entrance into these speculations ... ."

"But when I had obtained these results, I soon began to see that the terms $(1-x^2)^{\frac{0}{2}}$, $(1-x^2)^{\frac{2}{2}}$, $(1-x^2)^{\frac{4}{2}}$, $(1-x^2)^{\frac{6}{2}}$ could be interpolated in the same way as the areas they generate; and for this nothing more was necessary than the

* From this NEWTON derives his series for $\sin^{-1}X$ by $\sin^{-1}X=2\times$ area $(Oabc)-X(1-X^2)^{\frac{1}{2}}$, expanding the right side into an infinite sum-sequence.

[19] *CUL Add.* 4000: 18R—19V: "Having y^e^ signs of any angle to find y^e^ angle, or to find y^e^ content of any segment of a circle", with a draft in Add. 3958: 70—73.

[20] His notation, in particular, is strongly WALLISIAN in flavour. So he defines the general binomial coefficient $a_i=\binom{\frac{1}{2}}{i}$ which is to be inserted in the expansion of

$$\int_0^1 (1-x^2)^{\frac{1}{2}}\cdot dx=\lim_{n\to\infty}\sum_{0\leq i\leq n}\left[(-1)^i a_i\frac{x^{2i+1}}{2i+1}\right],$$

"yt is $\frac{0}{0}x-\frac{0}{0}\times\frac{1}{2}\times\frac{1}{3}x^3+\frac{0}{0}\times\frac{1}{2}\times\frac{1}{4}\times\frac{1}{5}x^5-\frac{0}{0}\times\frac{1}{2}\times\frac{1}{4}\times\frac{3}{6}\times\frac{1}{7}x^7\cdots$. This progression may be deduced from hence $\frac{0}{0}\times\frac{1}{2}\times\frac{-1}{4}\times\frac{3}{6}\times\frac{-5}{8}\times\frac{7}{10}\times\cdots$." The initial coefficient $a_0=$ "$\frac{0}{0}$" $(=1)$ is straight out of WALLIS' *AI*.

[21] Given more fully in the previous chapter.

omission of the denominators 1, 3, 5, 7 *etc.* in the terms expressing the areas ... $\left[\text{which are } (-1)^i \cdot \binom{\frac{1}{2}}{i} \cdot \frac{x^{2i+1}}{2i+1}\right]$ .... That is, the coefficients of the terms of the quantity to be interpolated ..., in general, $(1-x^2)^m$ arise from the continuous multiplication of the terms of this sequence,

$$\frac{m}{1} \times \frac{m-1}{2} \times \frac{m-2}{3} \times \frac{m-3}{4} \times \cdots \left[\cdot \equiv \cdot \binom{m}{i}\right]."$$

This is a most interesting point—NEWTON had to derive the binomial expansion in an integral form, $\int_0^x (1-x^2)^{\frac{1}{2}} \cdot dx$, before noticing that the same form is preserved in $(1-x^2)^{\frac{1}{2}}$ if we multiply each power of x, x^{2i+1}, by $\frac{2i+1}{x}$ (and so obtain the derivative, $(1-x^2)^{\frac{1}{2}}$, from the integral, $\int_0^x (1-x^2)^{\frac{1}{2}} \cdot dx$). It is significant, however, that NEWTON, aware that derivation by such a loose method of pattern-analogy may not be rigorous, checked the particular expansions $(1+x)^{-1}$, $(1-x^2)^{\frac{1}{2}}$ as equivalent, term by term, with the sequences arising from dividing and extracting the square root respectively in the standard way: but finally "when I had very clearly seen through these results, I ignored completely (WALLIS') interpolation of sequences, and applied only these operations as being more truly fundamental *[tamquam fundamenta magis genuina]*".

NEWTON, of course, did not (or more accurately perhaps could not[22]) publish this—in fact, though his general method was widely circulated in his (1669) *de analysi*, no account of it appeared in a printed text till 1685[23]. Meanwhile both MERCATOR[24] and BROUNCKER[25] had apparently rediscovered the BRIGGS expansion of $(1+\alpha)^{\frac{1}{2}}$, though all details of how or what they did seem to have

[22] His correspondence over the years 1671—1676 shows him trying desperately hard to have his research published either independently or in appendix to the projected English edition of KINCKHUYSEN's *Algebra*, but it appears that no publisher would print it—understandable if we remember that no advanced English mathematical texts at the time could command a sufficient audience to yield a profit unless the book were to be priced prohibitively high.

[23] When WALLIS, in his *Algebra*: ch. 91, gave an (adapted) extract from NEWTON's letter of 24 October 1676: and when JOHN CRAIG, in his *methodus figurarum lineis rectis et curvis comprehensarum quadraturas determinandi*, used particular examples of the binomial expansion "secundum methodum celeberrimi D. Isaaci Newtoni", for example, in his prob. 12: 14—15: *circuli quadraturam determinare*, where he gives the expansion of $r \cdot \left(1-\left(\frac{y}{r}\right)^2\right)^{\frac{1}{2}}$. DAVID GREGORY in his *exercitatio geometrica* of 1684: 19—21 (a work published specifically to give a permanent form to results derived by his uncle JAMES GREGORY) uses the expansions, $(1+\alpha)^{\frac{1}{2}}$, $(1+\alpha)^{\frac{1}{3}}$, but derives them by physically extracting the square and cube roots respectively.

[24] Compare COLLINS-GREGORY, 7 Jan 1668/9 ($\cdot \equiv \cdot$ GREGORY *TV*: 60) "Mr. Mercator hath often ... affirmed with much confidence that he hath now a series for the circle that shall make the sines of any arch and the converse, and give the area of any sector, segment or zone infinitely true". We know that MERCATOR and NEWTON corresponded in the 1670's (see NEWTON *PM* Book 3: prop. 17, theorem 15), and it would be interesting to know if the topic was introduced.

[25] Compare COLLINS-GREGORY, 2 Feb 1668/9 (GREGORY *TV*: 66): "... the Lord Brouncker asserts he can turne the square roote into an infinite series ..."

vanished. More important, DAVID GREGORY[26] asserted later, in sketching the evolution of the mathematical thought of his uncle, JAMES GREGORY, that JAMES had found the binomical expansion independently of NEWTON, "huic rei... intentus". Indeed, JAMES gave the binomial expansion in 1670 in its general (logarithmic) form:[27]

$$\log b + \frac{a}{c}\left[\log(b+d) - \log(b)\right]\left(= \log\left[b\left(1+\frac{d}{b}\right)\right]^{\frac{a}{c}}\right)$$
$$= \log\left[b\cdot\sum_{1\leq i\leq n}\left[\binom{a/c}{i}\cdot\left(\frac{d}{b}\right)^i\right]\right],$$

where n may be taken indefinitely great, having apparently derived it by use of his finite-difference interpolation formula (for unit-differences of the argument),

$$f(x_0+h) = f(x_0) + \binom{h}{1}\Delta^1 f(x_0) + \binom{h}{2}\Delta^2 f(x_0) + \cdots.$$

If this was GREGORY's derivation, he was on far firmer ground than NEWTON, who had derived the binomial expansion merely by noticing and formulating a general pattern which seemed to run through a sequence of particular expansions, and who could only justify such a generalisation by checking its consistency with results to be had by other procedures, particularly root-extraction—unfortunately, no convenient numerical p/q^{th} root extraction process existed which could check the general expansion of $(1+\alpha)^{p/q}$. GREGORY's derivation is more fundamental, and makes the binomial expansion only a particular case of a general (finite-difference) theorem [even though, very probably, he could give no better proof for it than NEWTON for his development—that is, by inducing a general law by analogy with particular (computed) instances]. Above all, the GREGORY approach is highly suggestive, leading straight to the formulation of the general "TAYLOR" expansion[28]—which is the limit form of the general finite-difference

---

[26] *op. cit.*, note [23].

[27] In an enclosure to his letter to COLLINS of 23 November 1670 ·≡· GREGORY *TV*: 131–132. The statement, given without any indication of proof, is followed immediately by an example, where $b = 100$, $d = 6$, $a = 1$, $c = 365$, and so

$$b\left(1+\frac{d}{b}\right)^{\frac{a}{c}} = 100\left(\frac{106}{100}\right)^{\frac{1}{365}},$$

which is treated by his general interpolation formula—compare GREGORY *TV*: 119–120. Taking $f(x_0+h) = b\left(1+\frac{d}{b}\right)^h$, then $hf(x_0) = b$ and $\Delta^i f(x_0) = b\left(\frac{d}{b}\right)^i = \frac{d^i}{b^{i-1}}$, and the binomial expansion is immediate.

[28] H.W. TURNBULL has, indeed, argued very plausibly that GREGORY was already using such an expansion by 1672. Compare GREGORY *TV*: 356ff. TURNBULL bases his argument on elaborate calculations for series expansions made on the back of GIDEON SHAW's letter to GREGORY of 29 January 1671: (p. 356): "... these sixteen mathematical items on this double-sheeted manuscript reveal the workings of a mind upon which the importance of a certain mathematical principle was dawning—the principle of successive differentiation ..."; and again (p. 357): "... Gregory was familiar with (the Taylor expansion) in the sense that he applied this rule to a wide variety of trigonometrical and logarithmic functions. In contrast to his interpolation formula ..., which he explicitly stated in general form in his letter to Collins of 23 November 1670, the Taylor series occurs only in applications, [but, if we deny that Gregory had found the Taylor expansion, we are] faced with the puzzling question how to account for the wealth of applications of a complicated theorem if the theorem itself were unknown to Gregory."

formula,

$$f(x_0+h)=f(x_0)+\binom{\frac{h}{H}}{1}\Delta^1 f(x_0)+\binom{\frac{h}{H}}{2}\Delta^2 f(x_0)+\cdots\left[=f\left(x_0+\frac{h}{H}\cdot H\right)\right]$$

(where the argument is given at equal H-intervals), and BROOK TAYLOR derived his expansion on that basis.[29] In fact, remodelling the finite-difference formula and assuming Δx-intervals,

$$\begin{aligned}f(x_0+h)&=f\left(x_0+\frac{h}{\Delta x}\cdot\Delta x\right)\\&=f(x_0)+\frac{h}{1!}\,\frac{\Delta f^1(x_0)}{\Delta x}+\frac{h\cdot h_1}{2!}\,\frac{\Delta f^2(x_0)}{(\Delta x)^2}+\frac{h\cdot h_1\cdot h_2}{3!}\,\frac{\Delta f^3(x_0)}{(\Delta x)^3}+\cdots\end{aligned}$$

where

$$h_i=\Delta x\left(\frac{h}{\Delta x}-i\right)=h-i\,\Delta x$$

and so, using $\lim\limits_{\Delta x\to 0}\Delta f^i(x)=0$, $\lim\limits_{\Delta x\to 0}(h_i)=h$, and

$$\begin{aligned}\lim_{\Delta x\to 0}\left[\frac{\Delta f^i(x_0)}{(\Delta x)^i}\right]&=\lim_{\Delta x\to 0}\left[\frac{\Delta f^{i-1}(x_0+\Delta x)-\Delta f^{i-1}(x_0)}{(\Delta x)^i}\right]=\frac{d}{dx}\left[\lim_{\Delta x\to 0}\frac{\Delta f^{i-1}(x_0)}{(\Delta x)^{i-1}}\right]\\&=\frac{d^i}{dx^i}\left(f(x_0)\right)=f^{[i]}(x_0),\end{aligned}$$

the Taylor expansion,

$$f(x_0+h)=\lim_{\Delta x\to 0}\left(f\left(x_0+\frac{h}{\Delta x}\cdot\Delta x\right)\right)=f(x_0)+\frac{h}{1!}f^{[1]}(x_0)+\frac{h^2}{2!}f^{[2]}(x_0)+\cdots,$$

is immediate.

However, when the binomial expansion had been accepted into mathematics, the way was clear for the production of an enormous number of particular sum-series. Beginning with the letters of GREGORY to COLLINS in the 1670's and the circulated NEWTON manuscript *de analysi*, there came forth a bewilderingly rich and complex collection—series for the lengths of ellipses, zones of circles, for trigonometrical and corresponding inverse functions. By their immediacy and constructibility infinite sum-sequences fired the imagination of the lesser mathematicians even more than the great few. Even NEWTON could be caught up in it all[30]: even when later, as an old man, he fell out of love with sheer numerical computation, he put infinite sum-sequences at the very basis of his mathematical method. Significantly, in the fluxional controversy he refused to allow LEIBNIZ

[29] In his *methodus incrementorum directa et inversa*, London, 1715, 21 ff.—compare A. PRINGSHEIM: *Zur Geschichte des Taylorschen Lehrsatzes*, Bibliotheca mathematica$_3$ **1** (1900—1901): 433—484, especially 433ff.

[30] As I have said above, the NEWTON manuscripts contain many drafts of logarithmic calculations, admittedly written in extreme youth, —*cf. CUL Add.* 3958: 77 Rff.; 4000: 20 Rff.; 4004: 81 Rff.—of which NEWTON could say in his letter to OLDENBURG of 24 October 1676: "I am ashamed to say to how many places of figures I carried through these computations, having then a great deal of leisure. For then, indeed, I took an excess of pleasure in these findings ..." (GERHARDT (*B*) **1**: 207). In fact, NEWTON's computations are rounded off variously at 47D—57D, the calculations themselves often filling a whole manuscript sheet for each case.

to separate the two methods of infinite series and fluxions because they were —for him, at least—inextricably involved with each other: nor is this an overstatement calculated to win support in the controversy—for NEWTON infinite series and fluxions became a single analytical method on which all analysis of the infinite is to be based.[31]

And there it rested for the 17th century English mathematician who, while he could marvel (on a numerical level) at the accuracy and flexibility of the infinite sum-sequence, would be therefore largely unconcerned with such theoretical functional considerations as uniqueness, periodicity and limit-convergence. The later 17th century was truly a period of frontier expansion in mathematical analysis when everything must bow to that felt need for widening factual knowledge: there were such rich vastnesses of virgin territory to be explored that, when and if the way became in any wise difficult, there was greater immediate profit to be had by changing direction towards an easier terrain than by carrying on through the roughnesses of obscurity and complexity.*

But in the mid 17th century before the flood of infinite series developments broke on the mathematical world, bringing with it a tidal wave of uncritical ideas, serious attempts had been made to formulate the concept of sequence on a strict basis and to set up concepts of (and indeed tests for) convergence.[33]

Let us return once more to the logarithm (and its geometrical model of hyperbola-area) to consider the point.

PIETRO MENGOLI, as we have seen[34], took his inspiration from the model of the area under $xy=1$, deriving therefrom sufficient defining conditions to allow

---

* It is entirely typical, for example, that NEWTON does not answer LEIBNIZ' serious reflection in 1677 that the transform of $f(x, y)=0$ into the explicit $y=g(x)$ (with real coefficients) cannot give imaginary roots of f, since $g(x)$, x real, converges to a real limit.[32]

[31] Compare his remark in JOSEPH RAPHSON'S *History of fluxions*, London, 1714 · ≡ · GERHARDT (*B*), 1, 287: "In my letter of the 13th of June 1676 I said that my method of series extended to almost all problems, but became not general without some other methods, meaning ... the method of fluxions and the method of arbitrary series [*sc.* NEWTON'S method, an improvement on VIETA'S, of extracting the explicit limit polynomial expansion $y=g(x)$ from the implicitly given $f(x, y)=0$ by substituting and comparing coefficients—to be equated to zero for each power of x—in $f(x, g(x))=0$] and now to take those other methods from me is to restrain and restrict the method of series, and make it cease to be general. In my letter of October 24 1676 I called all these methods together my general method." We can see NEWTON'S ideal worked out in some detail in *CUL Add.* 3960: Section 14 (to be dated about 1670), the tract printed as *geometria analytica* in S. HORSLEY: *Newtoni opera.* 1: 389—519.

[32] See LEIBNIZ' letter to OLDENBURG of 12 July 1677, GERHARDT (*B*) 1: 248—249.

[33] It is very tempting to equate the disappearance of such rigorous considerations with the sudden outpouring of the shakily-based series developments. Perhaps the sheer numerical weight of these new series expansions cheapened their individual value for the mathematician. Before, one forced out a particular expansion only with great mental labour and therefore did not leave it in a rough state, but polished it, tightened it up, defined convergence conditions, related it to known results. A further factor, however, must be that till the 1670's functions were very largely defined with respect to a suitable geometrical model—on such a well-tested and so strongly visual basis certain restraints of rigour must be automatically applied which have to be formulated explicitly in an analytically equivalent structure.

[34] Compare chapter 3.

a purely abstract, analytical treatment. Probably it was an attempt to apply an analytical convergence condition to $\int_1^x \frac{1}{x} \cdot dx$ which led him to consider the limit as n increases indefinitely of $\sum_{1 \leq i \leq n} (1/i)$. By an ingenious grouping he was able to show that the sum increases indefinitely with n,[35] but his treatment has an air of a trick well-done about it.*

Brouncker, however, in his consideration of convergence of the various sum-sequences he had developed for $\log(2)$,[36] on the basis of the same model of the area under the hyperbola $xy=1$ introduced a more obvious and more general technique. Having shown that

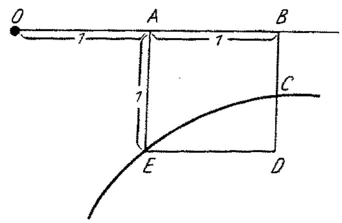


Fig. 20

$$\log(2) = \text{area}\,(ABCE) = \text{limit}\left(\frac{1}{1\cdot 2} + \frac{1}{3\cdot 4} + \frac{1}{5\cdot 6} + \frac{1}{7\cdot 8} + \cdots\right),$$

and similarly, that

$$1 - \log(2) = \text{area}\,(CDE) = \text{limit}\left(\frac{1}{2\cdot 3} + \frac{1}{4\cdot 5} + \frac{1}{6\cdot 7} + \cdots\right),$$

he states the sufficient condition for the convergence of each (monotonically increasing) sequence that limit $\left[\left(\frac{1}{1\cdot 2} + \frac{1}{3\cdot 4} + \cdots\right) + \left(\frac{1}{2\cdot 3} + \frac{1}{4\cdot 5} + \cdots\right)\right] = 1$, —a condition immediately derivable from the model, since area $(ABCE)+$ area$(CDE)=$area$(ABDE)$—and shows it true by splitting the general terms of the two series $\frac{1}{(2i-1)\,2i}$ and $\frac{1}{2i\,(2i+1)}$ into the part-fractions $\left(\frac{1}{2i-1} - \frac{1}{2i}\right)$ and $\left(\frac{1}{2i} - \frac{1}{2i+1}\right)$ respectively, so that

$$\lim_{n\to\infty} \sum_{1\leq i\leq n} \left[\frac{1}{(2i-1)\,2i} + \frac{1}{2i\,(2i+1)}\right] = \lim_{n\to\infty} \sum_{1\leq i\leq n} \left[\frac{1}{2i-1} - \frac{1}{2i+1}\right]$$
$$= \frac{1}{1} - \lim_{n\to\infty}\left(\frac{1}{2n+1}\right),$$

which tends to 1. Abstracting his convergence criterion from this, Brouncker has, in effect, two sequences (a_i), (b_i), where $a_i < A$, $b_i < B$ for all i, and states that $\lim_{i\to\infty}[(A+B)-(a_i+b_i)]=0$ is sufficient for $\lim_{i\to\infty}(a_i)=A$ (and $\lim_{i\to\infty}(b_i)=B$).

* Much as James Bernoulli in his independent rediscovery of the divergence used the inequality $\left(\frac{1}{2i-1}+\frac{1}{2i}\right) > \frac{2}{2i}\left(=\frac{1}{i}\right)$ Mengoli uses $\left(\frac{1}{a-1}+\frac{1}{a}+\frac{1}{a-1}\right) > \frac{3}{a}$; then grouping successively by threes, he derives $1+\left(\frac{1}{2}+\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}\right) + \cdots > 1+\frac{3}{3}+\left(\frac{3}{6}+\frac{3}{9}+\frac{3}{12}\right)+\cdots > 1+\frac{3}{3}+\frac{9}{9}+\frac{27}{27}+\cdots$.

[35] In his *novae quadraturae arithmeticae, seu de additione fractionum* ..., Bologna, 1650.

[36] Compare chapter 3.

BROUNCKER finds the extension to his more complex method of approximating by triangles less easy. Thus, he had developed the limit sum-sequence,

$$\text{area}\,(CDE) = \lim_{n\to\infty} \sum_{0\leqq r\leqq n} \sum_{1\leqq s\leqq 2^r-1} \left[\frac{1}{(\lambda-2)(\lambda-1)\lambda}\right],$$

where

$$\lambda = 2^r + 2s.$$

Clearly the sequence is monotonically increasing, and further, from the geometrical model, obviously the successive triangulations all lie inside area (CDE), since the hyperbola is everywhere convex in the range $x\in[1,2]$, but how is the convergence of the sequence in the limit to area (CDE) to be shown?—specifically, where $\mu_i = \sum_{0\leqq r\leqq i} \sum_{1\leqq s\leqq 2^r-1} \left[\frac{1}{(\lambda-2)(\lambda-1)\lambda}\right]$, how shall we prove $\lim_{i\to\infty}[\text{area}\,(CDE) - \mu_i] = 0$?

BROUNCKER'S solution develops an ingenious test, using the limit-sum of a geometrical progression as a comparison sequence.[37] Using the inequality

$$\frac{1}{4}\cdot\frac{1}{(a-2)\cdot(a-1)\cdot a} < \left[\frac{1}{(2a-2)(2a-1)\,2a} + \frac{1}{(2a-4)(2a-3)(2a-2)}\right]$$

we can show $\frac{1}{4}\bar\mu_i < \bar\mu_{i+1}$, and more generally $\frac{1}{4k}\bar\mu_i < \bar\mu_{i+k}$, where $\bar\mu_n = \mu_n - \mu_{n-1}$; so that

$$\lim_{j\to\infty}(\mu_j) = \mu_{n-1} + \lim_{k\to\infty} \sum_{n\leqq i\leqq k} (\bar\mu_i)$$
$$< \mu_{n-1} + \bar\mu_n \cdot \lim_{k\to\infty}\Big[\sum_{0\leqq\lambda\leqq(k-n)} [(\tfrac{1}{4})^\lambda]\Big] = \mu_{n-1} + \tfrac{4}{3}\bar\mu_n.$$

Clearly, this gives him an estimate for the error at the n^{th} term. Further, giving a very sketchy justification, BROUNCKER assumes in "WALLISian" manner (by inducing from numerical instances) that $\frac{\bar\mu_{n+1}}{\bar\mu_n} < \frac{\bar\mu_n}{\bar\mu_{n-1}}$ for all n, so that $\frac{\bar\mu_{n+\lambda}}{\bar\mu_n} < \left(\frac{\bar\mu_n}{\bar\mu_{n-1}}\right)^\lambda$. Finally

$$\lim_{k\to\infty}\left(\frac{\mu_k-\mu_{n-1}}{\bar\mu_n}\right) < \lim_{m\to\infty}\sum_{1\leqq\lambda\leqq m}\left[\left(\frac{\bar\mu_1}{\bar\mu_{n-1}}\right)^\lambda\right] = \frac{\bar\mu_n}{\bar\mu_{n-1}} \times \left(\frac{1}{1-\dfrac{\bar\mu_n}{\bar\mu_{n-1}}}\right)$$

or

$$\lim_{k\to\infty}(\mu_k) < \left[\mu_{n-1} + \frac{\bar\mu_n^2}{\bar\mu_{n-1}} \times \left(\frac{1}{1-\dfrac{\bar\mu_n}{\bar\mu_{n-1}}}\right)\right],$$

from which a second estimate for the error at the n^{th} term can be given. Together, the two attempts to use a comparison series are highly ingenious, and—despite the unjustified (but justifiable) assumption that $\bar\mu_{n+1}/\bar\mu_n$ decreases with increasing n—more soundly based than any later 17th century convergence investigation of a limit sum-sequence.

[37] Unconsciously following DESCARTES who had used such a device in treating the convergence of his method of isoperimetries—see *excerpta ex manuscriptis* ... in *opuscula posthuma, physica et mathematica*, Amsterdam 1701: pt. 6, no. 5. ·≡· DESCARTES: *Oeuvres* (Ed. ADAM & TANNERY) **10** Paris, 1908: 304; and compare EULER: *annotationes in locum quendam Cartesii ad circuli quadraturam spectantem*. Novae comm. Ac. sc. Petrop **8** (1760—1761): 157—168. ·≡· *opera* **15 1**, Berne, 1927: 1—15.

The sum-sequence, of course, represents the vast bulk of limit-sequences considered in the 17^{th} century. But, as we have seen, BROUNCKER had developed a general series of continued-fraction sequences[38], while product-sequences were not unknown. The supreme example of the latter in the period is WALLIS' product for $\frac{1}{2}\pi$, but an interesting case occurs in a letter of GREGORY to COLLINS[39], which in fact generalizes the well-known sequence, first given by VIETA,

$$\frac{2}{\pi} = \frac{\left(\sin\frac{\pi}{2}=\right)1}{\lim\limits_{i\to\infty}\left(2^i\sin\frac{\pi}{2^{i+1}}\right)} = \lim_{n\to\infty}\prod_{1\leq i\leq n}\left(\cos\frac{\pi}{2^{i+1}}\right)$$

$$= \sqrt{\tfrac{1}{2}}\times\sqrt{\tfrac{1}{2}+\tfrac{1}{2}\sqrt{\tfrac{1}{2}}}\times\sqrt{\tfrac{1}{2}+\tfrac{1}{2}\sqrt{\tfrac{1}{2}+\tfrac{1}{2}\sqrt{\tfrac{1}{2}}}}\times\cdots\text{*}$$

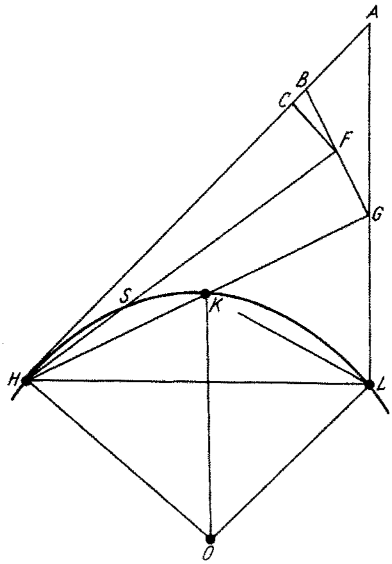


Fig. 21

Taking the general circle arc $\widehat{HKL}\ (<\pi)$, where AH, tangent at H meets AL, perpendicular to HL, in A; HK, through K (bisector of arc HL) meets AL in G, and GB is perpendicular to HG; HS, through S (bisector of arc HK) meets GB in F, and FC is perpendicular to HF; and so on through successive stages, we clearly have an operation-sequence which defines points $A, B, C\ldots$ successively on HA. It is obvious also that the limit-point λ defined in HA is such that $H\lambda = \widehat{HKL}$, since if $\widehat{AHL} = \delta = \widehat{HOK}$, then successively $\widehat{BHG} = \vartheta/2$, $\widehat{CHF} = \vartheta/2^2, \ldots$ and again, $HG\ (= HK + KL) = 2\cdot HK$, $HF = 2^2\cdot HS, \ldots$; so that

$$H\lambda = HO\times\lim_{n\to\infty}\left(2^{n+1}\times\sin\frac{\vartheta}{2^n}\right) = HO\times\lim_{\lambda\to 0}\left(\frac{\sin\lambda\vartheta}{\lambda\vartheta}\right)\times 2\vartheta \quad\text{where}\quad \lambda = \frac{1}{2^n}.$$

Further,

$$HG:HL = \sec\frac{\vartheta}{2},\quad HF:HG = \sec\frac{\vartheta}{2^2},\quad HC:HF = \sec\frac{\vartheta}{2^3},\ \cdots$$

and so

$$H\lambda = \lim_{m\to\infty}\prod_{1\leq\tilde{\lambda}\leq m}\left(\sec\frac{\vartheta}{2\tilde{\lambda}}\right)\times 2\,HO\sin\vartheta\text{**}$$

---

* Using the recursive scheme $\cos\frac{\pi}{4} = \sqrt{\frac{1}{2}}$, $\cos\left(\frac{\pi}{2^i}\right) + 1 = 2\cos^2\left(\frac{\pi}{2^{i+1}}\right)$.

** Since $HL = 2HO\sin\widehat{HOK}$.

[38] See chapter 2. No further continued-fraction limit-sequences were considered in the 17^{th} century, though ROGER COTES developed empirically the continued-fraction expansion of e: $e = (2, 1, 2;\ 1, 1, 4;\ 1, 1, 6;\ 1, 1, 8;\ \ldots)$ in his ***harmonia mensurarum, sivè analysis et synthesis per rationum et angulorum mensuras promotae***, Cambridge 1722: 7. EULER, of course, was to develop general techniques of examination in numerous papers spread fairly evenly throughout his life, but especially in the late 1730's.

[39] See GREGORY's letter to COLLINS, 15 February 1668/9, GREGORY *TV*: 68—70, especially 68—69, and compare CHRISTOPH J. SCRIBA: *James Gregorys frühe Schriften zur Infinitesimalrechnung* ·≡· Mitteilungen aus dem Mathem. Seminar Gießen. Heft 55: 65ff.

or, on reduction (by eliminating $H\lambda$),

$$\frac{\vartheta}{\sin\vartheta} = \lim_{m\to\infty} \prod_{1\leq\tilde{\lambda}\leq m} \left(\sec\frac{\vartheta}{2\tilde{\lambda}}\right).$$ [40]

This result is connected at a deep level with the convergent analytical sequences derived by GREGORY in his *VCHQ*[41] which are defined recursively from i_0, I_0; $i_{k+1} = (GM)\,(i_k, I_k)$, $I_{k+1} = (HM)\,(i_{k+1}, I_k)$. The convergence of these sequences is obvious from the particularised geometrical models given by GREGORY of the general sector of a central conic, which are parametrisable:

$$\text{(ellipse)} \qquad i_k = 2^{k-1}\sin\frac{\vartheta}{2^{k-1}}, \qquad I_k = 2^k\tan\frac{\vartheta}{2^k},$$

$$\text{(hyperbola)} \qquad i_k = 2^{k-1}\sinh\frac{\vartheta}{2^{k-1}}, \qquad I_k = 2^k\tanh\frac{\vartheta}{2^k}.$$

But in fact, GREGORY develops in *VCHQ*[42] a proof which shows convergence for *any* i_0, I_0 (and I think he intended deliberately to make his analysis general and independent of any particular model). Specifically GREGORY, setting up his two sequences (i_n), (I_n) in parallel columns, used the inequality $(i_{n+1} - i_n) < 4(i_{n+2} - i_{n+1})$ to compare convergence of (i_n) with the limit-sum of a geometrical progression

$$\begin{array}{cc} i_0 & I_0 \\ i_1 & I_1 \\ \vdots & \vdots \\ & (I)\,. \end{array}$$

In proof we have:

$$\frac{i_{n+1}-i_n}{I_n-i_{n+1}} = \frac{i_n}{i_{n+1}}, \qquad \text{since } i_{n+1}^2 = i_n \cdot I_n;$$

$$\frac{I_n-i_{n+1}}{I_{n+1}-i_{n+1}} = \frac{i_n+i_{n+1}}{i_n}, \qquad \text{since } I_{n+1} = \frac{2\,i_{n+1}\cdot I_n}{i_{n+1}+I_n} = \frac{2\,i_{n+1}^2}{i_n+i_{n+1}}$$

and

$$\frac{I_{n+1}-i_{n+1}}{i_{n+2}-i_{n+1}} = \frac{i_{n+2}+i_{n+1}}{i_{n+1}},$$

since $i_{n+2}^2 = i_{n+1} \cdot I_{n+1}$; so that, multiplying these ratios

$$\frac{i_{n+1}-i_n}{i_{n+2}-i_{n+1}} = \frac{(i_n+i_{n+1})\,(i_{n+1}+i_{n+2})}{i_{n+1}^2},$$

which GREGORY shows to be less than 4.*

* $I > i_{n+2}$ has $i_n \cdot I_n (= i_{n+1}^2) > i_n \cdot i_{n+2}$, or $i_n \cdot i_{n+2} + i_{n+1}^2 < 2_n i_{n+1}^2$. Again,

$$\frac{i_{n+1}-i_n}{I_{n+1}-i_{n+1}} = \frac{i_{n+1}-i_n}{I_n-i_{n+1}}\left[=\frac{i_n}{i_{n+1}}\right] \times \frac{I_n-i_{n+1}}{I_{n+1}-i_{n+1}}\left[=\frac{i_n+i_{n+1}}{i_n}\right]$$

$$= \frac{i_n+i_{n+1}}{i_{n+1}}, \; > 1\,.$$

Therefore, $(i_{n+1} - i_n) > (I_{n+1} - i_{n+1})$, $> (i_{n+2} - i_{n+1})$, since $I_{n+1} > i_{n+2}$, or $(i_n + i_{n+2}) < 2i_{n+1}$, and $i_{n+1}(i_n + i_{n+2}) < 2i_{n+1}^2$. Finally

$$(i_n + i_{n+1})\,(i_{n+1} + i_{n+2}) = (i_n \cdot i_{n+2} + i_{n+1}^2) + i_{n+1}\cdot(i_n + i_{n+2}) < 4\,i_{n+1}^2.$$

40 The particular case which is VIETA's, when $\vartheta = \pi/2$, or HL is the diameter of the circle, was given by DAVID GREGORY in his annotated account of his uncle's letters and manuscripts. See his *exercitatio geometrica de dimensione figurarum*, Edinburgh, 1684: 34–35.

41 See chapter 3.

42 *VCHQ*: prop. 15.

Clearly, GREGORY'S approach is powerful, and indeed he is able to derive several interesting corollaries. Thus[43] setting up the third comparison sequence

$$\begin{array}{cc} i_0 = & j_0 \\ i_1 = & j_1 \\ i_2 & j_2 \\ \vdots & \vdots \\ (I) & (J) \end{array}$$

(j_n) where $i_0 = j_0$, $i_1 = j_1$, $j_{i+2} - j_{i+1} = \frac{1}{4}(j_{i+1} - j_i)$, $i = 1, 2, \ldots$, he shows $i_1 - i_0$ $(= 4(j_2 - j_1)) < 4(i_2 - i_1)$, or $j_2 < i_2$. Similarly, $j_n < i_n$ $(n \geqq 2)$, so that

$$\begin{aligned} \lim_{n\to\infty} (j_{n+1} - j_1) &= \lim_{n\to\infty} \sum_{1\leqq\lambda\leqq n} (j_{\lambda+1} - j_\lambda) \\ &= (j_1 - j_0) \cdot \lim_{n\to\infty} \Big(\sum_{1\leqq\lambda\leqq n} (\tfrac{1}{4})^\lambda \Big) \\ &= \tfrac{1}{3} (j_1 - j_0). \end{aligned}$$

Or

$$\begin{aligned} I = \lim_{\mu\to\infty} (i_\mu) > \lim_{\mu\to\infty} (j_\mu) &= j_1 + \tfrac{1}{3} (j_1 - j_0) \\ &= i_1 + \tfrac{1}{3} (i_1 - i_0).^{44} \end{aligned}$$

A similar procedure[45] using two comparison sequences (j_n), $J_n)$ yields $I < i_0 + \frac{2}{3}(I_0 - i_0)$.[46] Thus, we define (j_n), (J_n) recursively such that $i_0 = j_0$, $I_0 = J_0$; and for all n

$$\begin{aligned} j_{n+1} &= (AM)(j_n, J_n) \left[= j_n + \tfrac{1}{2}(J_n - j_n) \right], \\ J_{n+1} &= (AM)(j_{n+1}, J_n) \left[= J_n - \tfrac{1}{4}(J_n - j_n) \right]. \end{aligned}$$

Then

$$\begin{aligned} j_1 &= (AM)(j_0, J_0) = (AM)(i_0, I_0) > (GM)(i_0, I_0) = i_1, \\ J_1 &= (AM)(j_1, J_0) > (AM)(i_1, I_0) > (HM)(i_1, I_0) = I_1; \end{aligned}$$

$$\begin{array}{cccc} i_0 & I_0 & j_0 & J_0 \\ i_1 & I_1 & j_1 & J_1 \\ \vdots & \vdots & \vdots & \vdots \\ & (I) & (J) & \end{array}$$

and in general, where $j_n > i_n$ and $J_n > I_n$, $j_{n+1} > i_{n+1}$ and $J_{n+1} > I_{n+1}$.* Further,

$$J_n - j_n = \tfrac{1}{4}(J_{n-1} - j_{n-1}) = (\tfrac{1}{4})^n (J_0 - j_0) = (\tfrac{1}{4})^n (I_0 - i_0),$$

and

$$\begin{aligned} j_n = j_{n-1} + \tfrac{1}{2}(J_{n-1} - j_{n-1}), &= j_0 + \tfrac{1}{2} \times (J_0 - j_0) \times \sum_{0\leqq\lambda\leqq n-1} (\tfrac{1}{4})^\lambda \\ &= i_0 + \tfrac{1}{2} \times (I_0 - i_0) \times \sum_{0\leqq\lambda\leqq n-1} (\tfrac{1}{4})^\lambda. \end{aligned}$$

* $j_{n+1} = (AM)(j_n, J_n) > (AM)(i_n, I_n) > (GM)(i_n, I_n) = i_{n+1}$,

and $J_{n+1} = (AM)(j_{n+1}, J_n) > (AM)(i_{n+1}, I_n) > (HM)(i_{n+1}, I_n) = I_{n+1}$.

[43] *VCHQ*: prop. 23.

[44] In its restriction to the circle-sector model it was given by HUYGENS in his *de circuli magnitudine inventa*, Leiden, 1654: prop. 5.

[45] *VCHQ*: prop. 21.

[46] First stated in the restriction to the circle-sector by WILLEBROD SNELL in *cyclometricus: de circuli dimensione secundum logistarum abacum*. Leiden 1621 (but not proved till HUYGENS gave several demonstrations in his *de circuli magnitudine inventa* (*op. cit.*, note [44]): especially prop. 5).

Finally,
$$J = \lim_{n\to\infty} j_n = i_0 + \tfrac{1}{2} \times (I_0 - i_0) \times \lim_{n\to\infty} \sum_{0\leq\lambda\leq n-1} (\tfrac{1}{4})^\lambda,$$
with
$$\lim_{n\to\infty}\Big(\sum_{0\leq\lambda\leq n-1} (\tfrac{1}{4})^\lambda\Big) = \tfrac{4}{3}$$
and $J > I$, since for all n $\begin{cases} j_n > i_n, \\ J_n > I_n. \end{cases}$

GREGORY'S analytical sequences, in fact, contain within their recursive definitions $\left(\text{in the parametrisation } i_n = 2^{n-1} \sin \frac{\vartheta}{2^{n-1}},\ I_n = 2^n \tan \frac{\vartheta}{2^n}\right)$ a sufficient basis on which to set up a general function theory of the circular functions; and $\left(\text{in the parametrisation } i_n = 2^{n-1} \sinh \frac{\vartheta}{2^{n-1}},\ I_n = 2^n \tanh \frac{\vartheta}{2^n}\right)$ also of the hyperbolic functions: the standard derivation technique would be by setting up suitable inequalities and comparison sequences. GREGORY had more than a glimmering of this richness and power, and tried to define by his sequences a problem which had taxed the ingenuity of mathematicians since Greek times—whether or not an analytical* quadrature of the circle is possible. After GREGORY ST. VINCENT'S gallant but feeble attempt[47], his is perhaps the first (and certainly in the 17th century the outstanding) attempt to prove that such analytical quadrature is impossible, as distinct from trying to isolate a particular rational number which shall be the ratio of circle circumference to diameter. GREGORY'S reasoning is most interesting and though inconsequential—it was justly if rather viciously attacked by HUYGENS[48]—cut away a lot of the deadwood of obsolete concepts which lay heavily but uselessly around the problem.

Interpreting his argument[49], let us consider the sequence (i_n), (I_n) whose common limit—when we take the model of the circle—is the quantity which we seek to derive analytically by some combination of members of (i_n), (I_n), say $n = 0, 1, 2, 3, \ldots, \lambda$ where λ is finite. GREGORY points out that if we can find an analytical function Φ such that $\Phi(i_n, I_n) = \Phi(i_{n+1}, I_{n+1})$, then $\Phi(i_0, I_0) = \lim_{n\to\infty} \Phi(i_n, I_n) = \Phi(I, I)$, and we could construct I analytically from i_0, I_0**: he therefore

* That is, in DESCARTES' sense of some combination of the four operations $\pm$, $\stackrel{\times}{:}$ together with root-extraction.

** He gives an example—in correction of a previous one whose inadequacy was pointed out by HUYGENS: Consider the sequences (a_n), (A_n), where
$$a_{n+1} = (HM)(a_n, A_n), \qquad A_{n+1} = (AM)(a_n, A_n):$$
$$\Phi(a_n, A_n) = a_n \cdot A_n \text{ has}$$
$$\begin{aligned}\Phi(a_{n+1}, A_{n+1}) &= a_{n+1} \cdot A_{n+1} = (AM)(a_n, A_n) \times (HM)(a_n, A_n)\\ &= [(GM)(a_n \cdot A_n)]^2 = a_n \cdot A_n \doteq \Phi(a_n, A_n).\end{aligned}$$
Therefore
$$\Phi(a_0, A_0) = \Phi(A, A), \quad \text{or} \quad A^2 = a_0 A_0, \quad \text{where } A = \lim_{n\to\infty} \begin{cases} a_n \\ A_n \end{cases}.$$

[47] In his *opus geometricum*, Antwerp, 1647—compare chapter 1.

[48] See E. J. DIJKSTERHUIS—who is perhaps overfair to HUYGENS in the squabble —*James Gregory and Christiaan Huygens*, ·≡· GREGORY *TV*: 478—486.

[49] *VCHQ*: prop. 11 and scholium. There is an interesting interpretation, which I do not wholly accept, in M. DEHN & E. HELLINGER: *On James Gregory's 'vera quadratura'* ≡ GREGORY *TV*: 468—478.

tries to argue that *no* such analytical function Φ can exist. We may assume that GREGORY tried many combinations of (AM), (GM) and (HM) to no effect before deciding that no such function Φ exists. In fact the functions Φ which satisfy must be transcendental since—in the particular case of the circle-area model—the sequence limit I (the general circle sector) can be shown to be non-expressible analytically (even more generally, algebraically) in terms of any set of members of the sequences (i_n) (I_n) *. Thus, perhaps the simplest function Φ which satisfies $\Phi(i_{n+1}, I_{n+1}) = \Phi(i_n, I_n)$ is $\Phi(i_n, I_n) \equiv I_n\left(\frac{i_n}{I_n - i_n}\right)^{\frac{1}{2}} \cos^{-1}\left(\frac{i_n}{I_n}\right)^{\frac{1}{2}}$, ** which is transcendental since $\cos^{-1} X$ is transcendental. (This function Φ gives an explicit value of the limit I of either sequence in terms of i_0, I_0—specifically $I = I_0\left(\frac{i_0}{I_0 - i_0}\right)^{\frac{1}{2}} \cos^{-1}\left(\frac{i_0}{I_0}\right)^{\frac{1}{2}}$.) *** The two geometrical models considered by GREGORY of $\left\{\begin{matrix}\text{ellipse}\\ \text{hyperbola}\end{matrix}\right.$ area arise by taking

$$\begin{cases} i_0 = \frac{1}{2}\sin 2\vartheta, & I_0 = \tan\vartheta \\ i_0 = \frac{1}{2}\sinh 2\vartheta, & I_0 = \tanh\vartheta \end{cases}$$

which induce the parametrisations

$$\begin{cases} i_k = 2^{k-1}\sin\dfrac{\vartheta}{2^{k-1}}, & I_k = 2^k \tan\dfrac{\vartheta}{2^k} \\ i_k = 2^{k-1}\sinh\dfrac{\vartheta}{2^{k-1}}, & I_k = 2^k \tanh\dfrac{\vartheta}{2^k} \end{cases}$$

which yields as the common limit of the sequences (i_n), (I_n)

$$I = \left\{\begin{matrix} \cos^{-1}\left(\dfrac{i_0}{I_0}\right)^{\frac{1}{2}} \\ \cosh^{-1}\left(\dfrac{i_0}{I_0}\right)^{\frac{1}{2}} \end{matrix}\right\} = \vartheta,$$

---

* The trigonometrical functions are transcendental.

** This follows by:

$$1.\quad I_{k+1}\left(\frac{i_{k+1}}{I_{k+1} - i_{k+1}}\right)^{\frac{1}{2}} = 2\, I_k \frac{i_{k+1}}{i_{k+1} + I_k} \times \left(\frac{i_{k+1} + I_k}{I_k - i_{k+1}}\right)^{\frac{1}{2}}, = 2\, I_k\left(\frac{i_k}{I_k - i_k}\right)^{\frac{1}{2}},$$

since

$$\frac{i_{k+1} + I_k}{I_k - i_{k+1}} = \frac{(i_{k+1} + I_k)^2}{I_k^2 - i_{k+1}^2} = \frac{(i_{k+1} + I_k)^2}{I_k(I_k - i_k)} = \left(\frac{i_{k+1} + I_k}{i_{k+1}}\right)^2 \times \frac{i_k}{I_k - i_k};$$

and

$$2.\quad \cos^{-1}\left(\frac{i_{k+1}}{I_{k+1}}\right)^{\frac{1}{2}} = \frac{1}{2}\cos^{-1}\left(2 \times \frac{i_{k+1}}{I_{k+1}} - 1\right) = \frac{1}{2}\cos^{-1}\left(\frac{i_{k+1} + I_k}{I_k} - 1\right)$$

$$= \frac{1}{2}\cos^{-1}\left(\frac{i_{k+1}}{I_k}\right) = \frac{1}{2}\cos^{-1}\left(\frac{i_k}{I_k}\right)^{\frac{1}{2}}.$$

$$*** \quad I_0\left(\frac{i_0}{I_0 - i_0}\right)^{\frac{1}{2}} \cos^{-1}\left(\frac{i_0}{I_0}\right)^{\frac{1}{2}} = \lim_{n\to\infty}\left[I_n\left(\frac{i_n}{I_n - i_n}\right)^{\frac{1}{2}} \cos^{-1}\left(\frac{i_n}{I_n}\right)^{\frac{1}{2}}\right]$$

$$= \lim_{n\to\infty}[I_n]\ (= I) \times \lim_{\lambda\to 1}\left[\left(\frac{\lambda^2}{1-\lambda^2}\right)^{\frac{1}{2}} \cos^{-1}\lambda\right]\ (= 1),$$

where $\lambda = (i_n/I_n)^{\frac{1}{2}}$.

since
$$I_0 \times \left(\frac{i_0}{I_0 - i_0}\right)^{\frac{1}{2}} = \left\{\begin{array}{c} \tan\vartheta\left(\dfrac{\sin\vartheta\cos\vartheta}{\tan\vartheta - \sin\vartheta\cos\vartheta}\right)^{\frac{1}{2}} \\ \tanh\vartheta\left(\dfrac{\sinh\vartheta\cosh\vartheta}{\tanh\vartheta - \sinh\vartheta\cosh\vartheta}\right)^{\frac{1}{2}} \end{array}\right\} = \left\{\begin{array}{c} 1 \\ i \end{array}\right\}. \star$$

It is interesting to see how GREGORY tries to prove the non-existence of (analytical) Φ by parametrising
$$\left\{\begin{array}{l} i_0 = a^2(a+b), \\ I_0 = b^2(a+b), \end{array}\right. \quad \text{which yields} \quad \left\{\begin{array}{l} i_1 = ab(a+b), \\ I_1 = 2ab^2, \end{array}\right.$$
and considering the functional equivalence which Φ has to satisfy, $\Phi\big(a^2(a+b), b^2(a+b)\big) \equiv \Phi\big(ab(a+b),\ 2\cdot ab^2\big)$. We realise that Φ cannot be a rational function, but it is difficult to see any further. GREGORY, however, tries to show that Φ cannot be a general analytical power-polynomial (where the coefficients may be general real), arguing on a basis of non-homogeneity—specifically, that the left side is a function of two binomials, while the right is a function of a binomial and a monomial. Such arguments, even when plausible—and of GREGORY's contemporaries HUYGENS at least would not allow even that—are difficult to check, while the property of homogeneity is not one which is, in general, unchanged by passage to the limit. We must therefore conclude that GREGORY, however verbally subtle, is not logically cogent.

These ideas of GREGORY's on sequence-convergence were not further developed in the period, and were not re-introduced into mathematical proof systematically till the rigorous reformulation of mathematics which began in the early 19th century. GREGORY himself, after his return to Scotland in 1669, forsook these methods for the more easily applicable ones afforded by the limit sum-sequence expansion.

The attitude typifies English mathematics from the early 1670's. The promising signs of birth of an analytical basis to function theory peter out, and the ease and rich suggestiveness of the new algorithmic methods flood everywhere. We now, however, pass on to an aspect of 17th century mathematics where, conversely, the very rigidity and power of its classically derived structure made the introduction of new concepts a slow and difficult process—geometry.

VI. The expanding concept of geometry

1. The synthetic approach

Elementary (EUCLIDEAN) geometry is, in a precise sense, more a psychological than a mathematical concept, appealing to some extent by its aesthetic purity but above all as an "obvious" abstraction from patterns apparent in sensed experience—an interpretation which agrees with its etymological derivation of "earth-measure". This abstraction has, at least from early Greek times, been increasingly elaborated and systematised till the present day, when we prefer, in exact treatment, to study the abstracted logical patterns in total disconnection from any consideration of the phenomena of physical reality, developing general sets of axioms which we hope, when operated on by appropriate deduction-rules, will consistently define an interesting geometry or topology. In the

⋆ The equivalence of the two parametrisations follows from
$$I_0\left(\frac{i_0}{I_0 - i_0}\right)^{\frac{1}{2}} \times \cos^{-1}\left(\frac{i_0}{I_0}\right)^{\frac{1}{2}} \equiv I_0\left(\frac{i_0}{I_0 - i_0}\right)^{\frac{1}{2}} \times \frac{1}{i}\cosh^{-1}\left(\frac{i_0}{I_0}\right)^{\frac{1}{2}}.$$

17th century the process was not very far advanced on its retreat from reality, and many particular geometrical concepts found difficult—especially that of continuity—were in fact justified by appeal to exactly these non-mathematical concepts of "smoothness", "unbrokenness" and the like which we prefer to reject as being insufficiently accurate for a mathematical treatment.

However, we can, I think, adopt a working definition which classifies as geometrical those 17th century studies which are more or less derivative from the classical (Greek) formalisation of EUCLIDEAN geometry and which have as typical undefined elements the ideas of "point", "line", "surface" and "volume". Further—following a traditional dichotomy—it will prove convenient to distinguish two important aspects as they crystallise out of a mass of inchoate material, partly original and partly an intellectual rediscovery of Greek geometry: the synthetic and the analytical. These aspects, however, in the ultimate must not be—and are not in my treatment—separated: each is a model of the other (and for any proof-structure in the one we can derive a corresponding proof-structure in the other) complementing it heuristically and as a matter of historical fact. In this chapter we discuss especially the synthetic aspect.

Perhaps most important in 17th century "pure" geometry are the freshly-studied projective concepts (developing for the most part out of 15th and 16th century perspective techniques in art), and the attempts made to prepare a sound theoretical basis for them by using aspects of classical Greek geometry, especially the APOLLONIAN derivation of the general conic as the cut of a plane with a (double-sheeted) cone, and the lemmas on cross-ratio developed by PAPPUS in Book 7 of his *Mathematical collection*[1]. Towards the middle of the century we find these systematised in the works of the Frenchmen G. DESARGUES[2], B. PASCAL[3] and, a little later, PH. DE LA HIRE[4]. but their inspiration, after a brief flowering, faded.[5] In contrast English geometry, isolated from the 16th century achievements in art (and in particular the theory of perspective), had little tra-

[1] See D. J. STRUIK's introduction to SIMON STEVIN's *De Deursichtige. The principal works of Simon Stevin II*, B. Amsterdam, 1958: especially 786ff.; J.L. COOLIDGE: *The mathematics of great amateurs*, Oxford 1949: especially chapters 3, 4, 5, and: *A history of geometrical methods*, Oxford, 1940: especially chapter 6: *Descriptive geometry*; and above all MICHEL CHASLES: *Aperçu historique ..., passim*. A recent review of relevant material is given by R. TATON: *La préhistoire de la géométrie moderne*, Révue d'Hist. des Sciences **2** (1949): 197—224.

[2] Compare R. TATON: *L'œuvre mathématique de Girard Desargues*, Paris 1951; and two interesting essays by WM. M. IVINS, Jr. in Scripta mathematica **9** (1943): 33—48; **13** (1947): 203—210, where he correlates DESARGUES' apparently esoteric terminology with technical terms used by 16th century Italian writers on perspective.

[3] Compare P. HUMBERT: *L'œuvre scientifique de Blaise Pascal*, Paris, 1947: especially 33ff.

[4] No adequate account is available of LA HIRE's work, but see R. TATON: *La première œuvre géométrique de Philippe de La Hire*, Révue d'Hist. des Sciences **6** (1953): 73—111.

[5] DESARGUES' treatises on theoretical geometry were largely ignored by his contemporaries in favour of his more practical works; PASCAL's projective treatment of conics were never published apart from the preliminary (privately circulated) handsheet of 1640, *Essay pour les coniques*, and are now otherwise completely lost except for a few notes taken by LEIBNIZ in the 1670's; while LA HIRE was admired more for his strictly APOLLONIAN study on conics (his *sectiones conicae*, Paris, 1685—*cf*. NEWTON: *Principia*, Book **1**: prop. 21, prob. 13) rather than for his little known work of 1673, the revolutionary *Nouvelle méthode en géométrie*.

ditional basis on which to develop projective concepts. Further, the standard English university course in mathematics of the mid 17^{th} century, in setting up EUCLID's *Elements* as a thought-structure to be viewed as an ideal of reasoned proof, tended rather to conceal the subtle mathematical concepts which lay embedded in it than to clarify them. BARROW—himself apparently self-taught—seems, in his public lectures at Cambridge and London from the 1650's, to have been the first university teacher in England[6] systematically to explore the riches of the Greek mathematical opus.[7] Significantly JAMES GREGORY, the greatest of the English geometers of the period apart from NEWTON and possibly WREN, received his main training under ANGELI during the four years of his stay in Italy, while NEWTON himself had BARROW for master. Other than by personal tuition there seemed little hope in England in the mid-century of gaining the adequate factual basis of knowledge which is necessary to complete comprehension and to further advance.

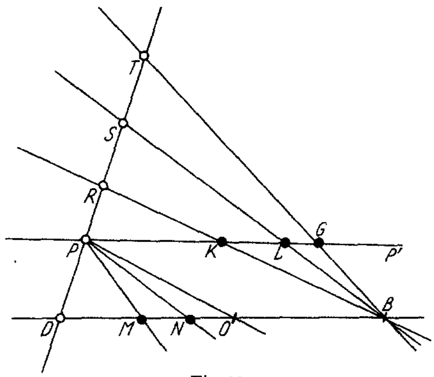


Fig. 22

Lack of a firmly-based tradition and standard text-book treatment implies almost inevitably an accompanying clumsiness in thought and expression—and so we find it, for example, in proof and application of the equivalent of the concept of cross-ratio invariance on a line-pencil. Thus BARROW in his *lectiones geometricae*[8] shows that where B is the centre of the pencil of lines BD, BR, BS, BT and PP', parallel to DB, is cut by BR, BS, BT in K, L, G, then $\frac{RD}{SD} = \frac{LG \times TD + KL \times RD}{KG \times TD}$ ★. The proof is immediate if we use the invariance of cross-ratio on a line-pencil:★★ for $B(TDSR) = B(G^{\infty}_{PP'}LK)$, which, expanded,

★ As we shall see later (chapter 10) BARROW requires this in the form

$$\frac{n}{DS} = \frac{m}{DR} + \frac{n-m}{DT},$$

where $LG:KG = m:n$.

★★ Not necessarily in exactly the modern, projectively suggestive form $(abcd) = (a'b'c'd')$ in which I give it, but also in any equivalent cross-product of line-segments in lines cut by the line-pencil—a form used by PAPPUS (in Greek times) and by BARROW's strict contemporary LA HIRE in an equally general way[9].

[6] FRANZ VAN SCHOOTEN had started such a systematic course at Leyden in the 1640's (of which CHRISTIAAN HUYGENS was the star pupil), and this could very well have inspired BARROW. HENRY BRIGGS in the early part of the century had attempted to inaugurate a stiff mathematical course at London and Cambridge, but the series quickly lapsed.

[7] These lectures were developed into his detailed if simplified and modernised texts of EUCLID (various editions of the *Elements* and "data" from 1655), but especially his *Archimedis opera, Apollonii Pergaei conicorum libri iiii, Theodosii sphaerica methodo nova illustrata et succincte demonstrata*, London 1675.

[8] *LG*: lectio 7: §§ 3—5 (§ 3 is the particular case where the point T is at infinity on the line DR).

[9] LA HIRE, of course, had published nothing in 1669, while the far-different proofs of cross-ratio invariance of PAPPUS' *Mathematical collection*, Book 7: props. 129, 136, 137, 140 and 142) suggest that BARROW was not familiar with PAPPUS' work (though COMMANDINUS had edited the full text in the later 16^{th} century, and his edition went through two printings).

gives $\frac{TS \times DR}{TR \times DS} = \frac{GL}{GK} = \frac{m}{n}$, so that $n\left(\frac{TD}{SD} - 1\right) = m\left(\frac{TD}{RD} - 1\right)$. BARROW, however, gives a long involved proof which reveals his lack of awareness of the significance of his result: so, taking PM, PN, PO parallel to BT, BS, BR respectively through P, the meet of PP' and DT, by similar triangles $DM \times TD = DN \times SD = DO \times RD = PD \times DB$: so that $DM \times TD = (DM + MN) \cdot SD$, or $DM \cdot (TD - SD) = MN \times SD$, and similarly $DM \cdot (TD - RD) = MO \times RD$, or $\frac{MN \times SD}{MO \times RD} = \frac{RD - SD}{TD - RD}$. Finally, $MN:MO = LG:KG$ has $LG \times SD \times TD + SD \times RD \times (KG - LG) = KG \times RD \times TD$, and the result follows. Clearly BARROW's result has as an immediate corollary the invariance of cross-ratio on the line-pencil, where the cross-ratio is defined as the cross-product of line-segments, but BARROW apparently failed to see it as more than a useful lemma invented to prove a tricky result, and certainly had no realisation that the theorem in fact defines an invariant of the point-correspondence cut on two arbitrary lines by his line-pencil.[10]

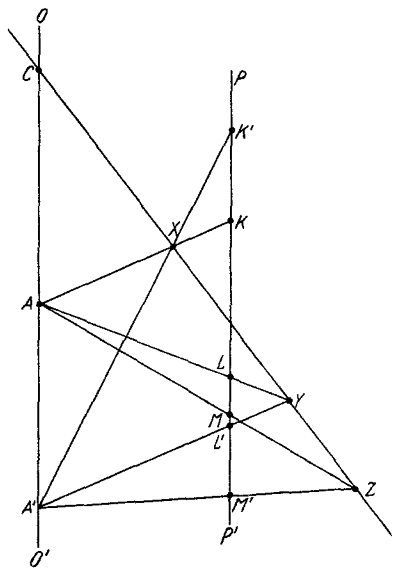


Fig. 23

A similar failure to abstract out any general concept of cross-ratio invariance may be found in WALLIS' *Angular sections*, where WALLIS gives his solution to a problem submitted to him in 1674 by GEORGE FAIRFAX[11]: where A is any point on the line OO' and X, Y, Z are three colline points, show that $KL:LM$ is constant, where K, L, M are cut out on an arbitrary line PP' by AX, AY, AZ. Again there is an immediate proof by cross-ratio by considering a second position A' of A and showing that $KL:LM = K'L':L'M'$, where K', L', M' are defined correspondingly*, and this is indeed WALLIS' approach. His proof, however, even more than BARROW's above, is a long, cumbersome essay on a grand scale in similar triangles and proportionality, and any general view is lost in a haze of particularities.

The general ideas which are lacking in BARROW and WALLIS had already been introduced in Greek mathematics—an aspect of the Greek achievement which has received too little credit. Much of this Greek work on general point and line correspondences is now—as it was in the 17th century—seemingly irretrievably lost, but its outline is clear whatever its particular historical forms may

* We have the perspectivities $A\,(\infty_{pp'}\, KLM) = (CXYZ) = A'(\infty_{pp'}\, K'L'M')$ where C is the meet of AA', XYZ.

[10] BARROW, in fact, wants the theorem only to give him relations between the subtangents BR, BS, BT of curves BR', BS', BT' tangent at B to each respectively. See *LG*: lectio 9: §§ 10, 12, 14; 73–74. It is significant that in 9: § 10, where he rejects the easy cross-ratio proof, BARROW's cumbrous alternative is invalid (see J.M. CHILD: *Geometrical lectures of Isaac Barrow*, Chicago, 1916: 107, note).

[11] Published with his *Algebra*, London, 1685: ch. 8 ≡ *opera mathematica* 2 (1693): 592–593.

have been. In common with other aspects of Greek geometry no adequate notation had formally been set up to deal with the concept of correspondence, but the general idea of a cross-product is already old with PAPPUS and the concept of pole-polar with regard to the general conic is fully developed in APOLLONIUS' *Conics*[12]. Further, as the lemmas in PAPPUS' *Mathematical collection* allow us to restore them, EUCLID's books of porisms[13] and several of the minor works of APOLLONIUS[14] but above all APOLLONIUS' *Conics* show that by the second century A.D. there had been obtained equivalents of the constancy of cross-ratio of the pencil formed by four fix-points on a conic and any fifth variable point on the conic as pencil-centre—specifically, the "locus ad tres et quatuor lineas"—, and of "DESARGUES'" theorem of the involution cut on a line by the four sides of a quadrilateral and the family of circumscribing conics[15], of DESARGUES' theorem that two triangles with corresponding vertices on copoint lines have the meets of corresponding sides colline[16], and of PASCAL's theorem on the colline meet of opposite sides of a hexagon inscribed in a conic in the degenerate case of a line-pair.[17] It was, however, the problem of the 3/4 line locus* which attracted most attention among 17th century geometers—probably in the first instance because it had gained the reputation of being supremely difficult and because in solving it one might gain insight into the methods of solution of the ancients[18] rather than through any consciousness of its fundamental importance.[19] DESCARTES, in a development confused by many modern historians, had reduced

* The point-set such that the product of its angled distances from two given lines has a constant ratio to the product of its angled distances from two further given lines (which may coincide).

[12] Especially Book **3**: props. 30—34.

[13] An admirable restoration is that of M. CHASLES: *Les trois livres de porismes d'Euclide* ..., Paris, 1860; and compare J. J. MILNE: *An elementary treatise on cross-ratio geometry* ..., Cambridge, 1911: especially appendix **1**: 114—129: *Pappus' account of the porisms of Euclid* ...; and CHASLES' *Aperçu historique* ..., Paris, 1889: 274—284: Note 3: *Sur les porismes d'Euclide.*

[14] Such as his (lost) works *On cutting off a space*, *On determinate section*, but especially the (extant) *On cutting off a ratio* (edited by HALLEY from an Arab manuscript, as *de sectione rationis*, Oxford, 1706).

[15] This is developed in APOLLONIUS: *Conics*: Book **3**: props. 16—23, and is a slight modification only of the constant cross-ratio property by suitably defining involution.

[16] DESARGUES gave this form of the theorem in A. BOSSE's *Pratique de la perspective*, Paris, 1648: 304 ff., but the PAPPUS form is stated in "porism" form (and not quite fully) but with an extension not given by DESARGUES. (See PAPPUS: *La collection mathématique* (ed. P. VER EECKE), Paris-Bruges, 1933: Book 7, introduction ·≡· **2**: 488.) The extended theorem survives in a badly mangled text, and its meaning was restored in modern times only by R. SIMSON—see *Pappi Alexandrini propositiones duae generales* ... *PT* **32** (1723): 330—340.

[17] PAPPUS: Book 7: props. 138, 139.

[18] An important reason for 17th century mathematicians who—not wholly wrongly—were convinced that the ancient Greeks had "analytical" methods of solution, not transmitted to modern times, which they had used to derive many of the results given in the often artificial and obscure forms of the extant texts.

[19] APOLLONIUS, in the preamble to his *Conics*, had introduced it as a problem whose general solution had baffled EUCLID, remarking intriguingly that its solution was a corollary to theorems given in his Book *3*. It is significant that NEWTON's solution depends on exactly those propositions of Book *3* which contain, implicitly, the definition of a conic as the point-set meet of corresponding rays of equi-cross line-pencils.

its solution with respect to an oblique coordinate system to a second-degree polynomial point-set in two variables (the corresponding coordinate lengths) and so showed the locus a conic; while PASCAL in his lost *Traité des sections coniques* claimed a synthetic solution (and, indeed, it is easy to reduce the locus-property to the condition of colline meets of opposite sides of a hexagon—PASCAL'S "hexagramma mysticum" condition, which shows the six vertices of the hexagon to be on a conic[20]), but the first extant synthetic solution is that given by NEWTON.[21]

Briefly, NEWTON, taking APOLLONIUS **3**: props. 16—23* as his starting-point, derives the easy generalization which is equivalent to DESARGUES' conic-involution theorem[22]: where $ABCD$ is a quadrilateral inscribed in a conic, and PQ, PR, PS, PT—the angled distances of P from AD, BC, AB, CD respectively—are drawn from any point P on the conic under given angles PQA, PRC, PSB, PTD, the cross-product $\frac{PQ\times PR}{PS\times PT}$ is constant. The 3/4 line locus is the easy converse of this. It is important, however, to notice that the condition $\frac{PQ\times PR}{PS\times PT}=\lambda$, constant, is strictly equivalent to the condition that the point set of P be defined by the constancy of the cross-ratio $P(ACDB)$**; and that therefore any treatment which introduces the one introduces the other in equivalent form. In fact, NEWTON uses his theorem to derive a whole sequence of propositions defining several types of point-correpondences, and we may fairly say that

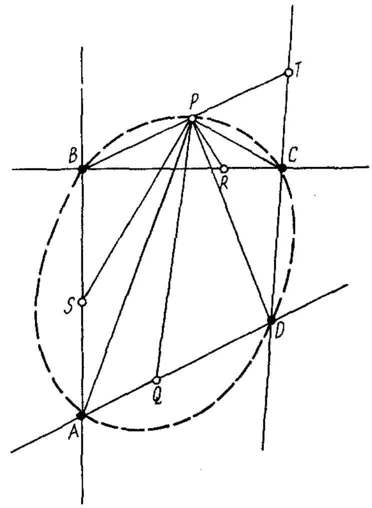


Fig. 24

* These propositions relate to rectangle-segments in a conic, and yield immediately DESARGUES' involution-theorem for a trapezium inscribed in a conic.

** $\Delta APD = q\times PQ\times AD = \frac{1}{2}\times PA\times PD\times \sin APD$, or $PQ = q'\times PA\times PD\times \sin \widehat{APD}\left(\text{where } q'=\frac{q}{2\cdot AD} \text{ is some constant}\right)$. Similarly

$$PR = r'\times PB\times PC\times \sin \widehat{BPC},$$
$$PS = s'\times PB\times PA\times \sin \widehat{APB},$$
$$PT = t'\times PC\times PD\times \sin \widehat{CPD},$$

or $\frac{PQ\times PR}{PS\times PT}=\lambda$ (by the locus condition) $=\mu\times\frac{\sin\widehat{APD}\times\sin\widehat{BPC}}{\sin\widehat{APB}\times\sin\widehat{CPD}}=\mu\times P(ACDB)$, $\mu=\frac{q'\times r'}{s'\times t'}$.

[20] A restoration on these lines of PASCAL'S solution (using the help of the LEIBNIZ notes on his *Conics*) is in an (unpublished) paper of mine, *Pascal's hexagramma mysticum*. For DESCARTES' solution see the next chapter.

[21] In the manuscript *de compositione locorum solidorum* (to be dated in the early 1670's ·≡· *CUL Add.* 3963: various drafts in 126R—149R, later published—not quite so fully—in *PM*: Book *1*: Section 5: lemmas 17—19.

[22] *Add.* 3963: 127R: cons. 2 = *PM* **1**: lemma 17.

NEWTON develops that sequence on a basis which involves the projective definition of a conic as the cut of equi-cross line-pencils (in equivalent form, at least).

In amplification of this point let us consider his manuscript prop. 3[23] which, slightly reformulated, proves: given fix-points B, C and fix-lines PR, PT, the point-set of all points D such that $PR:PT$ is constant, where BD, CD meet PT, PR respectively in T, R, is a conic; and conversely*. In proof, NEWTON takes $DHIG$ parallel to PT, DE parallel to PR with CP meeting DE in F. Then, $PQ:DE\ (=IQ)=PB:HB=PT:DH$, and $PR:DF=RC:DC=IG\ (=PS):DG$, so that $\frac{PQ\times PR}{PS\times PT}=\frac{DE\times DF}{DG\times DH}$, constant for D on the conic through the fix-points A, C, P; B; or, since PQ, PS are given in magnitude, $PR:PT$ is constant (which shows the converse—the theorem itself is immediate by reversing the argument).

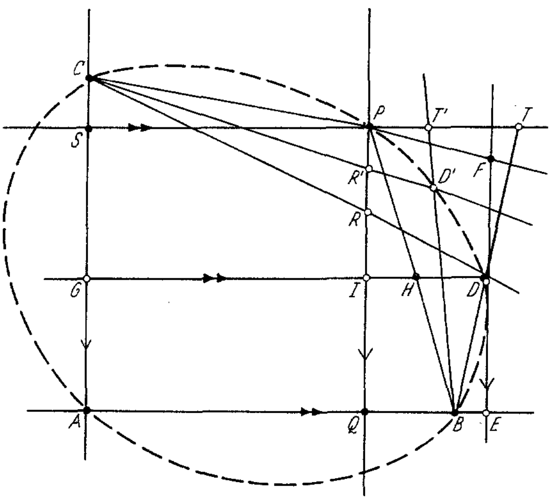


Fig. 25

This is a powerful porism in the EUCLIDEAN manner, but its significance tends to be hidden in a classically geometrical clothing. (The argument may, however, be neatly reduced to a form which reveals the implicit use of the cross-ratio invariancy property more clearly following each step of NEWTON's argument exactly.**) Indeed, he derives his "organic" construction of a conic almost in corollary[24]—specifically, if the given angles DBM, DCM rotate round fix-points B, C such that the meets of BM, CM are colline, then the point-set of all D is a conic. We have merely to take PR, PT through a fix-point P (defined by the organic construction from a corresponding fix-point N on the given generator-line NM) such that $\widehat{BPT}=\widehat{BNM}$, $\widehat{CPR}=\widehat{CNM}$: then the triangles NBM, PBT; NCM, PCR are similar, so that $PT:MN=PB:NB$, $PR:MN=PC:NC$, or $\frac{PT}{PR}=\frac{PB\times NC}{PC\times NB}$, constant—which shows that D, the meet of BT, CR is on a conic through B, C, P (and the meet A of the parallels through B, C to PT, PR respectively).

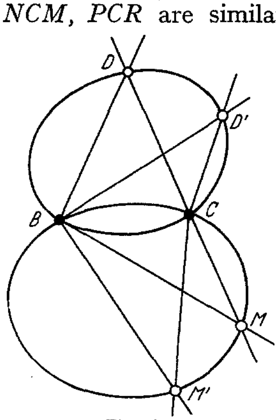


Fig. 26

* Clearly there is a unique point A on the conic which corresponds to R, T both at infinity—specifically, A is the meet of the parallels through B, C to PT, PR.

** Considering a similarly defined point D', $PT:PT'=PR:PR'$, and we show B, C to lie on a conic through P, A, D, D': for it is immediate that $(P\infty TT')=(P\infty RR')$ with $B(P\infty TT')=B(PADD')$ and $C(P\infty RR')=C(PADD')$, or $B(PADD')=C(PADD')$.

[23] *Add.* 3963: 128R $=$ *PM* 1: lemma 20.

[24] *Add.* 3963: Prop. 7: 130R–V $\equiv$ *PM* 1: lemma 21.

Conceptually, however, NEWTON'S attempt to show the converse is more revealing of the inadequate grasp even NEWTON had of the homographic definition of a conic which is implicit in the porism. Though in the manuscript version[25] NEWTON hints at the necessary and sufficient condition which would validate his argument, in the published *PM* version NEWTON is misled in showing the converse, by his not implausible conclusion that *only* colline points M, N will generate a point-conic through B, C, —in fact, any conic through B, C is transformed into a second conic through B, C under the organic construction.*[,26]

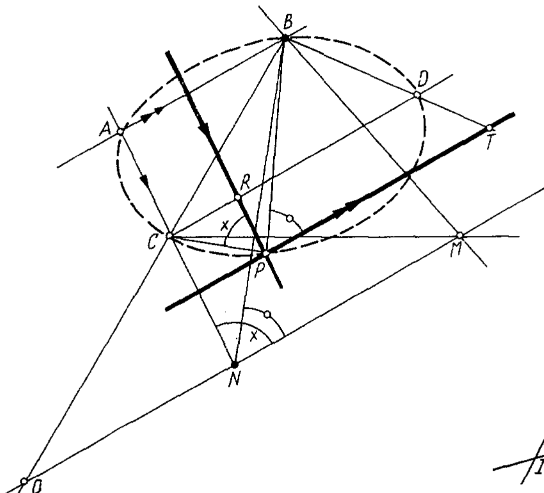


Fig. 27

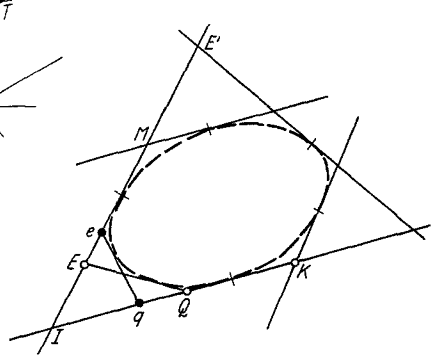


Fig. 28

Elsewhere in *PM*[27] NEWTON treats of a dual line-porism: if two fix-lines ME, KQ are given and fix-points M, K on them, and a correspondence between the points E, Q of the two lines is set up by the condition that $ME \times KQ$ is constant, then the line-set of the EQ envelopes a (line-) conic tangent to ME, KQ. (His proof is closely Apollonian in form, but then APOLLONIUS had, in his *Conics*, developed the basis for a general treatment of line-porisms at greater length than the corresponding one for point-porisms.) Together, as they are given in *PM*, these porisms are tied strictly to the easily provable corollaries which give constructions for conics through given points and tangent to given lines in various arrangements, and we could easily have the impression that they were thought up *ad hoc* during the period 1684—86 (when most of *PM* was written) expressly

* If point M defines a corresponding point D on the conic through B, C (where the angles DCM, DBM are constant), the line-pencils (CM) (BM) are transformed into the respectively equi-cross pencils (CD), (BD); together with, since the point-set of D is a conic through B, C, the condition that the pencils (CD), (BD) are equi-cross—which shows the pencils (BM), (CM) are equicross, or the point-set of M is a (usually) non-degenerate conic through B, C.

[25] *Add.* 3963: 130R. The condition there given which suitably restricts the converse is that some point O of the locus ONM be colline but not coincident with B, C—which implies that the locus ONM reduces to the line-pair $AB \times ONM$.

[26] The point was first made by J.L. COOLIDGE—see *A history of the conic sections and quadric surfaces*, Oxford, 1945: 46.

[27] *PM*: 1: lemma 25, which generalises APOLLONIUS *3*: prop. 42, his own lemma 24.

to prove such constructions. This is far from true[28] and a clearer view is obtainable from numerous manuscript drafts on geometry in the Portsmouth Collection[29]. In particular, the heading under which the propositions printed in *PM* were originally collected, *de compositione locorum solidorum*, indicates the deliberate intention to write a systematic treatise (never completed) on the Greek theories of point- and line-porisms. Striking confirmation is to be found in the manuscripts which he wrote at the end of his life (from about 1705) when interest in theories of correspondence and especially the Greek porism theory of point-correspondences was renewed.[30] These show that few exact thoughts crystallized out of a mass of fluid ideas which surged through his mind, but they yet remain tremendously suggestive for future developments.

It is clear that NEWTON was attempting a clarification and systemisation of basic concepts in geometry, particularly those of the point-set (locus, "locus punctorum") and line-set (envelope, "locus linearum") and the relationship between points and lines which correspond ("fratres sunt") or are "twin" ("quantitates gemellae").[31] In particular he elaborates the basic concept of porism (point-set) at some length[32]: "The curves ("lineae") on whose meets are the required points were called by the ancients the loci of these points, and they found other loci of the same kind by dropping one defining condition of the

---

[28] Examination of the handwriting style suggests that the tract cited above *de compositions locorum solidorum* was written in the early 1670's.

[29] Compare *Add.* 3963: Sections 1—5, 10, 12—14, but especially sheets 127—133, 135, 137, 141—144, 145—146; *Add.* 4004: 128—159, 183—185. A large part of these manuscripts are drafts, to be dated about 1705, of an intended treatise on geometry, of which perhaps the fullest draft (of Book 1 only) is *Add.* 4004: 128—159, to be collated with *Add.* 3963: 127—133. Many of the subsidiary tracts are specifically labelled "porismata".

[30] A manuscript quoted by S.P. RIGAUD in his *Historical essay on ... Sir Isaac Newton's 'principia'*, Oxford 1838: no. *23*: 79 shows that DAVID GREGORY in May 1701 had the intention of visiting NEWTON to talk among other things "about Euclid, especially the data; and if I should write a Preface, and what instances put in it" (his edition of the *data* came out in 1704 in his *Euclid*). Further, HALLEY, in his edition of APOLLONIUS' *de sectione rationis*, Oxford, 1706, gave a Latin translation of PAPPUS' description of lost Greek work on porisms in which he wrote of EUCLID's main porism—restored by SIMSON a few years later (see note[16] above—: "porismatum descriptio nec mihi intellecta nee lectori profutura; quid sibi vult Pappus haud mihi datum est conjicere". To NEWTON that could only have been a challenge to prove HALLEY wrong). NEWTON's work on porisms was based on a wide reading—*cf. Add.* 3963: 157L—of all available commentaries and attempts at restoration, but especially those of SNELL, VIETA, GHETALDI, ANDERSON and VAN SCHOOTEN, and of the rich if mangled text of PAPPUS' Book 7 itself (which is still our only source for information on the Greek theories). NEWTON's porism restorations anticipate to a surprising degree the later (and completely independent) work of MICHEL CHASLES published in his *Les trois livres de porismes d'Euclide*, Paris 1860, and that agreement in restoration must clearly give added weight to their plausibility.

[31] This concept is discussed in *Add.* 3963: Section 5: *regula fratrum* (rule of mates)—*cf.* 40R: "fratres voco puncta vel lineas quae eodem modo se habent ad conditiones problematis", and again "... quantitates gemellae, id est, quae eadem modo se habeant ad conditiones problematis, quaeque cognitam aliquam habeant relationem ad invicem: his non impono nomina, sed earum loco usurpo quantitates quae eodem modo se habeant ad utramque".

[32] *Add.* 3963: 17R.

problem and seeking the curve each one of whose points shall satisfy the remaining conditions. Then if each point of one curve satisfy all conditions but one, and each point of a second curve satisfy all conditions but a second one, their meets determine those points which satisfy (the union of) all the conditions." The natural way to develop this viewpoint is by introducing an analytical free variable to represent the set which satisfies all conditions but one—clearly, we have only to introduce some reference system, as Cartesian coordinates*—but using the pure geometrical model of the straight line we easily define correspondence conditions by restricting the line to joining corresponding points on given curves and then the whole field of elementary projective geometry lies open to investigation.

Of this, of course, the most important individual result will be the constancy of cross-ratio on a line-pencil, and we find that NEWTON gives more or less general (if differing) proofs in the manuscripts, showing for example[33] that, where any line through fix-point A meets the copoint lines Ef, Eg, Eh in B, C, D, then $(AB \times CD)$: $(AC \times BD)$:$(AD \times BC)$ are constant ratios—a theorem which corresponds exactly to our more sophisticated definition of cross-ratio, since the cross-products

$$\frac{AC \times BD}{AD \times BC} \ (= (ABCD))$$

and

$$\frac{AB \times CD}{AC \times BD} \ (= (ADBC))$$

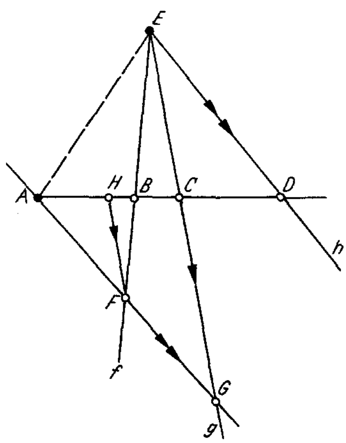


Fig. 29

are constant on the line-pencil. NEWTON gives, interestingly, a form of the PAPPUS proof which virtually projects D into infinity by taking AFG parallel to Eh, and again FH parallel to EG: then $AB:BD = AF:ED = AH:CD$, $AH:AC = AF:AG$, or $\frac{AB \times CD}{BD \times AC} = \frac{BD \times AH}{BD \times AC} = \frac{AF}{AG}$, constant—an argument exactly analogous to $(ADBC) = (A \infty_{Eh} FG)$, constant.[34]

The immediate application is to consider the PAPPUS lemma which is equivalent to DESARGUES' theorem on perspective triangles, and which NEWTON formulates[35]: where the fix-points A, B, C are colline and the point-sets F, D are fix-lines such that FD is through A, then the point-set E defined as the meets of BF, CD is a fix-line also (and passes through G, the meet of the point-

* This was, in fact, NEWTON'S basis for introducing his "independent" CARTESIAN coordinate system in treating the concept—see the next chapter.

[33] *Add.* 3963: 30R.

[34] A complementary analytical sketch, depending on a subtle analysis of conditions for 1, 1 correspondence between points on two lines, exists at 159Rff.; *de inventione porismatum*—see next chapter.

[35] *Add.* 3963: 29R: porism 12. As NEWTON shows by his figure of an alternative draft, he is aware that the point-sets E, F, D are copoint at G—a criticism which has been raised against the PAPPUS original, since the point is not made explicitly in the text or in figure.

sets (F), (D).[36] No proof is given, but the form in which the porism is given allows us plausibly to reconstruct it in equivalent form: specifically the line-pencils $C(E)=C(D)=A(D)=A(F)=B(F)=B(E)$, so that the point-set E is (part of) a conic through B, C; and this we easily show to be the line-pair $EE'\times BC$, where E', colline with B, C, is a point of (E).

But, more generally, NEWTON considers the correspondences set up by the meet of a line with higher curves[37] (perhaps on the model of APOLLONIUS: *On tangencies*, restored in printed form by VIETA and FERMAT, and in manuscript by TORRICELLI[38]). This leads easily to a general treatment of centres of similitude

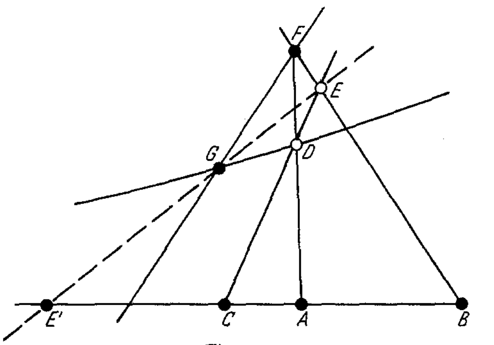


Fig. 30

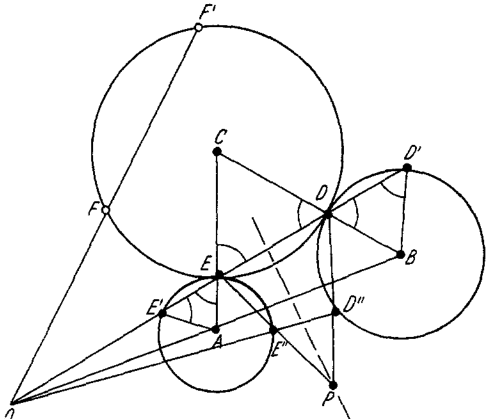


Fig. 31

with respect to pairs of circles. Thus, with regard to the circles (A), (B), consider the (external) centre of similitude O which is defined on AB by taking OA; OB in the ratio of the respective circle radii. Then for E on the circle (A) and D on OE such that $OE\times OD=OA\times OB$, we easily show D to be on the circle (B) and further that a unique circle (C) can be drawn touching the circles at D, E.* Again, given a point F on the circle (C), a second point F' on it (colline with O, F) is defined by $OF\times OF'$ $(=OE\times OD)=OA\times OB$, and from this NEWTON easily derives solutions of APOLLONIUS' problem to find the circle tangent to three given circles, any of which may degenerate.[39] But perhaps more important for NEWTON is that the "puncta gemella" D, E of the circles (A), (B) define, with

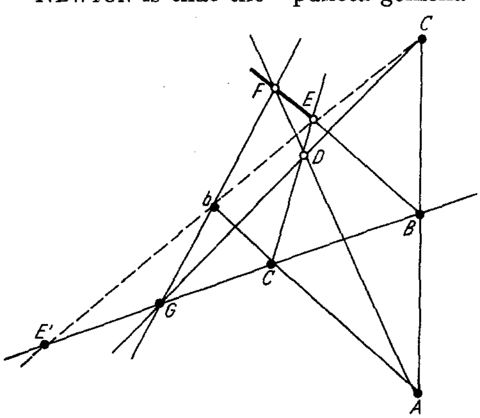


Fig. 32

* Since $\widehat{CED}=\widehat{OEA}=\widehat{D'DB}=\widehat{CDE}$, or $CE=CD$, where AE, BD meet in E.

36 Indeed, NEWTON adds the generalisation that the result holds for A, B, C, given in general position in the plane, provided that B, C and G are colline (in which case the point-sets (D), (E), (F) will not be copoint). [The proof follows immediately from PAPPUS' theorem on the hexagon $FBACDG$ inscribed in the line-pair FAD, BCG.] See *Add.* 3963: 29R.

37 *Add.* 3963: 40R–41V.

38 See E. TORRICELLI: *opera*, **1**. 1: 239–292: *de tactionibus*.

39 Considered in APOLLONIUS' (lost) treatise *de tactionibus*.

respect to given O, A, B, a "relatio" (our modern inversion correspondence) under which important circle properties remain invariant*; and, in this general viewpoint, the common tangent-circle is but one (simple) example of an element which remains invariant under the correspondence.

Conversely, the point-set of the circle—and more widely of the general conic—may be used to define correspondences in a given line, and NEWTON develops this aspect at some length.[40] Thus, for example, given a circle through fix-points A, B and the line $\alpha\beta$ (fixed likewise in position), NEWTON considers the point-sets (x), (y) which are cut in $\alpha\beta$ by the lines AZ, BZ drawn through

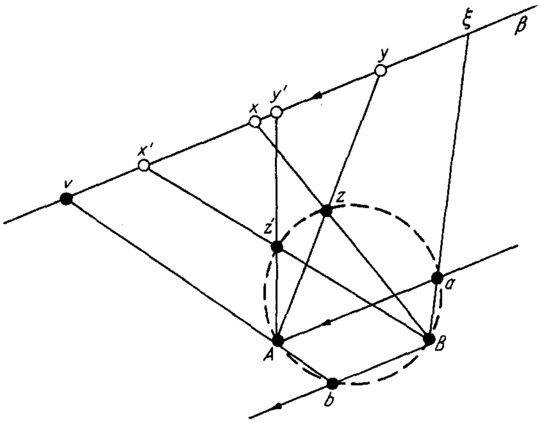


Fig. 33

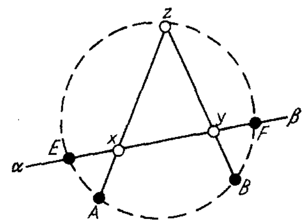


Fig. 34

an arbitrary point on the circle (and the points $y=\xi$, $x=\nu$ correspond respectively to the particular cases where x, y are at infinity), and states that the product $\xi x \times \nu y$ is constant. His justification depends on setting up a (detached) coordinate system in $\alpha\beta$ (where the coordinate line-lengths are defined by $\xi x = x$, $\nu y = y$), but we note that implicit is the definition of the circle as the meet of equicross (and indeed congruent) line-pencils**—a property stated explicitly in a second porism which follows immediately on: given a circle through fix-points A, B and the fix-line $\alpha\beta$ which cuts it in the fix-points E, F, if the lines AZ, BZ through any arbitrary point Z on the circle cut out the respective point-sets (x), (y) on $\alpha\beta$, then $Ex \times Fy : Ey \times Fx : EF \times xy$ are in given ratio.***

* So, if OE is tangent at E to (A), OD is tangent to (B) at D; and, again, corresponding circle chords EE'', DD'' have their meets P colline (on the radical axis (P), which is itself an invariant of the correspondence).

** Specifically, we take a second point z' on the circle to define a second pair of points x', y' in $\alpha\beta$; then $A(abzz') = (\infty_{\alpha\beta}\, yy') = \nu y' : \nu y = \xi x : \xi x' = (\infty_{\alpha\beta}\, xx') = B(ab\, zz')$, which shows that A, B are on a conic through a, b, z, z'.

*** In fact, the cross-product $\frac{Ex \times Fy}{Ey \times Fx} = (EFxy) = (EyFx) = \frac{EF \times xy}{EX \times Fy}$, and NEWTON's porism states equivalently, where the point-set (Z) is a circle, the constructed point-sets (x), (y) are such that $(EFxy) = (EyFx)$ is constant—or, alternatively, that for any Z on the circle $(EFxy) = Z(EFAB)$ is constant. A porism of a similar kind for the parabola had already been found by FERMAT no later than the middle 1650's, and first appeared in print in WALLIS' *commercium epistolicum* in 1658 (in Letter 47 ≡ FERMAT-KENELM DIGBY, 19 June 1658: 188). Both the circle and parabola forms appear (as porisms 3,2 respectively) in his posthumously printed tract on porisms. (See FERMAT's *varia opera*, 1679 ≡ *OE* **1** (1891): 76ff.)

[40] *Add.* 3963: 165 Rff.

None of this work of NEWTON'S on the concept of plane correspondences was published in his time—or, indeed, ever—and had no influence on his contemporaries. With NEWTON'S death the topic faded temporarily into oblivion.

The analogous concept, however, of 3-space correspondences, widely studied since Greek times, attracted wider attention—and especially that part which dealt in a general way with the continuous mapping of one surface into another. In particular, an offshoot of the growing science of cartography was the problem posed by the map-projection: how best shall we map the earth's surface (abstracted into the form of a sphere-surface) onto a plane? Clearly, a continuous mapping onto an infinite plane is possible where only one point on the sphere is not mapped onto a finite point in the plane (but no continuous mapping can map every point onto a finite point). A further important need in the (descriptive) map is that "shape" be preserved, that the mapping be conformal. Combining both advantages PTOLEMY[41] set up a perspective mapping of the sphere onto the equatorial plane from the south pole as perspective pole (known as "stereographic" projection of the sphere after D'AIGUILLON elaborated its theory under that name), and proved its conformality. With the pressing 16th century demand for a convenient navigating map, several projections were introduced but especially that of GERARD MERCATOR[42] (the "MERCATOR" projection) which, while non-perspective, was continuous, conformal and—most interestingly—direction-preserving, projecting meridians, parallels and loxodromes on the sphere into straight lines. In MERCATOR'S time the practical construction of the mapping $\left(\text{which involves an equivalent of } \int_0^{\vartheta} \sec\vartheta \cdot d\vartheta\right)$ was carried out by approximation, though the underlying theory was worked out only by JAMES GREGORY in 1668 $\left(\text{who in } EG \text{ gives the equivalent of } \int_0^{\vartheta} \sec\vartheta \cdot d\vartheta = -\log(\sec\vartheta - \tan\vartheta)\right)$, with later simplification of GREGORY'S complexities by BARROW and WALLIS.[43] HALLEY at the end of the century gave a discussion which neatly tied up the stereographic projection with the MERCATOR scheme,[44] showing that the stereographical projection of the loxodrome (the curve on the sphere which cuts all parallels at the same angle) must be the conformal curve which meets a family of concentric

---

[41] *Cf.* his *Geography*: 1, ch. 24 (*cf.* P. SCHNÄBEL: *Text und Karten des Ptolemäus*, Leipzig, 1939) though the theory is developed in his *planispherium* (Venice, 1558; Leipzig, 1907). Compare J.O. THOMSON: *History of ancient geography*, Cambridge, 1948, and D. J. STRUIK, *Outline of a history of differential geometry 1*, Isis 19 (1933): 92—120, especially 94ff. [Full bibliography in HOFMANN *1*: 188, col. 1.]

[42] See H. VON AVERDUNK & J. MÜLLER-REINHARD: *Gerhard Mercator und die Geographen unter seinen Nachkommen*, Gotha, 1914: 128ff.

[43] Detailed references are given in F. CAJORI: *On an integration ante-dating the integral calculus*, Bibliotheca mathematica$_3$ 14 (1913—1914): 312—318.

[44] In *PT 19* (1695): No. 215: *An easy demonstration of the analogy of the logarithmick secants to the meridian line* ... In outline the technique used by HALLEY was known in the 1670's—compare COLLIN'S letter to OLDENBURG (? 1670) (·≡· RIGAUD (*C*): 1, 142—147, especially 144) which apparently reports the manuscript on the "rhumb spiral" (now in the Royal Society Library) enclosed in GREGORY'S letter to him of 20 April 1670 (·≡· GREGORY *TV*: 93—96, especially 94). GREGORY'S solution, while not so precise as HALLEY'S is based likewise on the stereographic projection of the loxodrome ("rhumb-line") into the logarithmic spiral.

circles at that angle—that is, a logarithmic spiral. Proof of the conformality of stereographic projection is fundamental to the approach, and HALLEY substitutes a neat demonstration for PTOLEMY's complexities. Consider, then, the vertical section BPE of a sphere of centre C and south pole E: to show conformality it is sufficient to prove that the angle $\widehat{DPA}$ made by the tangent PD to the vertical PA at the sphere-point P projects into an equal $\widehat{dpa}$ in the equatorial plane CFd (where pd is the tangent to the projected curve at point p corresponding to P). Taking DA, da normal to the vertical plane PO perpendicular to BE with AK parallel to PO (meeting EP in K), we easily show $\widehat{AKP} = \widehat{OPE} = \widehat{APK}$, or $AK = AP$, so that $\widehat{dpa}$ ($= DKA$ since DKA, dpa are parallel planes) $= \widehat{DPA}$.*

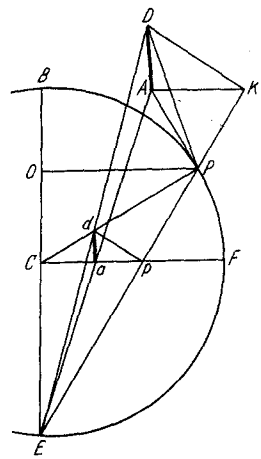


Fig. 35

But, of course, the most fully worked out case of a 3-space point-correspondence was the classical Apollonian construction of conics as the meet of a plane with a double-sheeted cone; or, restating it (but not wholly unclassically), the perspective correspondence which transforms any point-conic into any other (degenerate or otherwise), and conversely. In the Greek treatment, however, when the basic "symptoms" of the conics as plane curves had been derived, the point-correspondence on the cone was discarded, and the whole mass of Greek conic theory —and its systematised development in the 17th century[45]—had been elaborated as a plane curve theory, rather cumbersome and turgid in many ways, defined by "symptoms" with respect to a chord and a conjugate diameter. The especial difficulty of the purely plane approach was that definitions of many important elements, especially the focus and its polar (the directrix), had to be introduced in an entirely unobvious way as point-sets restricted by a condition involving unwieldy ratios of line-segments. (In comparison, DANDELIN's 19th century definitions of the foci as the contact-points of the plane which cuts a right circular cone with two spheres inscribed in the cone, and of the directrices as the meets

---

* The rest follows equally neatly by taking $\widehat{DPA} = \widehat{dpa} = \Phi$, constant (so that the angle between the radius vector Cp and tangent dp will also be Φ): it is immediate that $Cp = DE \times \tan \widehat{CEp} = CE \times \tan \frac{1}{2}(\frac{1}{2}\pi - \vartheta)$, where $\vartheta = \widehat{PCp}$, the angular height of P, so that the representing (polar equation) of the spiral will be

$$-\log\left(\frac{Cp}{Cp'}\right) = \cot \Phi \times \widehat{pCp'}.$$

Finally, taking $\Phi = \frac{1}{4}\pi$ and sphere radius $CE = 1 = Cp'$,

$$\int_0^{\vartheta} \sec\vartheta \cdot d\vartheta = \widehat{pCF} = -\log\left[\tan\tfrac{1}{2}(\tfrac{1}{2}\pi - \vartheta)\right].$$

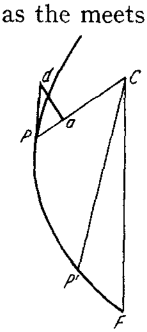


Fig. 36

[45] In such works as GREGORY ST. VINCENT's *opus geometricum*, Antwerp, 1647, and LA HIRE's *sectiones conicae*, Paris, 1685 (the first treatise on conics to absorb the newly found Books 5–7 of APOLLONIUS' *Conics*, rather badly published by BORELLI at Florence in 1661).

of the section-plane with the two planes through the respective circles of contact of cone and sphere are intuitively appealing.)

During the 17th century, however, we find a new and growing tendency to make the 3-space construction of the conic fundamental in its detailed treatment, a tendency which was to develop into the 19th century systematised treatment by synthetic methods of the geometry of conics and higher curves. Above all, the concept is introduced of invariance under optical projection from a point-centre (perspective invariance). Beginning with DESARGUES' (1639) *Brouillon proiect* ...[46] and PASCAL'S (1640) *Essay* and his lost treatises on conics[47] and continued in LA HIRE'S brilliant essay of 1673, *Nouvelle méthode en géométrie pour les sections des superficies coniques et cylindriques* ..., we have a rapidly growing study of such projective invariants as cross-ratio, involution, pole-polar correspondence and tangents, and of the corresponding non-projective properties which could now be seen as characterising the particular conic and differentiating it from conics of a different type. Nor was there any theoretical consideration which limited such an approach to conics, but historically the obstacle to such an extension was that, apart from a few properties of the corresponding CARTESIAN equation, little was known of the geometrical properties of the higher algebraic curves. NEWTON, in fact, was the first to carry through such an extension by classifying the various cubics into five distinct species, each of which is the set of possible optical projections of one of the five divergent parabolas, and then using analytical methods to separate out particular genera from each projective species.[48] (Presumably he could do so only after years of hard work spent in drawing innumerable particular cubics, and only gradually ordering and collating his crystallizing thoughts.)

At several places in his manuscripts[49] NEWTON has drawn up hurried drafts of the general basis on which such projective classification is extensible to n^{th} degree curves, but perhaps most interesting is his sketch[50] of how such an optical classification may be embedded in a general theory of 1, 1 point correspondences: "As we can from five simpler figures of the third order derive all figures of the same order, so we can all figures of higher orders from the simplest—and on that ground they can be differentiated into coordinate genera, positing that those are of the same genus which mutually transform into each other under projection. For that reason there is a single genus of second-order curves since they are all projections of the circle and of each other .... All those and only those which transform into each other under projection are cognate, and are

[46] See R. TATON (*op. cit.* note [2]).

[47] Whose contents can be reconstructed in a general way from the LEIBNIZ notes at Hanover—see PASCAL: *Oeuvres* (ed. BRUNSCHVIG & BOUTROUX), **2** (1908): 217—243, and especially 234—243: *generatio conisectionum.*

[48] This classification, as far as the handwriting of original drafts in the Portsmouth manuscripts can be dated, seems to have been carried out by NEWTON in the 1670's, though nothing was printed save the brief sketch (without proof-suggestions) of the *enumeratio linearum tertii ordinis*, London, 1704/1711.

[49] Compare *CUL Add.* 3961: 1ff.: *enumeratio curvarum trium dimensionum*; and 37ff.: *enumeratio curvarum secundi ordinis.*

[50] *Add.* 3963: 13ff.: *ejusdem ordinis lineae sic distinguuntur in genera coordinata oculo immoto* ... (the quotation is from 13R).

different in kind from those into which they do not transform. And so by the various cases of projection are families of curves to be split into species."

Such optical projection, to be fully effective in curve-classification, needed an accompanying construction technique which should derive the various projected forms in an analogously analytical way, and this NEWTON provides in his lemma 22 of *PM* Book 1[51]: *To transform figures into others of the same genus.* Taking as the figure to be transformed HGI, the point-set of G, we define the transform $(G) \to (g)$, the point-set hgi (in the figure-plane), in the following way: given the fix-point O (projection-centre) and fix-lines BH, BI, Bh $(=BH)$ given in direction, take $Od:OD=dg:DG$, where GD meets BI in D, OD meets BH in d and dg is drawn parallel to Bh. "By the same reasoning each point of the first figure will give a corresponding point of the new figure. Conceive then the point G as running with a continuous movement through all points of the first figure, and the point g with a like continuous movement will run through all points of the new figure and so describe it ...". Further, if the point G touches the first curve, we can see it as meeting the curve in two points, coincident in the limit, of which the corresponding two points of the transformed curve will also be coincident in the limit, and so the tangent to HGI at G transforms into a point-set—easily shown to be a line—tangent to the transformed curve.

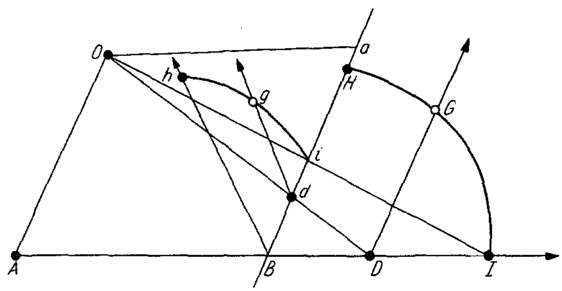


Fig. 37

With these preliminaries over, NEWTON comes to the point, showing that if the curves HGI, hgi are referred to respective ordinates GD, gd and abscissas AD, ad (where OA, Oa are parallel to DG, BD) and the "relatio" which relates the coordinate-lengths AD, DG is representable by an n-degree algebraic equation (in variables AD, DG), then the "relatio" which holds between ad, dg is also represented by a (different) n^{th}-degree equation (in variables ad, dg). For suppose $f(X, Y)=O$ is the n^{th}-degree polynomial which relates $AD=X$ and $DG=Y$: then $ad:OA=Od:OD=dg:DG=AB:AD$, and so $AD=\dfrac{OA \cdot AB}{ad}$, $DG=\dfrac{OA \cdot dg}{ad}$; or, where $ad=x$, $dg=y$; $OA=m$, $AB=n$, $f\left(\dfrac{m \times n}{x}, \dfrac{m \times y}{x}\right)=O$ relates ad and dg. Immediately, by multiplying through by x^n, this reduces to a new n^{th}-degree polynomial equation, so that "the curves defined by the points G, g are of the same analytical order".

This transform is, in more modern language, a 1, 1 point-correspondence $G \leftrightarrow g$, and therefore projective (though not simply perspective)—an aspect NEWTON introduces specifically[52]: "This lemma serves to resolve more difficult problems

[51] See H.W. TURNBULL: *The mathematical discoveries of Newton*, London 1945; 55—56; and J.L. COOLIDGE: *A history of the conic sections and quadric surfaces*, Oxford 1945: 46—47. The lemma occurs in *PM* (1687): 85—87.

[52] *PM* (1687): 87. The last part is slightly confused (in attempted clarification) in *PM* (1713) into "... may transform one of them, if an hyperbola or parabola, into an ellipse, and then the ellipse readily into a circle ...".

by transforming the given figures into simpler ones. So, any convergent right lines may be transformed into parallels by taking for first ordinate radius any line through their meet, for by that their meet is transferred into infinity.* This lemma is also of use in resolving solid problems, for as often as two conics occur by whose intersection a problem is to be solved we may transform one of them into a circle. Likewise a line and a conic ... may be transformed into a line and a circle". (Clearly the way is open for an elementary treatment of such projectively invariant concepts as the pole-polar relation.) In particular, any quadrilateral may be transformed into a parallelogram by taking the meets of opposite sides on AO (since the transform will project them both into infinity)—a property used in his immediately following propositions to derive simple constructions for conics through given points and touching given lines.

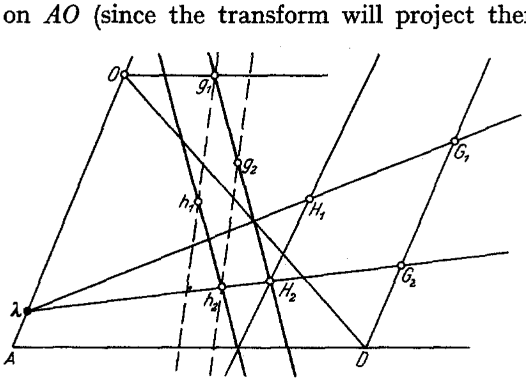


Fig. 38

As it stands in NEWTON'S form this transformation is a plane point-correspondence, seemingly detached from previous derivations of conic-projections as a 3-space point-correspondence on a cone-surface—indeed, its very baldness made it a thing little understood at the time, even by HALLEY, a mathematician in his own right.[53] It is tempting, however, for lack of direct evidence to connect it with a similar (1, 1) plane point-correspondence developed by LA HIRE in his (1673) *Les planiconiques.*[54] LA HIRE'S ideas are intimately connected with the standard (if

* Equivalently, we could show, where the points G_1, G_2 (G_1 in λH_1, G_2 in λH_2, λ in AO) are such that $G_1 G_2$ is parallel to BH_1H_2, their transforms g_1, g_2 are such that $g_1 g_2$ is parallel and equal to $h_1 h_2$—the first is trivial, and the second follows by:

$$\frac{H_1 H_2}{G_1 G_2} = \frac{od}{OD} = \frac{g_1 g_2}{G_1 G_2}, \; g_1 g_2 = H_1 H_2 = h_1 h_2.$$

[53] HALLEY wrote to NEWTON in the middle of checking proofs of *principia* (see HALLEY-NEWTON, 14 October 1686 ·≡· BALL: *An essay on Newton's principia* London 1893: 167—168, especially 167): "In your transmutation of figures according to the 22nd lemma ..., to me it seems that the manner of transmuting a trapezium [general quadrilateral] into a parallelogram needs some further explanation: I have printed it as you sent it, but I pray you please a little further to describe it by an example the manner of doing it, for I am not perfectly master of it: a short hint will suffice ..." In answer (NEWTON-HALLEY, 18 October 1686 ·≡· BALL, 168—169, or RIGAUD'S *Historical essay* ..., 43—47) NEWTON sketches the proof that the transform of a point G on OA is at infinity: "For the point G falling upon the line OA, the point D will fall upon the point A, and the line OD upon the line OA; and so, becoming parallel to aB, their intersection-point d will become infinitely distant, and so will its point g."

[54] Printed as pp. 73—84 of his *Nouvelle méthode en géométrie* ... The connection has been urged (in a slightly different way) by CHASLES in his *Aperçu historique* ... $_3$1889: Note **19**: 347—348: *Sur la méthode de Newton pour changer les figures en d'autres figures du même genre.*

unconventionally treated) derivation of conics in the proceding *Nouvelle méthode*... and without too much distortion we can conveniently abridge them as follows: Consider the (right) cone of base circle $hg'h'$ and vertex O (LA HIRE's "pole") cut by the plane hgh', and construct the parallel plane OLA (through O) which cuts the plane of the base-circle in the "directrice" line AL. Then, taking any line in the base-circle plane $g'K$ (which meets the "formatrice" hh' in K, and "directrice" AL in L) $g'A'A$ is the particular line which is perpendicular to hh', AL—we easily show that Kg drawn parallel to LO meets the generator-line Og' in a point g of the conic section.* If we now collapse the figure into the plane of the paper, lines and conics pass into lines and conics, and we have LA HIRE's plane transform: Taking $A'K$, AL as "formatrice" and "directrice", and fix-point O as "pole", any point g' is transformed into (unique) point g by drawing any line $g'KL$ through g' to cut $A'K$, AL in K, L and defining g as the meet of Og' with Kg drawn parallel to OL. (Clearly, the transform remains that of the 3-space perspective correspondence, lines passing into lines, conics into conics—and, indeed, n^{th}-order algebraic curves into n^{th}-order curves.) Finally, by introducing a few subsidiary lines into NEWTON's lemma we see how it may be reduced to LA HIRE's form. Visualising NEWTON's diagram in 3-space form, we consider the three (in

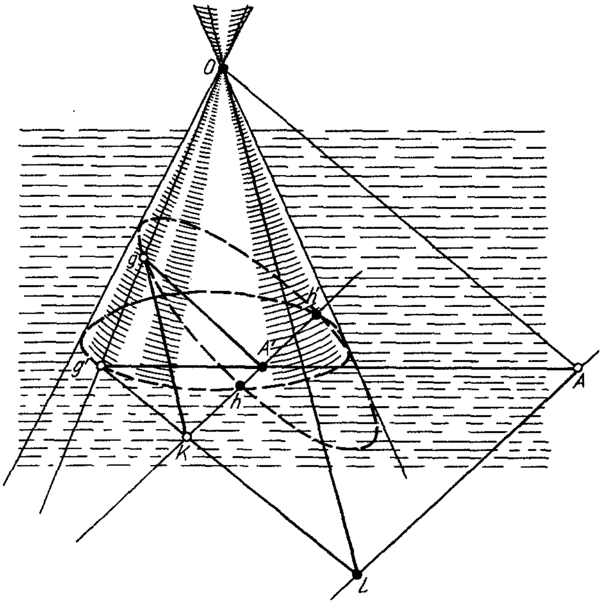


Fig. 39

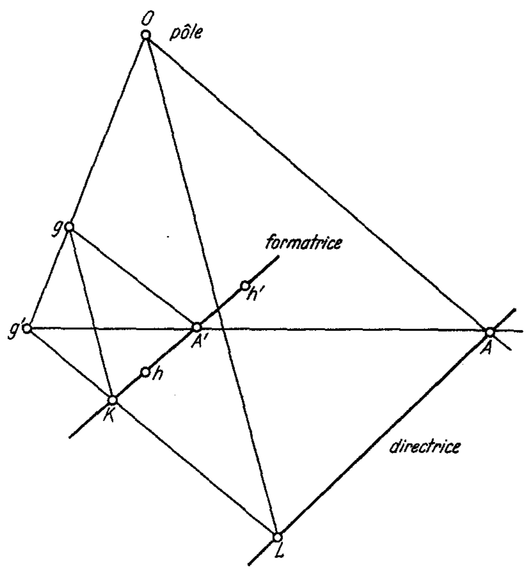


Fig. 40

* Since OLg' is a triangle with Kg parallel to OL, Kg does meet Og' (and in a unique point). But K is in the plane hgh' parallel to plane LOA, and so Kg meets the conic as well, and we can show one of the meets is g since the generator-line Og' meets the conic hgh' in (unique) point g.

general, oblique) planes *HBI*, *HBh*, *hBI* and some point-set *HGI* in the plane *HBI*. Define now, by a simple orthogonal transform (or rotation round *BI* when *HB* is perpendicular to *BI*) the curve *h'gI* in the plane *hBI* (where *g'* is defined from *G* by drawing *GD* parallel to *BH* and $Dg' = DG$ parallel to *Bh*). The point *g* thus becomes the meet of *Og'* with the plane *LBh* under the perspective transform of plane *IBh* into *HBh* which has *O* for optical centre, and we can then define the curve *hgi* as the (perspective) transform of *hg'I*. NEWTON's transform then becomes equivalent to LA HIRE's by viewing *O* as the "pole" and *hB*, *LA* as formatrix and directrix respectively.

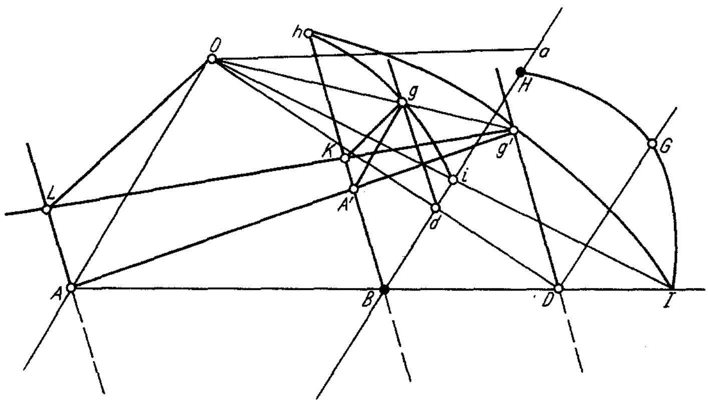


Fig. 41

Whether NEWTON did take his inspiration from LA HIRE's work is an interesting hypothesis, plausible if difficult to prove. (Certainly NEWTON had opportunity[55], but it is also fairly easy to derive the basic transform from first principles in a way analogous to LA HIRE's.) What is more important, however, is that NEWTON's treatment improves on LA HIRE's in its deeper consideration of the "relatio" which exists between ordinate and abscissa in both original and transformed curves. Above all it is significant that NEWTON used a method which is an equivalent of optical projection to justify his projective ordering of algebraic curves into (equivalence) classes—a rich basis for later elaborations in the theory of higher plane curves.

NEWTON was the supreme geometer of 17th century England, and it does not seem unjust to his contemporaries to dwell on the concepts which he introduced or refined. WALLIS, though he could introduce tesselations of the plane into his discussion of the geometrical flat-floor[56], was no pure geometer, and JAMES GREGORY—though he had a complete mastery of traditional techniques which enabled him, for example, to reduce ALHAZEN's problem of reflexion at a conical

[55] One of the rare copies of LA HIRE's *Nouvelle méthode* ... exists in the University Library, Cambridge. In other contexts, too, he had read and appreciated LA HIRE's work, particularly (1679) *Nouveaux élémens des sections coniques* ... (of which a copy exists in his library, now in Trinity College Library) and (1685) *sectiones conicae* (which he quotes approvingly in *PM* in the same section in which he gives his lemma 22).

[56] See his *mechanica, sive de motu* Oxford, 1670: Book *3* ch. 6: *de vecte* prop. 10.

mirror to that of drawing an ellipse whose foci are light-source and eye-point to touch the conic[57] (and later, in the 1670's, to solve the resulting quartic equation); and again to derive beautiful if miscellaneous propositions on lines, circles and conics in his *GPU*[58]—rapidly grew away from synthetic methods into analysis. NEWTON's teacher and friend at Trinity, ISAAC BARROW, has indeed a great many examples in his *LG* of elegant proof[59], but he remains—as HUYGENS' teacher PELL—only a thoroughly competent university don whose real importance lies more in his coordination of available knowledge for future use rather than in introducing new concepts.[60] Perhaps only CHRISTOPHER WREN, in the few years he spent at Oxford as Savilian Astronomer before directing himself to his life's work as an architect, can be set in comparison to NEWTON in creative originality. With his highly sensitive visual ability—he received training as an artist and draughtsman in his youth, and illustrated several contemporary medical and biological texts—he had a head start in an age when complex transformations were defined on a geometrical model and the faculty of mental visualization was a necessity for the geometer. Much of the work he did seems to have been lost, but the little which has been saved by WALLIS[61] is developed with an elegance which contrasts powerfully with WALLIS' own clumsinesses.

NEWTON himself, of course, had a thorough knowledge of classical geometry and contributed many elegant individual results[62] to the mighty (if slightly sterile) corpus of classical geometry. But, in comparison, none of this achievement has the richness and fertility of the new projective concepts, and is rather an ornament of an elaborated theory than the foundation of a fresh insight into the very concept of geometry itself which is the point- (and line-) correspondence. We have lost little in ignoring the one and emphasising the second—an approach which leads naturally into our next chapter: the introduction of analytical techniques into geometrical treatments, a topic which is unjustly lumped into the single vague idea of "CARTESIAN" coordinate geometry.

---

[57] Compare his *optica promota*, London 1663: prop. 34.

[58] See *GPU* (1668): 123–132, especially prop. 69: 128–130.

[59] Especially lectio *6* (ellipse and hyperbola properties), the appendices to lectiones *11* and *12*, lectio *13* (on general parabolas and hyperbolas).

[60] It is significant that the myth of BARROW's mathematical genius is the creation of WHEWELL in the 19th century and of J.M. CHILD in this: in contrast, MONTUCLA in his *Histoire* and CHASLES in his *Aperçu historique* place a lesser value on his mathematical pre-eminence.

[61] In WALLIS' *tractatus de cycloide ... de cissoide ...*, Oxford, 1659: especially (62–74) his work on cycloids, strictly comparable with PASCAL's similar work in *Lettres de A. Dettonville*, Paris, 1659; and (107ff.) his study of convolution transforms (with application to the study of the spiral forms of seashells in interesting anticipation of later studies of the logarithmic spiral in biological structures).

[62] Thus, for example, his treatment of general epicyclic forms in *PM* Book *1*: Section 10. But most important in its effects was his thorough knowledge of conic theory which allowed him, where others (including WREN, apparently) had failed, to furnish a proof that the conical path of a freely-failing body implies an attractive force directed towards a focus which varies inversely as the square of its distance from it (see *PM* Book *1*: Section 3—and the "Locke" proof, in slightly different form, of *CUL Add.* 3965: 1ff.).

VII. The expanding concept of geometry

2. The analytical approach

The development of analytical ("CARTESIAN") techniques is one of the more attractive aspects of 17th century geometry, but—despite a comparatively rich literature devoted to attempts at explication[1]—one not very well understood. Much of the difficulty of understanding derives from the misguided effort to read too many concepts which were developed later into the theory as it existed even in its late 17th century form, —probably under the impression that development from the 17th to 19th centuries was roughly an implementation and elaboration of existing concepts. But in its 19th century form analytical geometry is rather based on ideas of point-distance and invariance under transform to new axes conceived in the mid 18th century (and especially by EULER) than on original 17th century forms. Again, in previous historical evaluations many false trails have been laid which confuse the basic issues—in particular, a sterile search for anticipations and "pre-discoverers" has distorted a basic fact which should loom very large. Whatever the level to which the theory of latitude of forms had been advanced by the medieval calculators, and especially by ORESME, and whatever slight formulations are to be attributed to FERMAT in the same century, it is DESCARTES who, collating Greek coordinate systems with the analytical power of the free variable, which had been moulded in the 16th century to a fluid, usable state, laid the foundations of an analytical study of geometrical forms; and it was his *Géométrie*[2] which rapidly became standard in the new university mathematical courses in Western Europe from the middle of the century.[3] Nor did any contemporary mathematician—and least of all the great geometers NEWTON and HUYGENS—deny that fact.

To introduce the CARTESIAN viewpoint, then, I will consider in detail the problem which is basic to *Géométrie*, the solution of the Greek 3/4 line locus.[4] Given four fix-lines AB, AD, EF, GH meeting as shown in the figure[5] we wish to examine the nature of the point-set C such that, where CB, CD, CF, CH are drawn under given angles to them (meeting them in respective points B, D, F, H),

$$CB \times CF = CD \times CH.$$

---

[1] Compare, for example, the notes and bibliography in C. B. BOYER: *A history of analytical geometry*, New York, 1956.

[2] Published in appendix to his *Discours de la Méthode pour bien conduire sa Raison et chercher la Vérité dans les Sciences*, Leyden, 1637; but more importantly in the 1649 and the greatly augmented (2 volumes) 1659—1661 Latin editions. My argument is based on the original French version, edited by D.E. SMITH & M.L. LATHAM ($_2$1954).

[3] Thus WALLIS seems to have studied DESCARTES in the 1649 edition, NEWTON in both 1649 and 1659/1661 editions, while HUYGENS, under VAN SCHOOTEN's tutelage, used the original French.

[4] Introduced in Book 1: 304—314, but discussed in detail in Book 2: 324—335; compare G. MILHAUD: *Descartes savant*, Paris, 1921: ch. 6: 124ff.

[5] Slightly adapted and enlarged from DESCARTES'. It is a mistake common to all standard editions that the point-set of C, which should pass through G, the meet of AB, GH, does not (though the error is detected by DESCARTES himself in his letter to VAN SCHOOTEN in September, 1639). See DESCARTES: *Oeuvres* (ed. ADAM & TANNERY), 2: 574—582, especially 574ff. · ≡ · *Correspondence* (ed. ADAM & MILHAUD), 3: 315—320, 315ff.

DESCARTES begins[6]: "First I consider the thing already done, and to rid myself of the confusion of all these lines I consider one of the given lines and one of those we have to find, for example AB and BC, as the principal ones, and so try to refer all the others to them. Let the segment of the line AB between the points A and B be named x, and BC be named y, and let all the other given lines be prolonged till they cut these two, prolonged as far as necessary and if they are not parallels". Thus, take AB, BC meeting AD, EF, GH in A, R; E, S; G, T respectively. Then, where x and y are measured in the directions shown in the figure, since all the angles in the figure are given, we have, say $AB:BR=z:b$ (constant), or $RB=\frac{bx}{z}$ and $CR=y+\frac{bx}{z}$; and similarly, where $CR:CD=z:c$ (constant), $CD=\frac{c}{z}\left(y+\frac{bx}{2}\right)$. Further, where we denote the fix-lengths AE, AG by k, l, $EB=EA+AB=k+x$, and for $BE:BS=z:d$ (constant) $BS=\frac{d}{z}(k+x)$ and $CS=y+\frac{d}{z}(k+x)$, while for $CS:CF=z:e$ (constant) $CF=\frac{e}{z}\left(y+\frac{d}{2}(k+x)\right)$; and again for $BG:BT=z:f$ (constant), since $BG=l-x$, $BT=\frac{f}{z}(l-x)$ and $CT=y+\frac{f}{z}(l-x)$, while for $TC:CH=z:g$ (constant) $CH=\frac{g}{z}\left(y+\frac{f}{z}(l-x)\right)$. Finally the defining condition $CB\times CF=CD\times CH$ can be represented by

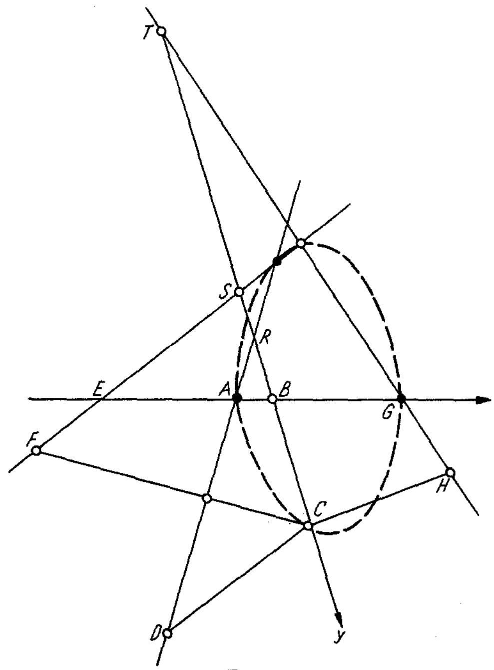


Fig. 42

$$y\times\frac{e}{z}\left(y+\frac{d}{z}(k+x)\right)=\frac{c}{z}\left(y+\frac{bx}{z}\right)\times\frac{g}{z}\left(y+\frac{f}{z}(l-x)\right),$$

which is a 2-degree equation in x and y,

$$\left(y+\frac{n}{z}x-m\right)^2=m^2+o\,x+\frac{p}{m}x^2,$$

where the constants are suitably defined.*

* In fact $z^2(ez-cg)\,y^2+z(dez+cfg-bcg)\,xy+bcfgx^2+z(dekz-cfgl)\,y-bcfglx=0$ (and clearly the point-set is through $\begin{cases}x=0\\y=0\end{cases}$ and $\begin{cases}x=l\\y=0\end{cases}$, or points A and G); so that

$$2m=\frac{cflg-dekz}{z^2(ez-cg)},\qquad \frac{2n}{z}=\frac{dez+cfg-bcg}{z^2(ez-cg)},$$

$$o=-\frac{2mn}{z}+\frac{bcfgl}{z^2(cz-cg)},\qquad \frac{p}{m}=\frac{h^2}{z^2}-\frac{bcfg}{z^2(ez-cg)}.$$

[6] *Géométrie*: 310.

Simplifying DESCARTES' rather unsure further argument, we can take this by $y'^2=\lambda(x'+\mu)(x'+\nu)$, where $y'=y+\frac{n}{z}x-m$, $x'=\frac{r}{z}x$, $\lambda=\frac{z^2p}{r^2m}$, and the constants μ, ν are found by equating coefficients. Returning now to the geometrical model, we can represent x', y' by IL, CL, where $IK(=AB):KL:IL=z:n:r$, and AI, IK are drawn parallel to BC, AB such that $AI=m$,* and we reduce the 3/4 line condition to the point-set (x, y) which satisfies the representing equation $y'^2=\lambda x'^2+\lambda(\mu+\nu)x'+\lambda\mu\nu$, where I is a fix-point in the plane and $IL=x'$, $LC=y'$ are given in direction.

Before sketching in DESCARTES' final solution (which shows that the point-set of C is a conic, possibly degenerate), let us consider in detail the ideas which DESCARTES has introduced. First, implicitly he has brought in the concept of dimension and the assumption that by choice of a suitable line-length AB a second line-length inclined at some angle through B can be made to pass through any point C in the plane[7]—an axiom which virtually defines the plane, and, as such, was assumed by all his contemporary mathematicians as well as DESCARTES as "self-evident".** In a straightforward way[8] DESCARTES supposes that we can attach (real-number) measures to both the line intervals AB, BC—defined in a suitable sense (indicated in my diagrams by an arrow pointing in the positive direction)—such that with respect to a conventional unit, this measure is the EUCLIDEAN length of the lines AB, BC: in effect, we define a 1, 1 correspondence between the points of a line extending to infinity in either direction and the numbers of the real interval $[-\infty, +\infty]$. This procedure yields, of course, the classical "CARTESIAN" order-

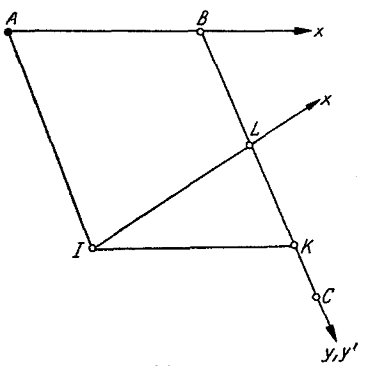


Fig. 43

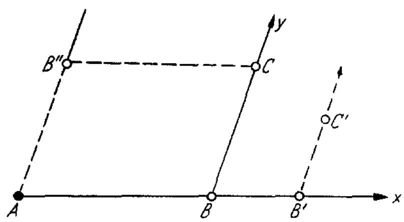


Fig. 44

* Since

$$BL=BK(=AI)-LK\left(=\frac{n}{z}\times IK\right)=m-\frac{n}{z}\times x, \quad \text{and} \quad CL=CB(=y)-BL.$$

** Probably this means little more than "consistent with the EUCLIDEAN scheme of geometry" (with the proviso that no other system of geometry is acceptable). The concept of dimension is, indeed, an extraordinarily difficult thing to pin down, and a suitable definition has to allow that 1, 1 point correspondence (though not 1, 1 correspondences of the ε-neighbourhoods of points) is not a dimensional invariant. We cannot, therefore, fairly attack DESCARTES for assuming what is, in fact, a possible definition of a RIEMANNIAN 2-space (one particular member of the family being the EUCLIDEAN plane).

[7] In the EUCLIDEAN scheme, of course, the axiom that no two distinct parallels can be drawn through the same point shows the uniqueness of the procedure.

[8] The idea is as old as cartography.

ing of 2-space (and by easy extension n-space) by the 1, 1 correspondence which exists between every point C and the (unique) values of the measures of $AB(x)$ and $BC(y)$—later to be denoted by the ordered pair (x, y). (The equivalent procedure of considering a second axis $B''A$ through fix-point A such that $B''A$ is parallel (and equal) to CB, and so AB to $B''C$, and the defining point C by the 1, 1 correspondence of $AB(x)$, $AB''(y)$ with the ordered number pair (x, y) came into general use only in the 18th century[9].)

None of this is new with DESCARTES, but more important there is his implying no limitation, geometrical or analytical, which restricts his coordinate system to being EUCLIDEAN. In modern treatments this restriction is introduced by defining the concept of "point-distance" by the analytical equivalent of "PYTHAGORAS'" Theorem: for given points $C_1 \equiv (a_1, b_1)$, $C_2 \equiv (a_2, b_2)$ the distance between C_1, C_2 is

$$\text{Dist}(C_1, C_2) = [(a_1 - a_2)^2 + (b_1 - b_2)^2 - 2(a_1 - a_2)(b_1 - b_2) \cdot \cos\vartheta]^{\frac{1}{2}}$$

where ϑ is the angle between AB and BC:

$$= [(a - a_2)^2 + (b_1 - b_2)^2]^{\frac{1}{2}},$$

where AB is normal to BC. DESCARTES, however, uses the somewhat different, if equivalent[10] concept of triangles given "in species"; that is, whose sides are given in direction, and so in proportion with the angles of inclination given in absolute magnitude (so that all members of the set of triangles given in the same species are similar)—a most important aspect of his procedure slurred over in modern accounts.

Next, taking "unknown" (free variable) quantities x, y for the line-lengths AB, BC, DESCARTES reduces a given defining equation on the point C, represented geometrically as some relation between line-lengths, to an equivalent analytical representing equation between x and y, say $f(x, y) = 0$, where the relation f is specified by reduction of the original condition into an analytical form: conversely, each particular relation $f(x, y) = 0$ connecting x and y defines a particular point with respect to coordinate line-lengths $AB(x)$, $BC(y)$ in a EUCLIDEAN plane.

Finally, in a beautiful generalisation, DESCARTES replaces the condition that each point so defined be restricted by $f(x, y) = 0$ by the free-variabled condition that the point-set whose members are the particular points defined is restricted in its analytical model by the representing equation $f = 0$ for all x, y.*

The concept of point-set as, virtually, the class of particular points which satisfy some restricting condition had, of course, been developed in classical Greek

* More formally, by $(x, y)\ (f(x, y) = 0) : x, y \in [-\infty, +\infty]$.

9 Particularly through the influence of EULER's *introductio*—*cf.* BOYER (*op. cit.* note 1). The concept was known, however, to the 17th century mathematician, and LA HIRE, for example, sets the construction up with a terminology of "tige" and "rameau". See his *Les lieux géométriques*, Paris, 1679: introduction.

10 As WALLIS showed in his *de postulato quinto et definitione quinta lib. 6 Euclidis disceptatio geometrica* (given originally as a lecture in the early 1660's but printed in his *opera mathematica* 2 (1693); 665—678—*cf.* UGO CASSINA: *Sulla dimostrazione di Wallis del postulato quinto d'Euclide*, Act. Congr. int. Hist. sc. (8): Roma, 1956: 33—38), the postulate of the existence of similar triangles is equivalent to that of the parallel postulate, and so defines the metric to be Euclidean.

geometry. Thus, the circle was seen as the set of points which are at a constant distance from a given point, and certain algebraic curves—notably the cissoid and the conchoid—had been so defined by simple line- and circle-intersection properties. More obscurely, in the development of "porism" theory[11] general sets of conditions on a point had been shown to imply that the point-set was a line or circle. But more important of all (and most generally in APOLLONIUS' *Conics*) the conic defined as the plane section of a two-sheeted cone, had been reduced to an equivalent plane defining condition, its "symptom": where A, D are fix-points on a given line and B a variable point on it with BC a line at a given

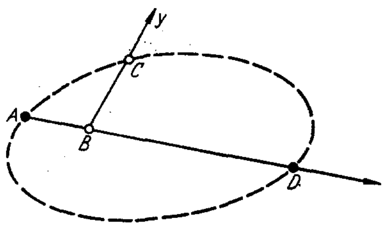


Fig. 45

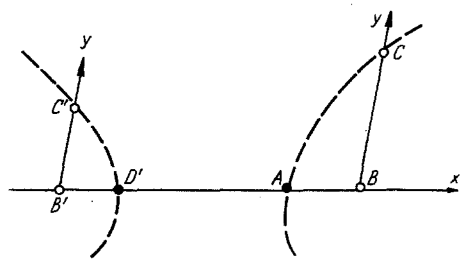


Fig. 46

constant angle to AB, then C is on a conic if the ratio $\frac{BC^2}{AB \times BD}$ is constant for all points C—an ellipse or hyperbola according as AB, BD are taken in the same or different senses $\Big($and the limiting case of each, where one of the fix-points, say D, is at infinity—or $\frac{BC^2}{AB}$ is constant—is a parabola$\Big)$.

Introducing CARTESIAN coordinates we see immediately, that, where $AB = x$, $BC = y$, $AD = a$, the defining analytical equations are respectively $y^2 = \lambda x(a - x)$, $y^2 = \lambda x(a + x)$ and $y^2 = \lambda' x$, but it is obvious that there are great difficulties in the way of such an interpretation till we have an adequate analytical concept of free variable—and the Greeks never departed from the purely geometrical model. DESCARTES was, in a worthwhile sense, lucky in that he could draw on just such an adequate concept of free variable for the basis of *Géométrie*—no analytical geometry was possible without it, but with it the development of an analytical theory of conics was immediate, merely requiring transposition of the Greek plane "symptoms" into a free-variabled algebraic form.*

Returning to DESCARTES' reduction of the 3/4 line locus to the defining equation $y'^2 = \lambda x^2 + \lambda(\mu + \nu)x' + \lambda\mu\nu_1 = \lambda x'' \times (x'' + \nu - \mu)$ where $x'' = x' + \mu$, it is now clear that, apart from degenerate cases, the locus is an ellipse or hyperbola according as λ is greater or less than zero (and a parabola when no term in x'^2 is present), and this is DESCARTES' solution[12]. As for the degenerate cases, $y'^2 = 0$

* In the circumstances, we can only be surprised that so much of *Géométrie* should be concerned with the analysis of equations if we accept a modern viewpoint which sees the procedures there developed as mere algebraic technique. Rather, at a deeper level, much of *Géométrie* is concerned with exploring bounding conditions on the general free-variable polynomial—a study directly related to the analogous theory of the geometrical point- (and line-) set.

[11] See previous chapter.

[12] In much more detail—see *Géométrie*: 327—333.

is clearly the (doubled) line $y' = y + \frac{1}{2}nx - m = 0$ (which we can then take as the general equation of a line in the plane) and, where $\mu = \nu$, $y'^2 = \lambda(x' + \mu)^2$, or $(y' + \lambda^{\frac{1}{2}}(x' + \mu))\,(y' - \lambda^{\frac{1}{2}}(x' + \mu)) = 0$, a line-pair (though DESCARTES admits only one line, apparently omitting the negative value of the square root[13]).

The attempt to apply a similar procedure to other problems treated in the *Géométrie*[14] lacks power in general, especially in the introduction of the unwieldy circle method for finding the subnormal at a point on the curve (and so indirectly the subtangent).[15] Yet a wealth of ideas and suggestions was put forward which hinted, for example, —not quite accurately—at a general classification of algebraic curves by the degree of the representing polynomial, and we can without exaggeration say that *Géométrie* was a rich store-house of thoughts awaiting verification and elaboration and extension in the learned commentary. In the half-century after it appeared the study of analytical geometry is largely the history of the improvement and, in some cases, considered rejection of ideas original with DESCARTES.

In England WALLIS was the first to expound the CARTESIAN method in his *de sectionibus conics*[16] perhaps indeed the first elementary textbook of conics treated by CARTESIAN methods, and his treatment, certainly in no way profound, had the virtue of being clear and simple. In 44 propositions (and 108 pages) conic theory was developed from a basic definition as sections of a right cone, geometrically reduced to the APOLLONIAN plane "symptom", into the analytical equivalents of the (easily manipulable) 2nd-degree equations, $e^2 = d\left(l - \frac{l}{t}d\right)$, $p^2 = ld$, $h^2 = d\left(l + \frac{l}{t}d\right)$ (where e, p, h are ordinates of the ellipse, parabola and hyperbola respectively with d, the abscissa, measured from coordinate origin at a vertex of the conic, l the latus rectum and t the length of the main diameter conjugate to the ordinate), with a brief consideration of the elementary defined concepts of tangent (and subtangent) and diameter. The work remains extremely readable, developing a firm basis for the ideas thrown out by DESCARTES in his resolution of the 3/4 line locus, but conceptually derivative. In an interesting appendix[17] however, WALLIS tries to extend the CARTESIAN approach to higher plane curves (and specifically[18] to the cubical paraboloid). Thus[19] where we take the point-set of P defined by $y^3 = a^2x$ with respect to rectangular coordinates $OX = x$, $PX = y$, he deals quite successfully with the problem of finding the subtangent $TX = t$ at any point on the curve.* Assuming that the cubical parabola is "everywhere" convex, he considers a second point P' on the curve (with corresponding abscissa $OX' = x'$) which he will take infinitely near to P. Let $P'X'$

* The approach is that catalogued as "FERMATIAN" in chapter X.

[13] Compare (p. 328): "... ce point C se trouveroit en une autre droite qui ne seroit pas plus mal aysée a trouver qu'*IL* ...".

[14] Especially in Book 2 (Book 3 is concerned with applications to the solutions of equations, and in particular the isolation of roots by interesting conics).

[15] See chapter X.

[16] *de sectionibus conicis nova methodo expositis tractatus*, dated on title-page 1655, but issued as part 2 of *operum mathematicorum pars altera*, Oxford 1656.

[17] *op. cit.* 104–112.

[18] *op. cit.* props. 46–47: 106–110.

[19] *op. cit.* prop. 46: 106.

meet tangent PT in X'': then $X'P' < X'X''$ with equality in the limit as $P' \to P$: or $(X'P)^3 < (X'X)^3$ (with equality in the limit); but

$$(X'P')^3 = y^3 = a\,x' = (x + \varepsilon)\,\frac{y^3}{x}$$

and

$$(X'X'')^3 = \left(\frac{X'T}{XT}\right)^3 PX^3 = \left(\frac{t+\varepsilon}{t}\right)^3 y^3, \quad \text{where} \quad \varepsilon = x' - x,\ ^{20}$$

so that in the limit as $P' \to P$ $(x' \to x,\ \varepsilon \to 0)$ we can equate these values, and have on reduction $t = \lim_{\varepsilon \to 0} x\left(3 + \frac{3\varepsilon}{t} + \frac{\varepsilon^2}{t^2}\right) = 3x$.

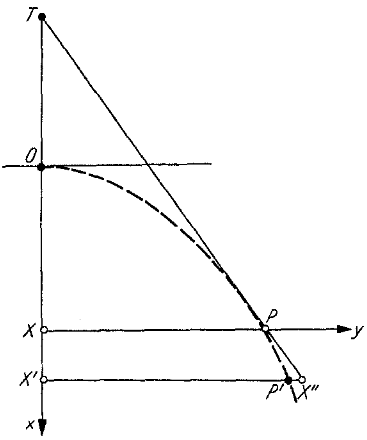


Fig. 47

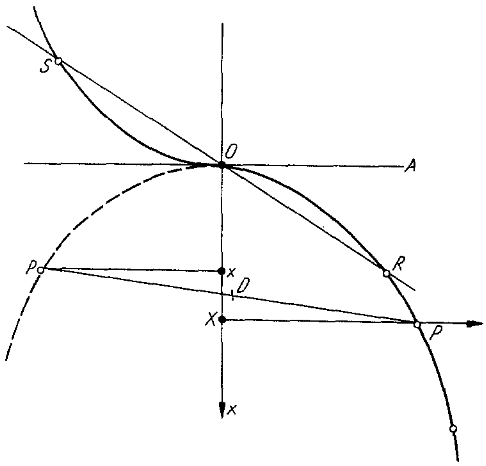


Fig. 48

In the following proposition[21], however, he loses control over the method, assuming (on the analogy of the APOLLONIAN parabola) that the curve continued past the vertex O to P will lie on the same side of OA, tangent at the vertex, and on this basis tries to develop the concept of diameter: specifically he assumes that any chord pP through two points of the curve will not meet it again (in a distinct point, at least), and so tries to find the point-set of D, the chord's mid-point—in the case of the simple parabola, of course, a parallel to the axis. In fact, WALLIS finds a cubic representing equation and concludes the cubic parabola has no simple diameter.[22] His mistake, of course, is that the curve continues past O on the *opposite* side of the vertex tangent, and so he recognizes it in the long dedication* of his *adversus M. Meibomii de proportionibus*[23]—that is, that a general line PRS may meet the curve in three points (of which two may not

* To BROUNCKER, who in private correspondence had pointed out a clear counter example.

[20] WALLIS uses the (confusing) FERMATIAN o in the original.

[21] *op. cit.* prop. 47: 107—110.

[22] *op. cit.* 110: "propterea ejusmodi parallelae diametri in paraboloeide cubicali non reperiuntur".

[23] *adversus M. Meibomii de proportionibus dialogum tractatus elenticus*, printed in *operum mathematicorum pars prima*, Oxford 1657.

"exist"). With the mistake acknowledged it is an easy step to set up (NEWTON'S) definition of the diameter of a cubic as the point-set of the mean of the (three) meets of a general line with it.[24]

Several important points arise out of this example. First and most obviously we realize how little the CARTESIAN coordinate framework was understood, that the very ease with which it could be used as an algorithm could hinder appreciation of its structure. Yet we must not make too much of this (and of the allied difficulty of the concept of a negative quadrant)—WALLIS' example shows how easily readjustment was made.* Indeed, too little advantage was to be derived from the imperfectly polished free-variable concept accepted as standard in the period with its restriction of the variable range to the positive interval $[0, \infty]$ (so that for $x \in [-\infty, 0]$, the clumsy transform $x = -y$, $y \in [0, \infty]$ had to be made). In the ensuing proliferation of particular cases and corresponding "tied" signs $\pm$ (where the top signs are to be taken together as, say, a positive range of the variable, the bottom as the complementary negative instance) the basic unity of the CARTESIAN framework was easily obscured—though, again, we must not insist too strongly on the point: the transition to the full variable range is a natural extension which merely absorbs the signs $\pm$ into the variable restricted to a positive range.

Further, we find the important idea—originally, if implicitly, in *Géométrie*—that the order of a curve can be defined by the (upper bound of) the number of its meets with a general line in the plane. While WALLIS uses the concept only to modify a false viewpoint, NEWTON was to make it basic in many applications, but especially in his classification of cubics[25], showing the close connection with the general cubic representing polynomial; and more generally MACLAURIN, professedly developing NEWTON'S ideas, was later to reveal[26] how NEWTON'S organic construction of a conic could be generalized into a mapping of combinations of algebraic curves into an algebraic curve whose degree is a simple function of the particular degrees of the defining curves (the precise nature of which varies, of course, with the type of mapping).

Above, all, WALLIS' treatment typifies a general lack of knowledge in the mid-century about the form of general algebraic (and transcendental) curves other than the conic (but including the line treated analytically). Quite suddenly the mathematical world had been presented with a powerful technique for examining curves of general form, only to find that there were few existing known higher curves on which to practise it (and those defined by non-general properties of products of line-segments). Inevitably increase in knowledge of the higher curves was slow-paced, even uneventful, and the atmosphere of the work carried through

* In fact, the concept of quadrant (in the CARTESIAN plane) did not really assert itself till the systematic introduction of coordinate-axes as reference-frame replaced the existing abscissa-ordinate construction.

[24] Developed in manuscript from the middle of 1664—compare *CUL Add.* 4004: 15V—27V. WALLIS himself, in fact, considers the set of parallels QRS defined by $y = rx + s$ (r constant, s free) whose substitution in $y^3 = a^2 x$ gives a cubic in x, and so three values for x (positions of P).

[25] Compare *CUL Add.* 3961: *passim*, and his printed *enumeratio* ... .

[26] In his *geometria organica, sive descriptio linearum curvarum universalis*, London 1720.

in analytical geometry in the half century is, at first glance, that of drab consolidation. With increasing facility the shapes of the more interesting higher curves, especially particular cubics and quartics, became familiar—is it a comment on the dreariness of this process that they were given such vivid names?—and were to some extent published in the many elaborate commentaries on the *Géométrie* which appeared in the period[27], but meanwhile the conic held the field, with the line more easily and more naturally, it was thought, treated by the methods of pure geometry.

In short, typical above all of the period is a growing systemisation of treatment of the general 2nd-degree polynomial, and especially the development of standard reductions to canonical form. Thus LA HIRE, in his *Les lieux géométriques*[28], gives a more or less simplified treatment of what DESCARTES had sketched. Typically, his example 2 considers[29] the equation $x^2+y^2-xy-ay=ab$, which he takes in the form $z^2+\frac{3}{4}v^2=\frac{1}{3}a^2+ab$, where $z=y-\frac{1}{2}x-\frac{1}{2}a$ and $v=x-\frac{1}{3}a$. Taking CARTESIAN abscissa $On=x$ and ordinate $nl=y$, with $OM=\frac{1}{3}a$ (or $Mn=x-\frac{1}{3}a=v$) defining point M on On, we draw the locus C—a line, in fact—such that $(y=)\ Cn=\frac{1}{2}On(=\frac{1}{2}x)$ and the parallel line AB such that BC (on ln) $=\frac{1}{2}a$ $\big($or $lB=ln-(BC+Cn)=y-(\frac{1}{2}a+\frac{1}{2}x)\big)$ then where MA is drawn parallel to ln (meeting AB in A) and supposing $AB:Mn=r:s$ (constant) $\left(\text{or } AB=\frac{r}{s}v=v', \text{ say}\right)$, we can take the equation as $z^2=\lambda(\alpha^2-v'^2)$, where $\lambda=\frac{3}{4}\left(\frac{r}{s}\right)^2$ and $\alpha^2=\frac{1}{\lambda}\left(\frac{1}{3}a^2+ab\right)$. Finally, seeing the fix-point A as the new CARTESIAN origin and $AB=v'$, $Bl=z$ as the new abscissa and ordinate connected by this representing equation, the locus of l is an ellipse whose centre is at A and parameter $2\alpha\lambda$ and whose diameter AB (of length 2α) lies along the line AB.*

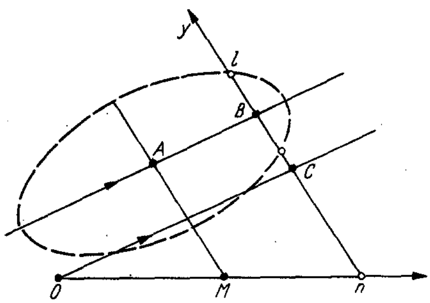


Fig. 49

LA HIRE states his method very clumsily, but the general pattern is clear, and the necessary last touches which complete study of the general 2nd-degree curve were given an exhaustive treatment by JOHN CRAIG.[30] Though CRAIG does not admit any influences his method is clearly a modification of DESCARTES' (and perhaps also of that taught at Cambridge by NEWTON in his Lucasian lectures of the 1670's[31]).

* Since the representing equation can be put in the form $\frac{z^2}{(\alpha+v')(\alpha-v')}=\frac{2\alpha\lambda}{2\alpha}$.

27 Compare BOYER (*op. cit.* note 1): ch. 6: 103—137.

28 Printed as Book 2: 179—293 of his *Nouveaux élémens des sections coniques, les lieux géométriques, la construction ou affectation des équations*, Paris 1679.

29 *op. cit.* 274—278.

30 In *nova methodus determinandi loca geometrica* printed as part 2: 62—76 of his *tractatus mathematicus de figurarum curvilinearum quadraturis et locis geometricis*, London, 1693.

31 These lectures are now in Cambridge University Library (*CUL Dd.* 9. 68), and formed the basic of his *AU*.

So, given the general 2nd-degree equation

$$A x^2 + 2H xy + B y^2 + 2G x + 2F y + C = 0,$$

we can (with NEWTON) put it in either of the equivalent forms,

$$z_1^2 = (A x + H y + G)^2 = (H^2 - AB)\, y^2 + 2(GH - AF)\, y + (G^2 - AC),$$

or

$$z_2^2 = (H x + B y + F)^2 = (H^2 - AB)\, x^2 + 2(FH - BG)\, x + (F^2 - BC),$$

and this is apparently CRAIG's basis for classification also. In particular, where $x' = \varrho x + \sigma$, for suitably chosen ϱ and σ the first form can be reduced to $z_1^2 = \lambda \cdot (\alpha^2 - x'^2)$, $z'^2 = \mu \cdot x'$, $z_1^2 = \lambda'(\alpha^2 + x^2)$ according as H^2 is greater than, equal to or less than AB, and the familiar test for conic-type is immediate*. CRAIG uses this idea in pursuance of his ideal: to give a systematic geometrical construction of every point-set which has a 2nd-degree algebraic representing equation. Thus, in his theorem 3 he develops[33] a general construction for those point-sets which have, in the above general 2-degree form, $H^2 < AB$ (and so are ellipses)—specifically he gives the derivable general equation

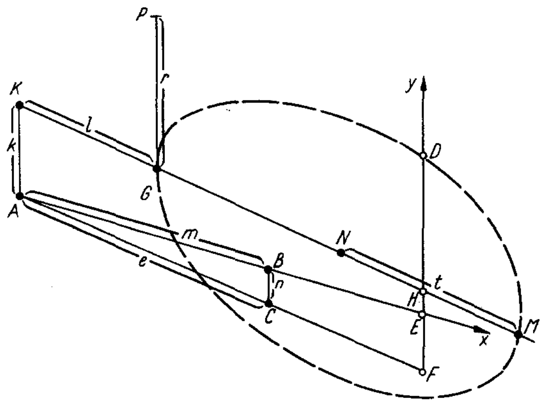


Fig. 50

$$\left(y + \frac{n}{m} x - k\right)^2 = \frac{r}{2t}\left(2t - \frac{e}{m} x + l\right)\left(\frac{e}{m} x - l\right),$$

or $\quad z^2 : (2t - x')\, x' = r : 2t,$

where $z = y + \frac{n}{m} x - k$ and $x' = \frac{e}{m} x - l$. Clearly this is an ellipse of transverse diameter $2t$ and parameter r, and CRAIG's construction of it closely follows LA HIRE: taking abscissa $AE = x$ and corresponding ordinate $ED = y$, construct the triangle ABC, where BC is drawn parallel to ED such that $AC : AB : BC = e : m : n$, and make AK, parallel to ED, $= k$, then, taking points G, N, M on the parallel through k to AC such that $KG = l$ and $GN = NM = t$, the required ellipse has centre N, transverse diameter $GM = 2t$ and parameter $PG = r$.** Finally CRAIG

* CRAIG, however, does not seem to know the test for degeneracy (a corollary of this approach)—certainly known to NEWTON in the 1670's—that the right side be a perfect square[32], *viz*: $(H^2 - AB)\,(G^2 - AC) = (GH - AF)^2$, or

$$(H^2 - AB)\,(F^2 - BC) = (FH - BG)^2 \left[\equiv \begin{vmatrix} A & H & G \\ H & B & F \\ G & F & C \end{vmatrix} = 0 \right].$$

** For, where ED meets GM in H, $DH = DE + EF - HF = y + \frac{n}{m} x - k$, and $GH = KH - KG = \frac{l}{m} x - l$, so that the equation is the analytical representation of the geometrical "symptom", $DH^2 : GH \times HM = r : 2t$.

[32] Compare *AU*: prob 57: 156—157, for example.

[33] *tractatus mathematicus*: 71—73.

expands into the full form,

$$y^2+2\frac{n}{m}xy+\left(\frac{n^2}{m^2}+\frac{re^2}{2tm^2}\right)x^2-2kxy-\left(2\frac{n}{m}k+\frac{re}{mt}l+\frac{re}{m}\right)x+$$
$$+\left(k^2+rl+\frac{rl^2}{2t}\right)=0,^{\star}$$

which is to serve as the canonical form of ellipse-construction, applicable to particular equations by suitable coefficient comparison.**

In all this there is nothing new conceptually, but it is a compact systematic exposition which remained standard during the next half-century[34] till EULER made a thorough restudy of the representation and construction of the 2nd-degree equation using coordinate axes.

What made all this detailed development comparatively easy was, of course, that in the APOLLONIAN theory of conics there was a basis already worked out which developed a corresponding geometrical approach with respect to analogously defined line-lengths and an equivalent point-set definition of the conic. No such geometrical basis existed for higher curves—nor indeed for the straight line—and it is a plausible hypothesis that the very elaborateness of conic theory was more of a hindrance than an aid to the formulation of general analytical treatments. As we have seen[35] NEWTON'S *enumeratio linearum tertii ordinis* was the first attempt—on the basis of a long experience of particular forms of the cubic—to classify the general 3rd-degree curve into species analogous to the three types of (non-degenerate) conic[36], but it is significant that only in a projective classification (distinguishing five projective classes) does he discuss the general cubic in terms analogous to those he uses with regard to the general conic.*** However NEWTON'S work was only a rough draft of a possible line of development which

---

* Clearly $H^2=\frac{n^2}{m^2}<\frac{n^2}{m^2}+\frac{r^2e^2}{2tm^2}=AB$.

** Typically, in his example 2 CRAIG considers the equation $y^2-2ay+x^2=0$: comparing coefficients $\frac{2n}{m}=0$ (or $n=0$ and we take $m=e$), $k=a$, $\frac{re^2}{2tm^2}=1$ (or $r=2t$), $\frac{re}{mt}l+\frac{re}{m}=0$ (or $l=-t$), and $k^2+rl+\frac{rl^2}{2t}=0$ (or $t=a$), so that the equation is $(y-a)^2=(x+a)(2a-(x+a))$; or AG is parallel to GM, and when $AG=0$, $DE=0$, $2a$ (and the ellipse is a circle when $\widehat{AED}$ is right).

*** Though he does, for example, outline how particular analogous concepts can fruitfully be isolated in the case of the cubic, and especially that of "diameter" which is the (provably linear) point set of the generalized arithmetic mean $\sum_{1\leqq i\leqq 3}(X_i X)=0$, where the X_i are the three meets of a co-parallel set of lines with the cubic.

[34] It is, for example, adopted by L'HOSPITAL in his *Traité analytique des sections coniques*, Paris 1707: 213ff.; and as late as 1748 by COLIN MACLAURIN in his *Treatise of algebra* ...: part 3: *Of the application of algebra and geometry to each other*, especially ch. 2: 325—352.

[35] See previous chapter.

[36] Compare H. HILTON: *Newton on plane cubic curves* in *Isaac Newton, 1642—1727*: 115—116; and especially W.W.R. BALL: *On Newton's classification of cubic curves*, Proc. London Math. Soc. **22** (1890): 104—143, where he examines the drafts of the *enumeratio* (more detailed than the printed version) which are to be found in *CUL Add.* 3961.

others—MACLAURIN in his (1720) *geometria organica* was perhaps the earliest[37]—were to elaborate into a general description of higher curves.

Analogous CARTESIAN treatments of EUCLIDEAN 3-space developed even more slowly—but earlier than many historians have allowed. In the few examples which exist in the period, the construction of the basic reference-system of co-ordinate line-lengths is an extension of the abscissa-ordinate one of 2-space: a general point is defined (uniquely) by the lengths of three (non-coplanar) line-lengths each given in direction, OX, XP, PQ, each of whose measures (x, y, z respectively) may vary over the real interval $[-\infty, +\infty]$. As before the assumption of triangles given in species (the postulate of similarity) implicitly restricts the space to being EUCLIDEAN. Very often the directions of OX and XP are seen as defining a (unique) plane OXP in which they lie: then the direction QP through a general point Q outside the plane will define a unique corresponding point P in the plane, and the treatment is suitably reduced to a more controlable treatment in the plane OXP. Finally, as with the CARTESIAN method in the plane, we note that little use is made of the equivalent concept of definition of the general point Q with respect to fixed coordinate-axis lengths OX, OP', OQ'.

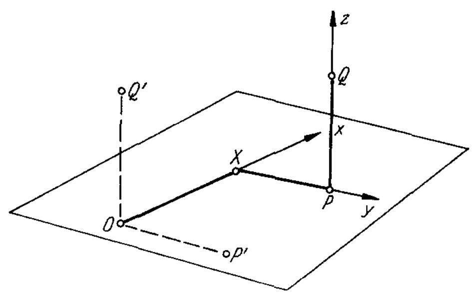


Fig. 51

What clearly hindered the rapid development of 3-space analytical methods was the perceptual difficulty of visualizing complex spatial structures, and an adequate analytical algorithm which could replace the psychological process of direct visualisation was not yet feasible. WALLIS' treatise on the cone-wedge ("cono-cuneus")[38] shows very well how far analytical techniques still depended on suitable preliminary geometrical reduction. In its most general form the cone-wedge was defined as the two sheeted surface which is the set of all lines as, as' constructed as follows: given two perpendicular diameters DD', EE' of the base-circle $DED'F'$ and equal line-lengths BD, $B'D'$ raised perpendicularly to the circle-plane, to point a, the meet of a second plane perpendicular to BB' with it,

[37] MACLAURIN takes his lead from a generalisation of NEWTON's organic description of conics, which virtually establishes a 1, 1 correspondence between the points of two conics (of which their intersection-points are invariants), one of which is conveniently assumed to degenerate into a line-pair. In his extension a 1, 1 correspondence is set up between the points of two n-degree curves, one of which degenerates into an (n — 1st)-degree curve and a line, or into an (n — 2nd)-degree curve and a line taken twice: on that basis he introduces an analytical treatment which allows him to make precise such ideas as nodes, double points and other now well-known defined concepts basic in the study of higher curves.

[38] First printed as *The shipwright's circular wedge* in appendix to his *Algebra* 1685, and republished as *conocuneus, seu corpus partim conum, partim cuneum repraesentans geometrice consideratum* ... in Latin translation in his *opera mathematica* 2 (1693): 681—704. As WALLIS outlines in introduction the work developed from the problem of sectioning designs for ships' hulls proposed in the early 1660's by Sir ROBERT MORAY and Sir WILLIAM PETTY.

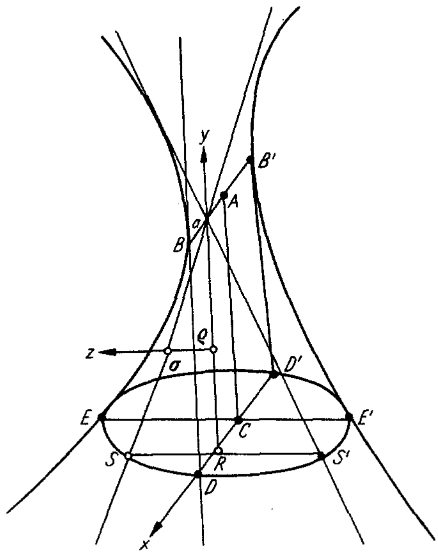


Fig. 52

correspond s, s' the meets of the base-circle with this perpendicular plane (so that ss' is normal to DD').*

Wallis' chief aim in his examination of the surface is to find plane-sections of it in a variety of ways, but in each case treated he neatly avoids using an equivalent of the analytical equation for the surface. Typically he considers a plane section through the vertex B which meets the base circle in a line srs' parallel to CE, and simplifies by noting that the curve of the meet of surface and cutting plane lies wholly in that plane Brs, and so can be given by a suitable "relatio" between abscissa $c\varrho = x$ and ordinate $\varrho\sigma = y$, where σ is a general point on the

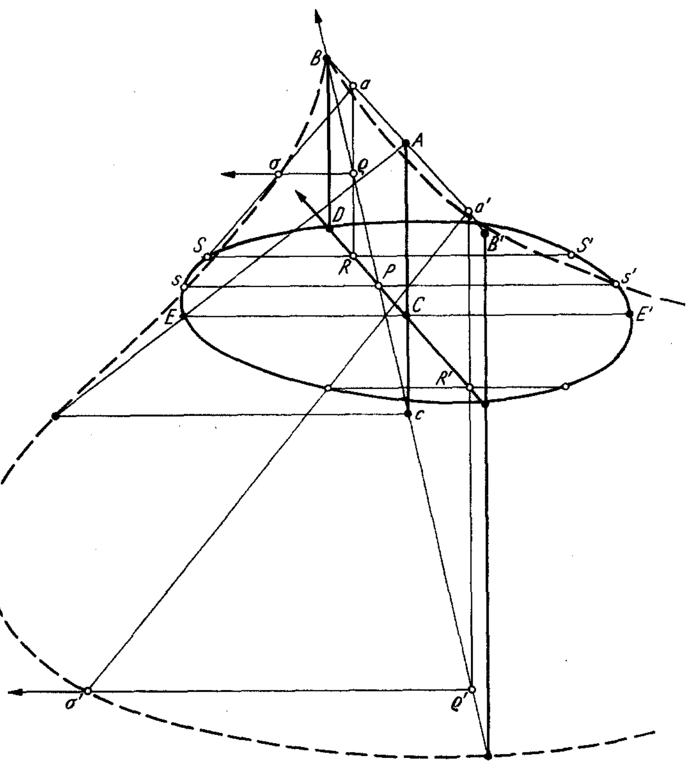


Fig. 53

* Analytically, where $CR = x$, $R = y$, $\varrho\sigma = z$, and using the proportion $\varrho\sigma : RS$ $(= (a^2 - x^2)^{\frac{1}{2}}) \doteq a\varrho\,(= b - y) : aR\,(= b)$, its representing equation can be taken as $b^2 z^2 = (b - y)^2 \cdot (a^2 - x^2)$, where the circle-radius is a, and $BD = B'D' = b$.

meet and $\varrho\sigma$ is taken parallel to rs. In fact, taking $AB=a$, $AC=b$, $Cc=c$ (which is sufficient to fix the plane Brs) and $BC:BA=1:\lambda\left(=(a^2+(b+c)^2)^{\frac{1}{2}}:a\right)$, we find that $Aa=\lambda x=CR$, and so $RS=\left(RD\times(DC+CR)\right)^{\frac{1}{2}}=(a^2-\lambda^2x^2)^{\frac{1}{2}}$, with $BA:Ba=(Bc:B\varrho=)\ Ac:a\varrho$, or $a\varrho=(b+c)\times\frac{a-\lambda x}{a}$; so that $a\varrho:aR\ (=AC)=\varrho\sigma:RS$ yields $\left(\frac{a-\lambda x}{a}\right)\times(b+c):b=y:(a^2-\lambda^2x^2)^{\frac{1}{2}}$, or the point-set of σ in the plane Brs is the tear-shaped single-looped quartic, $a^2b^2y^2=(a^2-\lambda^2x^2)(a-\lambda x)^2\times(b+c)^2$. Without deriving a similar representing equation it is extremely difficult to visualize the section of the surface by a general plane, but using analogously derived equations WALLIS is able to sketch a large number of particular sections for varying values of c, and for differently situated sections. The treatment carries over, too, in WALLIS' suggested extension in which similar conics through the points a, a, s' are to be substituted for straight lines, but it breaks down completely when the sections are no longer plane.*

A more general approach to plane sections of surfaces appeared a little later, in which there is a firmer grasp of the principle that parallel sections by their "motion" generate families of curves as their meets with the surface. This is, of course, obvious in the case of the two-sheeted cone where parallel sections cut off families of the same species of conic, but it is interesting to trace the approach in the case of the hyperboloid of revolution.

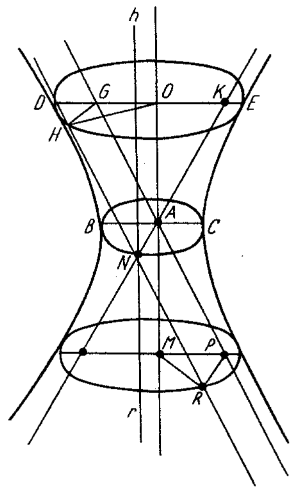


Fig. 54

WREN had defined the hyperboloid of revolution[40] by rotating the hyperbola DB round OAM, normal to the transverse axis BC through the centre A. Taking an asymptote GAP and any D on the hyperbola to be defined by $OD^2-OG^2=AB^2$, where DGO is drawn parallel to BA, we see that a plane section through the asymptote AG perpendicular to the hyperbola plane DBC meets the surface in a line (a "generator" of the surface).** Therefore, inverting the procedure, it is clear that a line HNR, inclined at some constant angle GAO to the perpendicular hNr to the circle-plane BNC, will by its rotation round the axis OA generate the hyperboloid of revolution.

* Since the section-curve is no longer definable in a CARTESIAN reference-framework by a "relatio" between two free variables, but now needs three.

** For, taking plane sections through DOE, BAC perpendicular to the hyperbola plane, these will be circles on DE, BC as diameters; and so $HG^2=HO^2(=DO^2)-GO^2=BA^2=AN^2$ or GH will be equal and parallel to AN for all lines DGO—that is, the point-set of H will be a line (and similarly for the perpendicular plane section through the second asymptote KA).

39 WALLIS describes it less accurately: "On a plain base which was ... a circle (like that of a ... Cone or Cylinder) stood an erect solid whose altitude (being arbitrary) was there double to the radius of that quadrant; and from every point of its perimeter straight lines drawn to the vertex met there not in a point (as is the apex of a cone) nor in a parallel quadrant (as in a ... cylinder) but in a straight line or sharp edge, like that of a wedge or cuneus".

40 In *PT* 4 (1669): 961—962: *generatio corporis cylindroidis hyperbolicis elaborandis lentibus hyperbolicis accommodati.*

A problem arises: what are the other plane sections of the surface, that is, those which are not through the hyperbola centre A normal to the circle-plane or are not parallel to the asymptotes GA, KA? To this NEWTON gave an answer in the early 1670's[41]. Let the surface be defined (see Fig. 53) by the rotation of the line XY round axis AB, where CD is the perpendicular distance between them and XY is inclined at a given angle $\widehat{LDF}$ to the plane DCB, and consider the section by some plane $QLKN$, inclined at given angle $\widehat{BHQ}$ to the axis AB: further, draw DF parallel to AB and LG, LF, LM (all of which will be coplane)

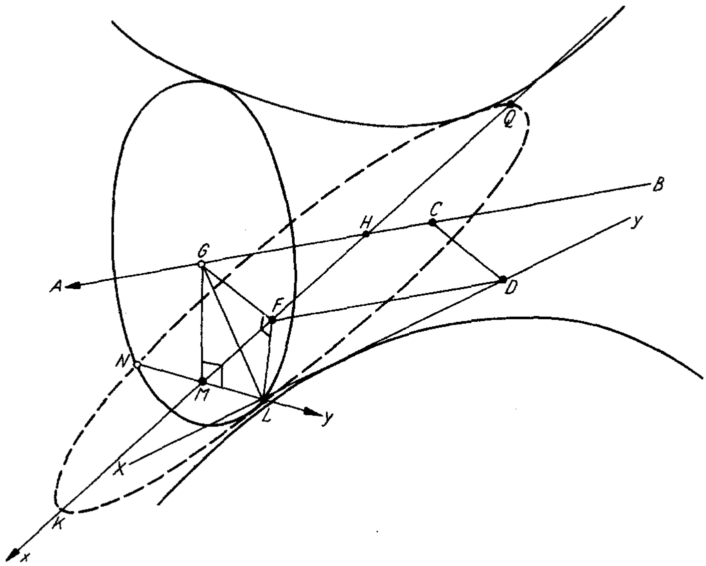


Fig. 55

perpendicular to AB, DF, HO respectively. NEWTON then needed only to consider the 2-space curve cut out by the section-plane by a preliminary geometrical treatment. Denoting $HM = x$, $ML = y$, and $CD = a$, $CH = b$, $MH{:}HG$ ($=\sec \widehat{GHM}$) $= d{:}e$ (constant) with $FD{:}FL$ ($=\tan \widehat{FLD}$) $= g{:}h$ (constant, then DF ($= CG = CH + GH$) $= b + \frac{e}{d}x$ and $FL = \frac{h}{g}\left(b + \frac{e}{d}x\right)$; so that

$$y^2 = ML^2 = GL^2 - MG^2 (= HM^2 - HG^2) = GF^2 (= CD^2) + FL^2 - (HM^2 - HG^2)$$
$$= a^2 + \frac{h}{g}\left(b + \frac{e}{d}x\right)^2 - x^2 + \left(\frac{e}{d}x\right)^2$$
$$= \frac{e^2(h^2+g^2) - d^2 g^2}{(dg)^2}x^2 + 2\frac{h^2 b e}{dg}x + \left(a^2 + \frac{h b^2}{g}\right),$$

which is a conic, and in particular $\left(\text{since } \frac{h^2+g^2}{g^2} - \frac{d^2}{e^2} = \frac{LD^2}{FD^2} - \frac{MH^2}{HG^2}\right)$ an ellipse, parabola or hyperbola according as $\frac{LD}{FD}$ is less than, equal to or greater than $\frac{MH}{HG}$ (or as angle $\widehat{LDF}$ is less than, equal to or greater than $\widehat{MHG}$).

[41] Printed in AU (1707): prop. 19: 141–142.

NEWTON keeps his length HC constant, but if we were to vary it we would, in effect, define the surface by a representing equation in x, y and b of the form $y^2=\lambda' x^2+2\mu' xb+\varrho' b^2+a^2$. This final step was taken by SLUSIUS and TOWNELEY* in a treatment[42] whose form clearly shows total independence of NEWTON. Adapting this (and a very ill-drawn figure in RIGAUD) we take (Fig. 54) the base hyperbola LO ($L'O'$) of centre I, transverse diameter OO' and asymptotes DI, dI, and rotate it round the axis IK perpendicular to OO', to form the surface. This we cut by some plane $N'NX$ at some constant angle NXO to OO', where $N'N$ is perpendicular to the hyperbola-plane $LOO'L'$. Then, drawing TI parallel to

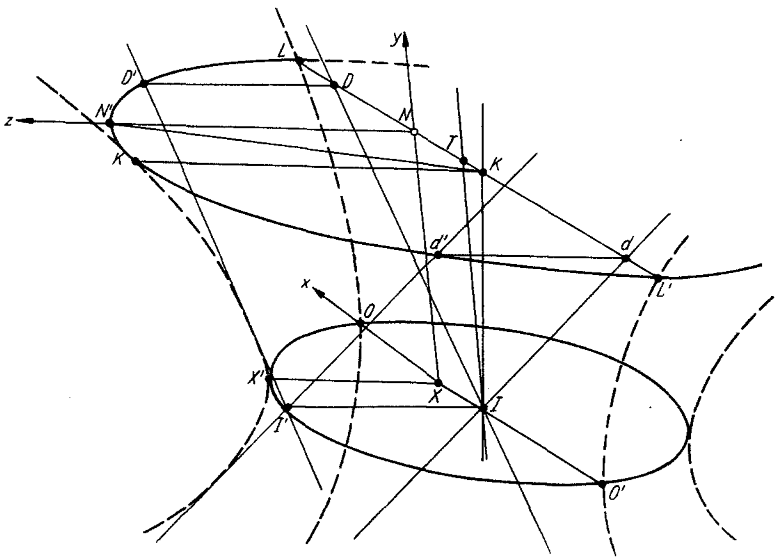


Fig. 56

NX and $LDNTKdL'$ parallel to OO' (meeting as shown) and denoting $IX=x$, $XN=y$, $NN'=z$ with $OI=b$ and $DK:TK:ID=\lambda:\mu:\nu$ (constant), we take the hyperbola to be defined by $LK^2-DK^2=OI^2$. Thus $DK=\frac{\lambda}{\nu}y$, $TK=\frac{\mu}{\nu}y$ and so $NK=XI+TK=x+\frac{\mu}{\nu}y$, and $LK^2=\left(\frac{\lambda}{\nu}y\right)^2+b^2=N'K^2$ (since $LN'L'$ is a semicircle) with $N'K^2=N'N^2+NK^2=z^2+\left(x+\frac{\mu}{\nu}y\right)^2$. Finally, equating we have $b^2+\left(\frac{\lambda}{\nu}y\right)^2=z^2+\left(x+\frac{\mu}{\nu}y\right)^2$ as the representing equation of the hyperbolic space, with the important corollary that any particular value of x in $[-\infty, +\infty]$ gives a plane-section of the surface as a 2^{nd}-degree "relatio" connecting

---

* A minor English geometer and a friend of JOHN KERSEY.

[42] See TOWNELEY's letter to COLLINS of 13 May 1672, RIGAUD (C). **1**: 190—195. TOWNELEY's treatment is apparently partly original, partly suggested by SLUSIUS. He writes (p. 191): "After M. de Sluse had proposed to me the solution of Dr. Wren's problem more generally ... he writ that the hyperbolical cylindroid might be so cut as to give all the sections both of cone and cylinder, and withal acquainted me with the property of an hyper. he had used to find them, and proposed to me the finding them, which I thus proceeded ...".

$XN=y$ and $NN'=z$, and so a conic—that is, for suitable parameter x, the equation represents a family of such conics. In particular, the family of conics whose plane is parallel to an asymptote is given by $\lambda=\pm\mu$, or is represented by $x^2+2\frac{\mu}{\nu}xy+z^2-b^2=0$, and the generator lines $I'D'$, $I'd'$ are the member which has $x=0$, or $z=\pm b$. For the family of plane sections parallel to the axis IK (a case added by SLUSIUS to TOWNELEY's results[43]) we have $\mu=0$ and so $b^2+\left(\frac{\lambda}{\nu}\right)^2 y^2=z^2+x^2$ (a family of hyperbolas): for the member through the hyperbola vertex O $x=\pm b$, or $z^2=\left(\frac{\lambda}{\nu}\right)^2 y^2$ (a second line-pair), while for other points X "this equation gives ... two constructions of an hyperbola" (for $x^2>b^2$ and $x^2<b^2$).

All this could, of course, be derived from the NEWTON example by letting $CH=b$ and the angle $\widehat{GNM}$ (or, equivalently, the ratio $MH:HG=d:e$) vary, but the extension made by SLUSIUS and TOWNELEY is a major conceptual advance. In effect, their approach sketches in the principle of continuity for conics defined on the hyperboloid of revolution, just as KEPLER had outlined it for plane sections of a two-sheeted cone half a century earlier[44]. Specifically, both show that by continuously varying the position of the section-plane the conic-meets also vary continuously, and in this general treatment line-pairs, circles and parabolas appear clearly as degenerate and limiting cases of the general conic (and not as specific curves in their own right). In general, any plane section of a quadric surface is a conic and a similar "KEPLER" law of continuity holds, but it is interesting to note how slowly it was to be realized that the conic is more general than its classical definition as the plane cut of a cone. Once again, apparently, we have an example of a case where an overelaborate Greek treatment became a block to further progress—significantly, no major analysis of the general quadric surface was made till MONGE.[45]

In all this rich confusion of developing procedures one basic aspect tended to be lost sight of: the idea of defining a point-set as a point-correspondence—or rather perhaps the CARTESIAN approach which set up the general point-set by its correspondence with two ordered line-lengths tended, by its successful and fruitful elaboration, to obscure less developed correspondences. NEWTON, as always, is the proving exception. In his later undergraduate years he had toyed with a bipolar coordinate reference-system (on the model of the central conics defined by $x\pm y=\lambda$, where x, y are the distances from a general point on the curve to the foci)[46], and later in life he came to consider more general

---

[43] Compare *op. cit.*: 194.

[44] In his *ad Vitellionem paralipomena quibus pars optica traditur*, Frankfurt, 1604: ch. 4: *de coni sectionibus*.

[45] A. LALOVÈRE had, however, in his *quadratura circuli et hyperbolae segmentorum ex dato eorum centro gravitatis*, Toulouse, 1651: Book 5, defined the hyperboloid of one sheet geometrically by plane (conic) sections.

[46] As we see from the *Waste Book*, *CUL Add.* 4004: 1V (miscellaneous calculations dated 1664, September) and 50Vff. (more systematic treatments of 1665 and 1666). General tangent treatments of curves defined by bipolar analytical coordinates are given in *CUL Add.* 3960: section 14 (to be dated around 1670—1672), which is HORSLEY's *geometria analytica* and (in English) COLSON's *Method of fluxions and infinite series*, London, 1736.

aspects of point-correspondence, slowly unloosing himself from the CARTESIAN idea of having a single fix-point as origin in a coordinate-system.

Already in the 1670's, as we have seen[47] NEWTON had used 1, 1 point-correspondences to define conic point-sets—in particular, his "organic construction" virtually defines a 1, 1 correspondence between the points of two conics, one of which is a degenerate line-pair. Problem 53 of his *AU*, probably dating from the same period[48], elaborates a corresponding analytical treatment (after some preliminary geometrical simplification). Specifically, where two fix-poles A, B round which rotate given angles CAD, CBD are such that the meet of AD and BD is on the fix-line EF, he wishes to examine the corresponding meet of the

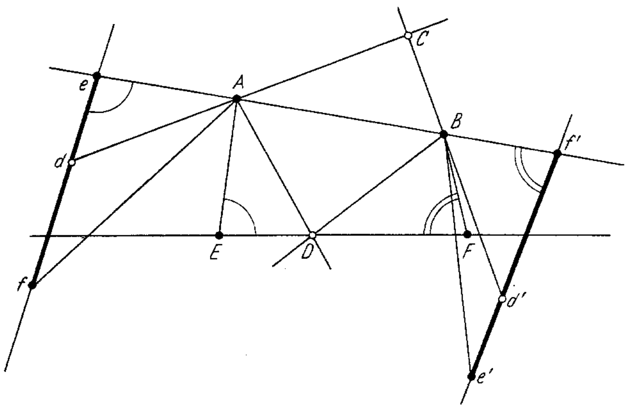


Fig. 57

other two arms AC, BC, giving first a neat geometrical reduction which virtually straightens out the angles $\widehat{CAD}$, $\widehat{CBD}$: as the line EF rotates round A through the angle $\widehat{dAD}=\pi-\widehat{CAD}$ into the line ef, let point E of EF pass into e, the meet of ef and AB; similarly, as EF rotates round B through the angle $\pi-\widehat{CBD}$ into the line $e'f'$ let point F of EF pass into f', the meet of $e'f'$ and AB; and finally, let F pass into f on ef, and E into e' on $e'f'$. Then, clearly, ef $(=EF)=e'f'$, and to any point D in EF there correspond points d, d' in ef, $e'f'$ such that $ed:df$ $(=ED:DF)=e'd':d'f'$ (or $ed=e'd'$, $df=d'f'$), further, the point-set C is given as the set of meets of the lines dA and $d'B$. Reformulating we can state the equivalent problem: given two equal line-lengths ef, $e'f'$ and the point-correspondence defined between them such that, where d is in ef, d' in $e'f'$, $ed:df=e'd':d'f'$ (or $ed=e'd'$, $df=d'f'$), what is the point-set of C, the meet of the lines dA, $d'B$, where A, B are two fix-points in ef'? In answer NEWTON introduces a CARTESIAN coordinate-system, taking CH, CK (through the locus-point C) parallel respectively to ef, $e'f'$ (and meeting AB in H, K) and denoting $BK=x$, $KC=y$ with $AB=m$, $Ae=a$, $Bf'=c$, $ef=e'f'=b$ and $CK:CH:HK=d:e+f$ (a constant ratio since the triangle CHK is given in species): then $BK:KC=Bf':f'd'$, or $f'd'$ $(=fd)=c\,\frac{y}{x}$,

[47] See previous chapter.

[48] It exists in his Lucasian lectures of the 1670's, see *CUL Dd.* 9.68 $\cdot\equiv\cdot$ *AU* (1707): 207—209.

$=ef-ed$, or $ed=b-c\frac{y}{x}$; again $CH=\frac{e}{d}y$ and $HK=\frac{f}{d}y$, or $AH=AK-HK=(m-x)-\frac{f}{d}y$; and finally the proportion $AH:HC=Ae:ed$ yields

$$\left(m-x-\frac{f}{d}y\right):\frac{e}{d}y=a:\left(b-c\frac{y}{x}\right),$$

or

$$bdx^2+(ae+bf-cd)\,xy-cfy^2-bdmx+cdmy=0,$$

a conic. Thus NEWTON has shown that the restricted homographic correspondence of CAd, CBd' defines a conic point-set.[49]

In manuscript papers dating from around 1680[50] NEWTON took the further radical step of modifying the CARTESIAN reference-scheme by separating the coordinate line-lengths on which his analytical theory of point-correspondence is to be defined.

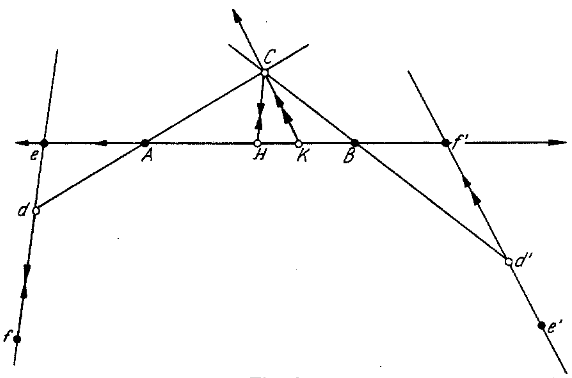


Fig. 58

Consider for example, what is the simplest general case where the 1, 1 correspondence is to be made between the points of two lines.[51] (NEWTON'S development is given, a little artificially, apparently in the reverse order of his original sequence of ideas*, but for the moment we will follow his exposition.) Where AC and BD are fix-lines on which are located respectively the fix-points A and B, while a third fix-point E is given in general position in the plane, NEWTON considers the point-correspondence set up in the two lines by their meet with a general line through E. Specifically, let the line ECD set up the correspondence $C\leftrightarrow D$ (where C is in AC, D in BD), and consider the two directed line-lengths $AC=x$, $ED=y$: what is the "relatio" which connects them? In a preliminary investigation NEWTON clarifies the conditions which must hold in the correspondence. Clearly, since two lines meet in a unique point, any point C defines a unique line BC through E, and therefore a unique point D in BD, and this must be incorporated in the "relatio"—that is, each value of x in the relatio must yield a unique value of y, and conversely, or the most general form of "relatio" must, for x constant, yield a linear equation in y, and conversely.

---

* In the tradition, in fact, of the classical Greek synthetic proof.

[49] MACLAURIN in his (1720) *geometria organica* (which generalizes the organic construction) gave an equivalent analytical treatment of the more general 1, 1 correspondence set up between two n-degree curves, one of which is allowed to degenerate suitably.

[50] The manuscript *de inventione porismatum* (*CUL Add.* 3963: 159—160)—with several slightly variant minor drafts—sketches verbally the concept of "porism", including in that concept several particular types of correspondence and showing, in particular, how knowledge of suitable corresponding points allows suitable restrictions to be put on the correspondence, listing several examples (without proof). An analytical basis, however, is given in the later propositions of the manuscript *de compositione locorum solidorum* (*Add.* 3963: 126—149, especially 132Lff.).

[51] *Add.* 3963: prop. 11: 132R.

NEWTON now sets this up more precisely on his geometrical model. Draw EH through E parallel to BD (meeting AC in H) and AL perpendicular to AC (meeting EH in L): then, taking $AL=c$, $EB=b$, $AF=a/b$ and $AL:AH=d:e$ (constant) we have the proportions $CH:CK=HL:LA$ and $HL:AH=FC:EK$, or $CK:EK=HL\times CK$ $(=AL\times CH):HL\times EK$ $(=AH\times FC)$, $=EB:BD$; so that, substituting the analytical measures of line-length, $(EB:BD=)\ b:y=\left(\frac{AL}{AH}(AH+HC):FC=\right)\left(c+\frac{d}{e}x\right):\left(\frac{a}{b}-x\right)$, and finally $a=bx+cy+\frac{d}{e}xy$, the most general form which a 1, 1 correspondence between x and y can take. Clearly, the contrived nature of the procedure and the simple (and elegant) form of the result show that, in fact, a "relatio" of this form was hypothesized and then the given values of the given line-lengths calculated,* and indeed NEWTON takes care to assert it.[52]

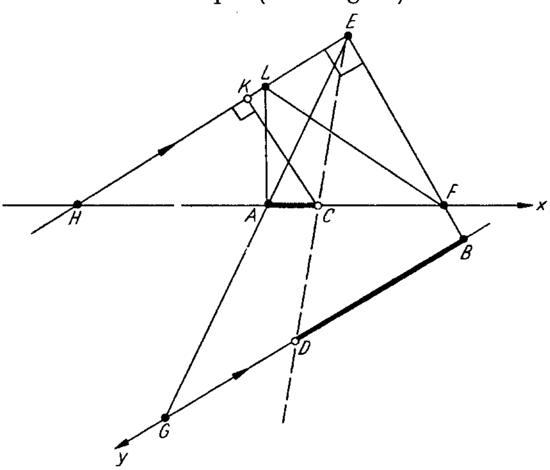


Fig. 59

What NEWTON has shown is that any 1, 1 correspondence between two variables x and y must take the general form $\alpha xy+\beta x+\gamma y+\delta=0$, where $\alpha, \beta, \gamma, \delta$ are constants to be chosen to fit. In short, NEWTON has the analytical basis on which to raise an analytical theory of cross-ratio, involution and homographic transform exactly as CHASLES was to do later,** and it is very tempting—though there seems no explicit attempt so to do in the

* Thus, assuming x and y connected by a "relatio" of the form $a=\lambda x+\mu y+\nu xy$, we have when $y=0$, $x=AF=\frac{a}{b}$, $=\frac{a}{\lambda}$ (or $\lambda=b$); when $x=0$, $y=BG=EB\times\frac{AF}{AC}=\frac{a}{c}$, $=\frac{a}{\mu}$ (or $\mu=c$); and when $y=\infty$, $x=AH=-\frac{e}{d}c$, $=-\frac{c}{\nu}\left(\text{or } \nu=\frac{d}{e}\right)$.

** Cross-ratio invariance on a line-pencil, for example, follows immediately by seeing the pencil as setting up a 1, 1 correspondence between the points x_i, y_i of any two transversals, or, considering four pairs of corresponding points $x_i \leftrightarrow y_i$, $i=1, 2, 3, 4$, the syzygy-set $(\alpha x_i+\beta y_i+\gamma x_i y_i+\delta=0)$ yields, on elimination of the constants $\alpha, \beta, \gamma, \delta$, the cross-ratio equality, $\frac{(x_1-x_3)(x_2-x_4)}{(x_1-x_4)(x_2-x_3)}=\frac{(y_1-y_3)(y_2-y_4)}{(y_1-y_4)(y_2-y_3)}$.

[52] Compare *Add.* 3963: 132R: "assumatur plenissima quaevis relatio quantitatum quae ad invicem per simplicem geometriam determinabiles sunt, qualis est haec $a=bx+cy+\frac{d}{e}xy$, ubi $a, b, c\ldots$ denotant quantitates datas cum signis suis + et— affectas, et x et y quantitates incertas ex quarum alterutra cognita supponitur posse determinari per simplicem geometriam". (He has defined a "simply geometrical" procedure at 131R: "per geometriam simplicem determinabiles esse intelligo quae per ductum ... linearum sine adminiculo circuli vel anguli dati—hoc est per additionem, subductionen et inventionem quartae proportionalis, vel, ut jam loquuntur geometrae, per multiplicationem et divisionem sine extractione radicis—determinari possent". In other words, "Simple geometry" is the geometrical equivalent of an "analytical" sequence of operations in the restricted Cartesian sense.)

manuscripts—to assume that NEWTON used such an analytical basis in deriving the thoughts on general m, n correspondences given in *de inventione porismatum*.

To point this, let us consider[53] his theorem 2 which states that if, in the $(1,1)$ correspondence $x \leftrightarrow y$, $\infty \leftrightarrow Y$ and $X \leftrightarrow \infty$ then $(x-X)\times(y-Y)$ is constant—a

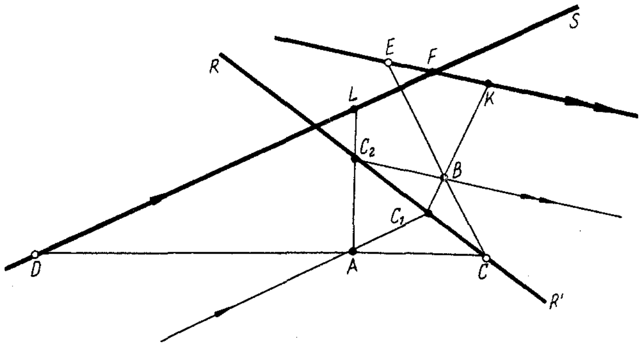


Fig. 60

result which is an immediate corollary of the general 1, 1 form above (since the conditions give respectively $\alpha Y+\beta=0$, or $\beta=-\alpha Y$, and $\alpha X+\gamma=0$, or $\gamma=-\alpha X$, so that $0=\alpha xy-\alpha Yx-\alpha Xy+\delta$, or $(x-X)(y-Y)=\frac{\alpha XY-\delta}{\alpha}$). The theorem is easily and neatly applied to a wide range of point-correspondences and NEWTON gives[54] a choice few. Thus[55] his porism 6 is a simple line-model: given three lines RR', FS, FT and two fix-points A, B, let any point C on RR' define corresponding points D in FS and E in FT, where D, E are the meets of AC with FS, BC with FT respectively. In this correspondence $D \leftrightarrow E$, let $\infty_{FS} \leftrightarrow K$, $L \leftrightarrow \infty_{FT}$, then by the theorem $DL \times EK$ is constant. More interestingly, a second application[56] of the theorem yields a point-correspondence proved by NEWTON in his *PM* on classical lines: where the lines AG, AH are the tangent-pair from a fix-point A in general position to a given conic, consider the 1, 1 correspondence set up in them between their meets B, D with RR', tangent to the conic. Clearly, where $B \leftrightarrow D$, $E \leftrightarrow \infty_{AG}$ and $\infty_{AH} \leftrightarrow F$, or $EB \times FD$ is constant $(=EH \times FA, =EA \times FG)$.

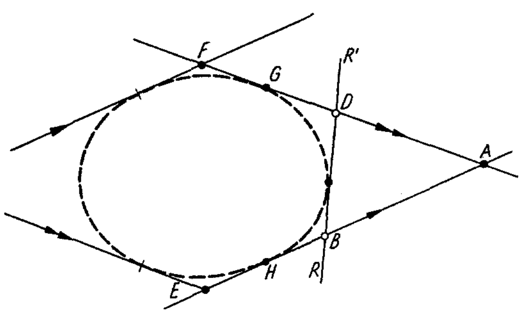


Fig. 61

The similar consideration of general m, n, correspondences becomes rapidly unwieldy, especially in a verbal exposition. NEWTON, in fact, goes on to consider some 2, 1 and 2, 2 correspondences set up by conic-tangents in fix-lines in the plane, but though the applications he makes in the *de inventione* are clearly

[53] *Add.* 3963: 159R.
[54] *op. cit.* 159Rff.
[55] *op. cit.* 160R.
[56] *op. cit.* porism 2: 159V; and compare previous chapter, note [27].

the product of hard thought they remain little more than a sketch, unimplemented by a systematic exposition. But the whole subject of m, n correspondence was a mighty thing for a man no longer young to handle and it is to his credit that he did so much, and a comment on those who later looked over the manuscripts that they failed to penetrate his ideas. It was, however, an unfortunate result that the manuscripts ultimately passed into oblivion while the ideas he had introduced had to be painfully rediscovered much later.

In general summary of the later 17th century attitude to analytical techniques in geometry, it is important to stress how little general method there was. Each proof depended fundamentally to a greater or less degree on a preliminary geometrical reduction to a form where existing techniques could be applied. Indeed it seems true to say that CARTESIAN analysis, while accepted as a useful form of proof, was looked upon as essentially eliminable by the substitution of an exactly corresponding synthetic form. NEWTON'S appendix to his *arithmetica universalis*[57] —an eternal worry to those historians who have tried to read 19th century attitudes into 17th century mathematics—essentially summarises a prevailing attitude: "Equations are expressions of arithmetical computation and properly have no place in geometry except in so far as truly geometrical quantities (that is, lines, surfaces, solids, and proportions) are thereby shown equal, some to others. Multiplications, divisions and computations of that kind have been recently introduced into geometry, unadvisedly and against the first principle of this science .... Therefore these two sciences ought not to be confounded, and recent generations by confounding them have lost that simplicity in which all geometrical elegance consists."

A thin framework for future development had been more or less tentatively and unsystematically established, but a very great deal remained to be done before any fully analytical treatment of geometrical concepts was possible. In historical fact, the process took another century of effort, and it would be fairer to cite EULER and MONGE as creators of our modern form of analytical theory—if, that is, there were any real point to making the claim at all.

VIII. Calculus

1. *Indivisibles and the arithmetick of infinites*

More so than any other branch of mathematics, the differential and integral calculus has been seen as the triumph of 17th century exact thought and, indeed, as one of its most attractive facets. A long historiographical tradition[1] has sketched the immense amount of work—developed largely in the geometrical models of curve-tangent and curvature, and of area, surface, curve-length and volume,

[57] *AU*.(1707): 282.

[1] C. B. BOYER in his *Concepts of the calculus* (New York, $_1$1939, $_2$1947, $_3$1959) includes a massive bibliography which, in regard to secondary works, is fairly complete up to about 1940. The not inconsiderable amount of work published since is to be most conveniently found listed in the monthly abstract, Mathematical Reviews. BOYER'S work itself is typical of the dangers inherent in a set attitude to the subject: approaching his subject with an ideal of rigour which seems to be that of the early 19th century formulations, he tends to make earlier investigations stand or fall by that criterion, and in particular misses much of the rich significance of geometrical treatments, widespread throughout the 17th century.

rather than as an abstract theory of derivative and integral—which preceded the more sophisticated treatments of later centuries. One might wonder how anything could now be said which has not been said many times before. Unfortunately, in the past too little has been said at too great length and too glibly. The source material available has been little studied*, and a great deal, both printed and in manuscript, which deserves to be better known still lies dusty and undisturbed. It is in a deliberate effort to bring to light some of these unexplored but richly significant calculus procedures that the artificial division of the remaining four chapters is made, though I hope at the same time to portray a range of thought typical in a real sense of the 17th century achievement in arithmetising the infinite.

In this first chapter, in particular, some account will be given of the CAVALIERI-TORRICELLI theory of indivisibles, and its English offshoots—an aspect whose complexity has been little appreciated.[2]

We can perhaps, tentatively, isolate three formative aspects which coalesced into the rambling, loosely connected set of concepts treated by CAVALIERI in his treatises on indivisible methods.[3] Of these the most immediately obvious influence is that of the numerical techniques for measuring area, volume and surface—we may name them "gauging" methods—a collection of often rough and ready approximative formulisations which yet had within them the germs of ideas basic to the concept of integration. Such methods, of course, date back to beyond recorded history in their simplest examples (among the Egyptians and the Babylonians) and, though have only a few extant arithmetical texts, such as HERON'S *metrica*, from which to argue, they must have been a large part of Greek practical mathematics. By the early 17th century these techniques had reached a certain level of refinement in the hands of such men as STEVIN[4] but especially KEPLER who in his *nova stereometria*[5] made general application of the gauging method of approximating to areas and volumes by suitably drawn sections. Thus, where we need to approximate to the area of the figure shown which is cut off between two parallel sections AB, $A'B'$, we split the area into

* Where, to name but a few of the more important figures, are the authoritative evaluations of the work of PIETRO MENGOLI, ANTONIUS LALOVERA, JOHN WALLIS, NEWTON?

[2] Though C. B. BOYER in *Cavalieri, limits and discarded infinitesimals*, Scripta mathematica **8** (1941): 79—91, has emphasised several errors in the conventional account—notably that CAVALIERI'S procedures for the most part (and exclusively in the early work) compare the limit of two "indivisible" sequences rather than calculate numerically a single limit-aggregate. Indeed, CAVALIERI'S thought in detail is unbelievably rich—he had read widely in ARCHIMEDES, STEVIN, KEPLER and others (and had absorbed the medieval theory of latitude of forms, especially the geometrical aspects developed by ORESME), and his ideas are an amalgam of what he had read and of the thoughts that reading inspired.

[3] Specifically, *geometria indivisibilibus continuorum nova quadam ratione promota*, Bologna, 1635 (which is fundamental); and *exercitationes geometricae sex*, Bologna, 1647. Compare, too, BOYER (*op. cit.*, note 1): 117—123.

[4] See H. BOSMANS: *Le calcul infinitésimal chez Simon Stévin*, Mathesis **37** (1923): 12—18, 55—62, 105—109; and *Sur quelques exemples de la méthode des limites chez Simon Stévin*, Annales de la Soc. sc. de Bruxelles **37** (1913): 171—199.

[5] *Nova stereometria doliorum vinoriorum*, Linz 1615.

sections by further parallels $A_i B_i$, $i=1, 2, 3, \ldots, n$, and then consider either of the summations,

$$\text{area } AA'B'B \approx \begin{cases} AB \times AA_1 + A_1B_1 \times A_1A_2 + \cdots + A_n B_n \times A_n A' \\ A_1B_1 \times AA_1 + A_2B_2 \times A_1A_2 + \cdots + A'B' \times A_n A'. \end{cases}$$

In particular, where we take the section-parallels at equal distances, we have what is usually thought of as the typical form of indivisible-process—merely by increasing the number of section-parallels indefinitely we trace a general sequence whose limit as the number of parallels becomes indefinitely large yields the required area (to any degree of approximation, at least).* But we do well to notice that what is important in this extension—the introduction of a limit-consideration—is a theoretical advance on (and a redefinition of) the practical gauger's idea, which remains a mere numerical approximating technique. Indeed, in the whole of 17[th] century mathematics there seems only one example where a practical approximation gave rise to a serious mathematical investigation—that of the integral $\int_\beta^\alpha \sec x \cdot dx$—and there circumstances were quite exceptional.[6]

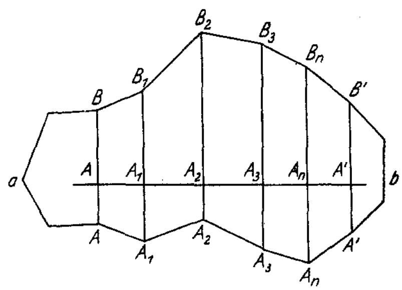


Fig. 62

Analogous concepts had existed in Greek mathematics, but their significance was disguised and distorted by the forbidding logical form in which they were stated, the exhaustion-method.[7] No contemporary mathematical work examined the nature of this method of proof, and its rigour, while accepted at a mathematical level, appeared artificial and over-precise. However, several standard results, proved rigorously by an exhaustion-proof, became the basis of many of CAVALIERI's indivisible-comparison theorems (and the exhaustion-proof was accepted as their ultimate theoretical justification); while later TORRICELLI in

---

* In fact, as we shall see, this explicit process is not to be found in CAVALIERI's *geometria indivisibilibus* ..., but is a simplification introduced in the 1630's by several mathematicians including FERMAT and ROBERVAL.

[6] The integral appears in the construction of the MERCATOR map, and for a century after the projection was introduced was tabulated by the inequalities,

$$\sum_{0 \leq n \leq N-1} [\sec(n \,\Delta\vartheta) \cdot \Delta\vartheta] < \int_0^\vartheta \sec x \cdot dx < \sum_{1 \leq n \leq N} [\sec(n\,\Delta\vartheta) \cdot \Delta\vartheta] \qquad (\text{where } \Delta = 1/N),$$

which can be made as narrow as we wish by decreasing the tabulation-interval $\Delta\vartheta$ (since the difference of the two bounds is $\sec\vartheta \cdot \Delta\vartheta$, which can be made as small as we wish by decreasing Δ). Such a table, calculated at 1′ intervals, $x \in [0°, 45°]$ had been given by EDWARD WRIGHT in 1599, and it was by comparing this table with a table of logarithmic tangents that HENRY BOND in the 1640's made the hypothesis that the integral is some log tan function—proved formally by JAMES GREGORY in 1669. It remains a historical curiosity that a table of "logarithms" should exist before NAPIER or BÜRGI published their canons (see F. CAJORI: *On an integration antedating the integral calculus*, Bibliotheca mathematica $_3$**14** (1913—1914): 312—318).

[7] See chapter 9.

his printed works and manuscripts[8] examined more closely the interconnection of the "exact" (exhaustion-method) and indivisible proofs.

Above all, however, close examination of CAVALIERI's indivisible theories shows the unmistakable influence of medieval ARISTOTELIAN treatments of such limit-concepts as instantaneous speed and continuous variation. Most obviously, in many places he takes over much of the scholastic terminology of the "calculators" in developing his own ideas on continuity and on continuously varying quantities[9], but more deeply he gives closely argued verbal justification of his indivisible theories in the medieval manner despised (and so ignored) by 19th century historians. Only by stripping away this verbal justification are we left with the travesty of his theory which is put forth by many historians. Rather CAVALIERI's treatment has, implicitly, many clarifications of the underlying concepts on which previous analysis of the infinite had been based, and in defining them both strengthened them and facilitated their use.

Mathematically CAVALIERI develops two major new concepts, that of powers of line-elements and that of coordinate directions (which are used to derive theorems which compare powers of variously defined line-elements). The departure-point for introducing the former is the concept of similarity and of being similarly situated: two figures (in two or three dimensions in those considerations developed with strict reference to a geometrical model, but generally in n-dimensions in the more analytical theory later given) are defined to be similar if to any point in the one corresponds a unique point in the other such that the distance between two points of one figure bears a constant ratio to the distance between the two corresponding points in the other.[10] On that basis he sets up the concept of power of a line-(area-, volume-)element.

Consider, for example, the two similar square pyramids $O{:}PQRS$, $o{:}pqrs$, and set up corresponding (square sections) parallel to the respective bases $ABCD$, $abcd$. CAVALIERI visualises these similarly-situated cross-sections as generating the respective solids, arguing that in some valid sense the solids are made up of the limit-sums of these cross-sections when the distance between two adjacent cross-sections becomes indefinitely small. In his mathematical treatment he is not concerned with the theoretical difficulties inherent in such a limit-procedure (never making it explicit, for example, whether he sees the limit-process as being actual or potential in the Aristotelian sense), but treats it only as an "artificium" which works and for which, presumably, a theoretical justification is possible—as such its nature need not be clarified, and in particular the question whether an indivisible had thickness in the limit could be left

[8] See his *opere* (ed. G. LORIA & G. VASSURA). Faenza, 1919; *passim*; but especially the *de dimensione parabolae* (included in his *opera geometrica*, Florence, 1644), where he contrasts numerous proofs of the same result (the quadrature of a parabola segment), clearly being more interested in the method used than in what it derived.

[9] Compare especially Book 5 of *exercitationes geometricae sex*: 321—422: *in qua de uniformiter difformiter gravibus per indivisibilia instituitur contemplatio*, where he derives a concept of indivisibles of weighted elements in which the weighting function is expressed in "gradus gravitatis" and defined by a latitude of forms variation pattern.

[10] Compare *geometria indivisibilibus* ...: Book 1: 11 ff.

undecided[11]. In some sense, then, we have $\sum(\text{area } ABCD) = \text{pyramid } (O{:}\,PQRS)$, and $\sum(\text{area } abcd) = \text{pyramid } (O{:}\,pqrs)$, and this he sets up as the proportion

$$\begin{aligned}\textstyle\sum(\text{area } ABCD) : \sum(\text{area } abcd) &= \text{pyramid } (O{:}\,PQRS) : \text{pyramid } (o{:}\,pqrs)\\ &= QR^3 : qr^3 \quad \text{(by their similarity)}.\end{aligned}$$

Now project each element parallel to the bases onto the triangular faces OQR, oqr: then, since the sections are similarly situated in similar solids, area $ABCD$: area $abcd = BC^2 : bc^2$, so that $\sum(BC^2) : \sum(bc^2) = (QR)^3 : (qr)^3$, where the limit-summation is made over corresponding lines BC, bc in the similar triangles

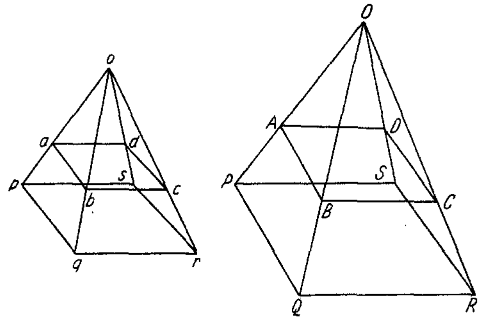


Fig. 63

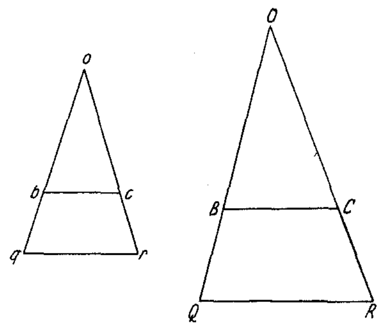


Fig. 64

OQR, oqr. In a similar way $\sum(BC) = \sum(bc) = \text{area } OQR : \text{area } oqr = (QR)^2 : (qr)^2$. By analogy the general pattern is suggested that

$$\textstyle\sum(BC^n) : \sum(bc^n) = (QR)^{n+1} : (qr)^{n+1},$$

and apparently it was in setting up a recursive way of verifying this for integral powers of n that CAVALIERI first introduced the concept of coordinate-direction, though the concept is given a general treatment later in *geometria*[12] independently of the particular application made of it in Book 1.

Where OX, OY are two (non-parallel) fix-lines given in direction, consider the area-segments $ABCcba$, $EFGgfe$ cut off from two given areas by AE, CG parallel to OY. Taking a third parallel BF to OY (which is CAVALIERI's "regula") cutting these segments in Bb, Ff respectively, we can denominate the general parallel BF by its outpoint x with OX* (and in particular AE, CG are X_1, X_2 respectively): then, viewing the areas as the limit-sum of the segment-lengths

* This is stated only verbally by CAVALIERI without any free variable denomination of the general parallel BF, but I introduce this adaptation to clarify his treatment.

[11] Though in Book 3 of his *exercitationes geometriae sex: in qua discutiuntur ea quae a Paulo Guldino .... in ejusdem rentrobaryca praefatae geometriae indivisibilium objicientur* he says that, if we wish, we may substitue for the indivisibles small elements of area, volume, as ARCHIMEDES had done, and gives (pp. 240–241) the analogy of the parallel threads in a piece of cloth which fill up the whole area of the weave, or again that of the parallel pages in a book which fill up its thickness. Elsewhere he uses the NEWTONian idea that the element generates the whole by a parallel motion, in which scheme his indivisibles are limit-motions.

[12] In *geometria* ...: Book 7 and *exercitationes* ...: Book 1.

cut off from the general parallel BF (where the distance between two adjacent segments becomes indefinitely small), we can introduce the following symbolism to clarify CAVALIERI's verbal treatment*:

$$\text{area } ABCcba = \sum_{X_1 \leqq x \leqq X_2} [\alpha(x)], \quad \text{area } EFGgfe = \sum_{X_1 \leqq x \leqq X_2} [\beta(x)],$$

where $\alpha(x)$, $\beta(x)$ are the respective lengths of the line-segments cut out of the respective areas by the general x-parallel BF. (There is an immediate extension to the model where volumes are cut by parallel planes, each parallel to a plane "regula".)

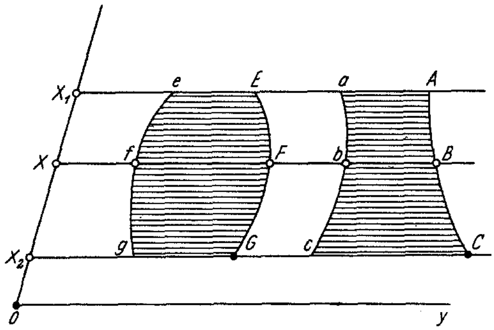


Fig. 65

On these two bases, the concept of similarity and the concept of a coparallel set of cutting lines (planes), CAVALIERI develops general "indivisible" techniques. Specifically he isolates two complementary approaches, his collective theory of indivisibles[13] and his distributive theory.[14]

In illustration we can compare the areas $ABCcba$, $EFGgfe$ above in two conceptually distinct ways: "collectivè, hoc est comparando aggregatum ad aggregatum", that is, by straightforwardly finding each of $\sum_{X_1 \leqq x \leqq X_2} [\alpha(x)]$, $\sum_{X_1 \leqq x \leqq X_2} [\beta(x)]$ separately and then comparing their proportion; and "distributivè, sc. comparando singillatim quamlibet rectam figurae ABC $[\alpha(x)]$ ... cuilibet rectae figurae EFG $[\beta(x)]$ ... in direction [on the x-parallel, that is] existenti", that is, we derive the proportion $\alpha(x):\beta(x)$ for each position of the x-parallel, and then (presumably using an averaging technique in the general case, though CAVALIERI considers only the case where their ratio is constant) to derive $\sum_{X_1 \leqq x \leqq X_2} [\alpha(x)]$, $\sum_{X_1 \leqq x \leqq X_2} [\beta(x)]$. Where, for all x, $\alpha(x):\beta(x)$ is a constant ratio $\lambda:\mu$ we have $\sum_{X_1 \leqq x \leqq X_2} [\alpha(x)] : \sum_{X_1 \leqq x \leqq X_2} [\beta(x)] = \lambda:\mu$, which is "CAVALIERI's" Theorem—an approach developed later, especially by GREGORY ST. VINCENT[15] but also by WALLIS, JAMES GREGORY, BARROW and other exponents of geometrical integration techniques, into a general method of geometrical transformation, the "ductus plani in planum". Thus, where a general plane x (moving parallel to some regula-plane) cuts off rectangles $ABCD$, $abcd$ from two solids such that always $AB \times BC : ab \times bc = \lambda:\mu$ = area $ABCD$: area $abcd$, then the respective volume-segments cut off between two particular planes X_1, X_2 are also in the ratio $\lambda:\mu$. GREGORY ST. VINCENT (and others after him) sees this as a transform,

* CAVALIERI uses the unwieldy verbal concept of "omnes lineae (omnia plana) ... juxta regulam (OY) assumptae (assumpta) ...".

[13] Very roughly this is developed in *geometria* ...: Books 1—6, with additions in *exercitationes* ...: Books 2ff.

[14] Given a detailed treatment in *geometria* ...: Book 7, *exercitationes*: Book 1.

[15] See his *opus geometricum*, Antwerp, 1647: Book 7: 703—864: *de ductu plani in planum*. WALLIS translated the transform into equivalent analytical form in his *AI* (1656): 60ff., which is equivalent to defining an integral transform.

defined by $AB \times BC : ab \times bc = \lambda : \mu$ of one volume (of element $ABCD$) into a second volume (of element $abcd$) which multiplies the measure of the volume by μ/λ (and in particular, when $AB \times CD = ab \times cd$, preserves the measure of the volume); and it very quickly became an elegant method of reducing geometrical problems of volume-measure to a more easily workable form.[16]

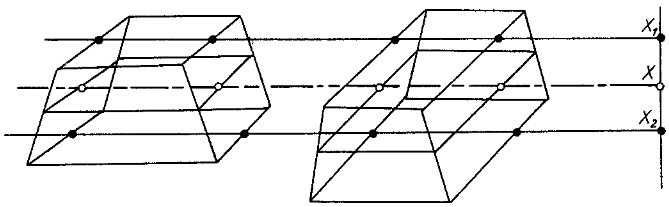


Fig. 66

CAVALIERI himself was content to sketch in a few elegant examples of its use[17] but in contrast developed the collective approach in minute detail.[18] To illustrate his general approach, consider[19] the parallelogram $ACGE$ in which AC, EG are parallel to the regula OY, BF bisects AC, EG, DH bisects AE, CG, and CE is a diagonal. Denote the parallels AC, DH, EG by their meets X_1, X_0, X with the "denoting" line OX (through o), and in this correspondence denote the general parallel $RSTU$ by its meet x. Then $RV^2 = 4SU^2 = RT^2 + TU^2 + 2RT \cdot TU \,(= 2(SU^2 - ST^2))$, or $2SU^2 = RT^2 + TU^2 - 2ST^2$; so that, by CAVALIERI's theorem,

$$2 \times \sum_{X_{-1} \leqq x \leqq X_1} (SU^2) = \sum_{X_{-1} \leqq x \leqq X_1} (RT^2) + \sum_{X_{-1} \leqq x \leqq X_1} (TU^2) - 2 \times \sum_{X_{-1} \leqq x \leqq X_1} (ST^2).$$

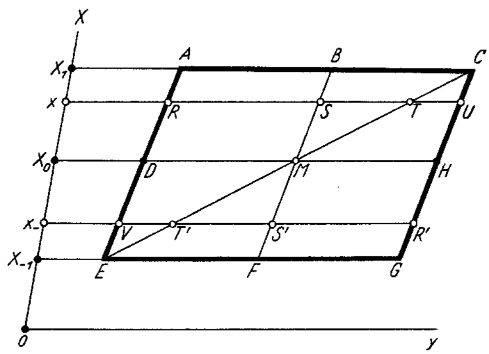


Fig. 67

Now consider the symmetrically situated parallel denoted by the meet x where $xX_0 = X_0 x_-$: by symmetry

$$\sum_{X_{-1} \leqq x \leqq X_1} (RT^2) = \sum_{X_1 \leqq x \leqq X_{-1}} (TU^2) = \sum_{X_{-1} \leqq x \leqq X_1} (TU^2);$$

and again

$$\sum_{X_{-1} \leqq x \leqq X_1} (RT^2) : \sum_{X_0 \leqq x \leqq X_1} (ST^2) = AC^3 : BC^3 = 8:1;$$

---

[16] JAMES GREGORY was a past master in its use—see *GPU passim*, but especially *EG*: 14—21: *analogia inter lineam meridianam planispherii nautici et tangentes artificialis geometrice demonstrata* .... In general, its use corresponded to treatments which involve transform of double integrals (with appropriate variable changes).

[17] In *geometria* ...: Book 7: 17—80. Thus his Theorem 8, Prop. 8: 33 is a proof that a cylinder has triple the volume of the cone of the same height and standing on the same base.

[18] Over some 500 pages in Books 1—6 of *geometria* ... .

[19] *geometria*: Book 1: prop. 24: 78ff.

therefore $$2\times \sum_{X_{-1}\leqq x\leqq X_1} (SU^2) = 4\times \sum_{X_0\leqq x\leqq X_1} (SU^2),$$
$$= 2\times \sum_{X_{-1}\leqq x\leqq X_1} (RT^2) - 2 \sum_{X_{-1}\leqq x\leqq X_1} (ST^2)$$
$$= 2\times 8\times \sum_{X_0\leqq x\leqq X_1} (ST^2) - 2\times 2\times \sum_{X_0\leqq x\leqq X_1} (ST^2)$$

or
$$\sum_{X_0\leqq x\leqq X_1} (SU^2) : \sum_{X_0\leqq x\leqq X_1} (ST^2) = \sum_{X_{-1}\leqq x\leqq X_1} (RU^2) : \sum_{X_{-1}\leqq x\leqq X_1} (RT^2) = (16-4):4 = 3:1$$

—a theorem which has an immediate application to all kinds of conic problems (and CAVALIERI develops the aspect very fully in Books 2 to 6 of his *geometria*). Here the $\sum(ST^2)$ is taken over a triangle (that is, ST varies linearly with the line-segment xX_0) and clearly the result is equivalent to

$$\int_{X_0}^{X_1} (X_1 X_0)^2 \cdot d(xX_0) : \int_{X_0}^{X_1} (xX_0)^2 \cdot d(xX_0) = 3:1,$$

or, by taking $xX_0 = x$ with $X_1X_0 = 1$,

$$(1=) \int_0^1 1^2 \cdot dx : \int_0^1 x^2 \cdot dx = 3:1 \qquad \left(\text{or } \int_0^1 x^2 \cdot dx = \tfrac{1}{3}\right).$$

The generalization of this approach is sketched by CAVALIERI in his *exercitationes*[20] but was given a thorough and exhaustive treatment by MENGOLI[21]. However—what is significant in a discussion of the development of indivisible theories in England—simplified and more accessible treatments of many of the basic theorems were given by TORRICELLI (who, as CAVALIERI's pupil, knew his work at first hand). Thus, TORRICELLI gives[22] an inverted treatment of CAVALIERI's result $\sum_{X_0\leqq x\leqq X_1} (SU^2) : \sum_{X_0\leqq x\leqq X_1} (ST^2) = 3:1$, deriving it from a Greek standard result (proved by an exhaustion-method in EUCLID): where the parallelogram $BCMH$ is a rectangle*, rotate it round BM as axis and consider the cylinder and inscribed cone traced out by the rectangle $BCMH$ and triangle BCM. Then

$$\sum_{X_0\leqq x\leqq X_1} (SU^2) : \sum_{X_0\leqq x\leqq X_1} (ST^2)$$
$$= \sum_{X_0\leqq x\leqq X_1} (\text{circle of radius } ST) : \sum_{X_0\leqq x\leqq X_1} (\text{circle of radius } SU)$$

= cone with axis BM and base radius BC : cylinder with axis BM and base-radius BC

$= 3:1$, by the standard Greek result.

* This involves no loss of generality since we need, in CAVALIERI's result, only to consider a general parallel STU. CAVALIERI, of course, used the analytical result in proof of the geometrical one.

[20] In Book 4: *de usu eorundem indivisibilium in potestatibus cossicis*: 243ff., where he also sketches an analytical approach suggested by BEAUGRAND (who may very well have communicated hints given him by FERMAT)—one which more closely follows what is accepted conventionally as CAVALIERIAN indivisible treatment.

[21] See his *geometria speciosa*, Bologna, 1659: Books 2, 3 and especially 6.

[22] In his lemma 20 of *de dimensione parabolae*: 57—58.

WALLIS, when he entered on his mathematical career in the early 1650's, derived his knowledge of CAVALIERI's theory of indivisibles in the first instance from TORRICELLI's *opera geometrica*, only later being able to read CAVALIERI's account in his *geometria*[23], and it was the experience of reading and digesting TORRICELLI's treatment which hardened the vague, unformed thoughts which had already come to him through his reading of EUCLID, APOLLONIUS and especially ARCHIMEDES. Specifically interested, as GREGORY ST. VINCENT before him, in circle-quadrature, WALLIS developed his ideas on the processes underlying existing indivisible theory (and the classical Greek exhaustion proofs) very much with that ideal before him. In particular, since, in considering the general line-element of an area to be evaluated, it is often possible to compare this with some power of a line-element already known and so to derive the numerical value of the ratio of the aggregate areas, he hoped to find some general method which would be applicable to the line-element of the circle, and so lead to quadrature. *

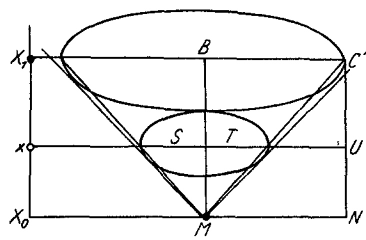


Fig. 68

Much as ROBERVAL and FERMAT had already found—though they had not published their results—WALLIS noticed that the CAVALIERI approach could be simplified by considering an analytical model of the limit-sum of the n^{th} powers of integers[24] and this he elaborates at great length in his *AI*[25]. Thus, in his prop. 19 he gives the theorem

$$\sum_{0\leqq i\leqq n} (i^2) : \sum_{0\leqq i\leqq n} (n^2) \left[= (n+1)\, n^2\right] = \frac{1}{3} + \frac{1}{6n},$$

and this, where n becomes indefinitely large, is equal to $\frac{1}{3}$ in the limit (since $\frac{1}{6n}$ becomes zero). Application to the CAVALIERI rectangle $BCHM$ is made by supposing the denoting segment $X_1X_0 = L$ to be divided into n equal parts L/n, with the general parallel STU cutting off a segment xX_0 which has λ

---

* In modern terms, since $\int_0^1 (1-x^2)^{\lambda}\cdot dx$ is easily calculable (by multiplication and integration) where λ is positive integral, he hoped to be able to calculate $\int_0^1 (1-x^2)^{\frac{1}{2}}\cdot dx$, which yields the quadrature of the circle quadrant. See chapter 4.

[23] WALLIS gives a detailed account of his mathematical development up to 1655 in the introduction to his *AI* (1656): iiff.: "at the end of 1650 I fell on Torricelli's mathematical writings (which, being otherwise occupied, I did not open till the following year 1651): there among other things he expounds CAVALIERI's *geometria indivisibilium*. CAVALIERI's work itself I had not at hand nor could I find it in the booksellers, but his method, as Torricelli expounds it, was the more pleasing to me because I had been turning something of the kind over in my mind ever since I first paid my respects to mathematics almost ...".

[24] A treatment considered (briefly) by CAVALIERI only in his (1647) *exercitationes*, which it is doubtful if WALLIS ever saw (compare previous note).

[25] See *AI*: Prop. 1ff.; and compare J.P. SCOTT: *The mathematical work of John Wallis*, London 1938, ch. 4, especially 27—49, and BOYER *(op. cit.)*: 141ff.

of them. Then

$$\sum_{X_0\leqq x\leqq X_1}(SU^2)=\lim_{n\to\infty}\sum_{0\leqq\lambda\leqq n}\left(n\times\frac{L}{n}\right)^2,$$

and

$$\sum_{X_0\leqq x\leqq X_1}(ST^2)=\lim_{n\to\infty}\sum_{0\leqq\lambda\leqq n}\left(\lambda\times\frac{L}{n}\right)^2,$$

so that

$$\sum_{X_0\leqq x\leqq X_1}(SU^2):\sum_{X_0\leqq x\leqq X_1}(ST^2)=\lim_{ln\to\infty}\left[\sum_{0\leqq\lambda\leqq n}\left(n\times\frac{L}{n}\right)^2:\sum_{0\leqq\lambda\leqq n}\left(\lambda\times\frac{L}{n}\right)^2\right]$$

$$=\lim_{l\to\infty}\left[\sum_{0\leqq\lambda\leqq n}(n^2):\sum_{0\leqq\lambda\leqq n}(\lambda^2)\right]=3:1$$

as before. More generally,

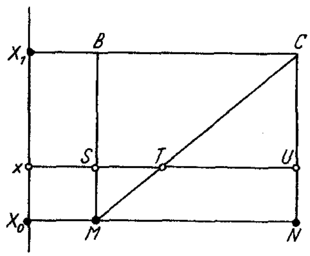


Fig. 69

$$\sum_{X_0\leqq x\leqq X_1}(SU^r):\sum_{X_0\leqq x\leqq X_1}(ST^r)$$

$$=\lim_{n\to\infty}\left[\sum_{0\leqq\lambda\leqq n}(n^r):\sum_{0\leqq\lambda\leqq n}(\lambda^r)\right],$$

and this WALLIS shows in a similar way for small values of r to be $(r+1):1$* (but extended by "analogy" to general r). (Later in AI[26] WALLIS states a recursive process of deriving formulae for sums of the n^{th} powers of integers, showing that

$$\sum_{1\leqq\lambda\leqq n}\left[\frac{\lambda}{1}\times\frac{\lambda+1}{2}\times\cdots\times\frac{\lambda+m-1}{m}\right]=\frac{n}{1}\times\frac{n+1}{2}\times\cdots\times\frac{n+m-1}{m}\times\frac{n+m}{m+1}.$$

An easy general proof of the theorem follows, though WALLIS contents himself with particular examples.)

All this is not new with WALLIS (though it had never been published before), nor does he claim originality in his application of it to finding the area under the ARCHIMEDEAN spiral[27] and the general parabolas $y=x^r$, r positive integral[28]. What is exciting, however, is his derivation in his props. 55 to 57 of the area under $y=x^{1/r}$, which he develops on a geometrical model from the allied rule for the area under $y=x^r$. Specifically his prop. 55 considers the problem of showing that $\sum_{0\leqq\mu\leqq m}(\mu^r):\sum_{0\leqq\mu\leqq m}(m^r)=1:(1+r)$ in the limit as m becomes illimitably great and in the particular case $r=\frac{1}{2}$. Take the parabola $AO'O$ defined by $D'O'^2=K\times AD'$, where

* So, where $r=3$, WALLIS uses the result that $\sum_{0\leqq\lambda\leqq n}(\lambda^3):\sum_{0\leqq\lambda\leqq n}(n^3)=\frac{1}{4}+\frac{1}{4n}$.

26 AI: prop. 182. The theorem had already been found by FERMAT in 1636 and used for the same purpose, together with the suggestive inequality,

$$\sum_{0\leqq\lambda\leqq n-1}(\lambda^m)<\frac{n^{m+1}}{m+1}<\sum_{1\leqq\lambda\leqq n}(\lambda^m),\qquad m>1.$$

27 AI: prop. 24. The theorem had been given both in CAVALIERI's *geometria* and TORRICELLI's *opera geometrica*.

28 Given both an indivisibles and an exhaustion-method treatment in various manuscripts of TORRICELLI, ROBERVAL and FERMAT. Treated by WALLIS in AI: prop. 23: 42ff.

$D'O'$ is a general ordinate to the abscisse AD'. Dividing $AT=L$ into equal sections $\frac{L}{n}$, of which AT' has λ, it is clear that $AD'=\frac{1}{K}\times\left(\lambda\times\frac{L}{N}\right)^2=T'O'$ or area $AT'TOO'$: area $ATOD=\operatorname{limit}_{n\to\infty}\left(\sum_{0\leq\lambda\leq n}(\lambda^2):\sum_{0\leq\lambda\leq n}(n^2)\right)$, $=1:3$. Therefore area $AO'OP'D$: area $ATOD=2:3=1:\frac{3}{2}$. But this is also $\lim_{m\to\infty}\left(\sum_{0\leq\mu\leq m}(D''O''):\sum_{0\leq\mu\leq m}(DO)\right)$, where we divide $AD=L'$ into m equal sections $\frac{L'}{m}$, of which AD'' has μ, and so $D''O''=K^{\frac{1}{2}}\left(\mu\times\frac{L'}{m}\right)^{\frac{1}{2}}$; or

$$\lim_{m\to\infty}\left(\sum_{0\leq\mu\leq m}k^{\frac{1}{2}}\left(\mu\times\frac{L'}{m}\right)^{\frac{1}{2}}:\sum_{0\leq\mu\leq m}K^{\frac{1}{2}}\left(m\times\frac{L'}{m}\right)^{\frac{1}{2}}\right)$$
$$=\lim_{m\to\infty}\left(\sum_{0\leq\mu\leq m}(\mu^{\frac{1}{2}}):\sum_{0\leq\mu\leq m}(m^{\frac{1}{2}})\right),$$
$$=1:\tfrac{3}{2}\left(=1:\left(1+\tfrac{1}{2}\right)\right).$$

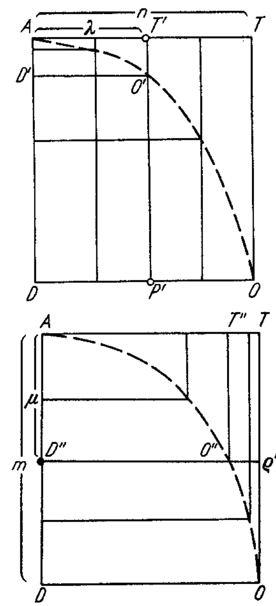


Fig. 70

More generally the result

$$\lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq n}(\lambda^r):\sum_{0\leq\lambda\leq n}(n^r)\right)=1:(1+r)$$

yields the corresponding result

$$\lim_{m\to\infty}\left(\sum_{0\leq\mu\leq m}(\mu^{1/r}):\sum_{0\leq\mu\leq m}(m^{1/r})\right)=r:(1+r)=1:\left(1+\frac{1}{r}\right).$$

This general result,

$$\lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq n}(\lambda^r):\sum_{0\leq\lambda\leq n}(n^r)\right)=1:(1+r),$$

where r is any real number, is strictly equivalent, where the integration interval $[O, X]$ is divided into n "indivisibles" X/n, to

$$\int_0^X x^r\cdot dx:\int_0^X X^r\cdot dx=1:(1+r),$$

or

$$\int_0^X x^r\cdot dx=\frac{1}{r+1}\int_0^X X^r\cdot dx=\frac{1}{r+1}X^{r+1}$$

—the form in which Wallis prefers to use the theorem in the later, more individual propositions of AI (but especially in the proposition which led up to his interpolation of $\square=\frac{4}{\pi}$, where $\square$ is, in equivalent form $\frac{1}{\int_0^1(1-x^2)^{\frac{1}{2}}\cdot dx}$[29]), though he uses a very cumbrous notation and no hint is given of the quantification of these definite integrals into any kind of indefinite forms.[30]

[29] See chapter IV.

[30] The introduction of a free variable upper bound in the Wallis integral is one of the improvements introduced by Newton in his manuscript annotations of *CUL Add.* 4000: 17ff. Significantly three years afterwards, in 1668, Mercator in his *logarithmotechnia* (and Wallis commenting on it) still use the rigid definite integral forms of the unmodified Wallis integral.

With WALLIS indivisible theory had reached, perhaps, its full power, but its mathematical heyday was inevitably short. In particular, the rigidity of its structure (with an implicit base-interval equisection) was an important defect, and it is significant that in some of the general theorems of his (1670) *mechanica* —in his examination, for example, of the general cycloid and cissoid areas— WALLIS himself was already rejecting the indivisible approach in favour of more general methods. WALLIS' *AI* had in many ways a tremendous influence on the later development of mathematical analysis, but more especially, perhaps, for the results contained in it than the methods by which they were derived. Inevitably in the rapid progress from 1660 onwards WALLIS' (and more generally CAVALIERI'S) indivisible method rapidly became obsolete.

Its two especial inadequacies were, first, that it presupposed (as CAVALIERI'S collective indivisible theory) an equisection of the base-interval, without which the method ceased to be rigorous; and, again, that the fundamental theorem on which its practical application was based,

$$\lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq n}(\lambda^r):\sum_{0\leq\lambda\leq n}(n^r)\right)=1:(1+r)\left(=\int_0^1\left(\frac{\lambda}{n}\right)^r\cdot d\left(\frac{\lambda}{n}\right)\right),$$

was restricted to a variable-range over the whole interval $\lambda\in[0,n]$ $\left(\frac{\lambda}{n}\in[0,1]\right)$, and no corresponding theorem could be proved to hold for the general subinterval $[O,\lambda]$ $\left(\text{which yields the indefinite integral, or the definite integral with free variable upper bound } \int_0^{\lambda/n}(\lambda/n)^r\cdot d(\lambda/n)\right)$. Of course, in many cases neither restriction mattered, and in other cases existing proofs could be modified to conform to the requirements of an indivisible proof. Thus, where WREN had rectified the general cycloid arc virtually by using a section of the base-interval in geometrical proportion[31], WALLIS reconstructed[32] an indivisible rectification of the whole cycloid arc (and, indeed, extended it to treat the more general contracted and protracted cycloids which had been shown by WREN and PASCAL the previous year to be transformable by length-preserving transforms into ellipse arcs).

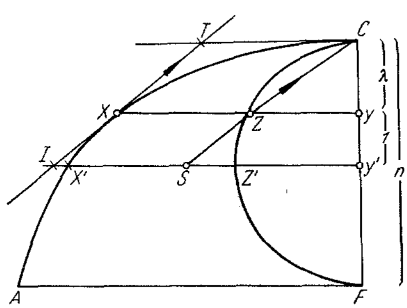


Fig. 71

Taking the cycloid arc $CXX'A$ and generating circle $CZZ'F$ WALLIS uses the property that the cycloid tangent XT at X is parallel to CZ, where C is the vertex and XY, moving parallel to the base AF, cuts the generator circle in Z. Consider the n-section of $CF=L$ in which CY has λ parts, and, where Y' is the next $(\lambda+1)^{\text{th}}$ section-point, draw $IX'SZ'Y$ through it parallel to the cycloid base: then

$$CY=\lambda\times\frac{L}{n},\quad ZC=(CY\times CF)^{\frac{1}{2}}=(\lambda n)^{\frac{1}{2}}\times\frac{L}{n},$$

[31] See next chapter.

[32] In his *tractatus duo; prior de cycloide ..., posterior ... de cissoide ...*, Oxford, 1659: 115ff.

and
$$CY:ZC = YY'\left(=\frac{L}{n}\right):ZS, \quad \text{or} \quad ZS(=XI) = \left(\frac{n}{\lambda}\right)^{\frac{1}{2}} \times \frac{L}{n};$$

so that, since in the limit as n becomes indefinitely large (and so YY', XI indefinitely small) we can take XI as the element XX' of the cycloid arc,

$$\text{cycloid arc } \widehat{CXA}:CF = \lim_{n\to\infty}\left(\sum_{0\leqq\lambda\leqq n}(XX' =)\left(\frac{n}{\lambda}\right)^{\frac{1}{2}}\times\frac{L}{n} : \sum_{0\leqq\lambda\leqq n}(YY' =)\frac{L}{n}\right)$$
$$= \lim_{n\to\infty}\left(\sum_{0\leqq\lambda\leqq n}(\lambda^{-\frac{1}{2}}) : \sum_{0\leqq\lambda\leqq n}(n^{-\frac{1}{2}})\right)\left(=\int_0^1\left(\frac{\lambda}{n}\right)^{-\frac{1}{2}}, d\left(\frac{\lambda}{n}\right)\right)$$
$$= 1:\left(1+\left(-\tfrac{1}{2}\right)\right) = 2:1$$

which shows $\widehat{CXA} = 2CF$.

It is clear that the indivisible theory is in transition to a more general form which considers equisections of the line-intervals (here IXT) not parallel to the basic integration-interval (CF). But no modification of the proof as it stands can show the more general result that cycloid arc $CX = 2\times CZ$—which follows as an immediate corollary of WREN's exhaustion treatment—because the basic indivisible theorem $\left(\text{that, equivalently, } \int_0^K \left(\frac{\lambda}{n}\right)^{-\frac{1}{2}}\cdot d\left(\frac{\lambda}{n}\right) = 1:\frac{1}{2},\ K=1\right)$ cannot be modified to the case where the upper bound is allowed to vary freely in 0, 1.

Apparently WALLIS himself was not aware of this limitation of his indivisible method. In a series of letters to him in the 1690's LEIBNIZ tried to suggest these restrictions but could not make WALLIS—then an old man—see his point[33]. In his letter of 19 March 1696/7 LEIBNIZ pinpoints the difficulty[34]; "I wish there were someone to carry through your method (of the arithmetick of infinites) to higher and more composed lines. For it does not lack usefulness. Since I see that ... I said* the method could not be extended to quadratures of segments, but merely covered whole quadratures, ... I wished to look more closely into the matter in the case of the cissoid ... And it seemed to me that its application to segments did not lack difficulty, because there collections of numbers into a single (limit) number have no easy place". LEIBNIZ then sketches[35] WALLIS' interpolation of the sequence of integrals defined virtually between fixed bounds x = diameter length, $x=0$, "by whose help is neatly found the area of the whole cissoid space, assuming the quadrature of the circle. But, in general, in the case of the partial segment two terms cannot be added into a single number, and so that elegant progression of numbers added into a single (limit) number on which the interpolation depends seems to cease to be of use in considering the general partial segment".

---

* In a review of volume 1 (1695) of WALLIS' *opera mathematica* printed in *AE* (June 1696).

[33] See WALLIS: *opera mathematica*, **3** (1699): 672ff.

[34] *opera* **3**: 673.

[35] See chapter 3.

LEIBNIZ suggests an ingenious extension. What has prevented application of the general indivisible theorem, $\lim_{n\to\infty}\Big(\sum_{0\leqq\lambda\leqq n}(\lambda^r):\sum_{0\leqq\lambda\leqq n}(n^r)\Big)=1:(1+r)$, to some particular case $\lim_{n\to\infty}\Big(\sum_{0\leqq\lambda\leqq m}(\lambda^r):\sum_{0\leqq\lambda\leqq m}(n^r)\Big)$, $m<n$, is that the limit-summation does not take place over all the variable range $\lambda\in[0,n]$. If, then, we can so define the function λ^r in terms of a new variable μ ranging over $[0,m]$, we can "compress" the integration-interval $[0,n]$ into that of $[0,m]$ by $\lambda=f(\mu)$ and consider the limit-form

$$\lim_{n\to\infty}\Big(\sum_{0\leqq\mu\leqq m}(f(\mu))^r:\sum_{0\leqq\mu\leqq m}(f(n))^r\Big),^{*}$$

in which we may be able to apply the indivisible theorem (with, perhaps, the help of a sequence of WALLIS interpolations—"nescio an tunc facile futurum sit pervenire ad progressiones numerorum aptas interpolationi").

Such a programme, if feasible, would be immensely cumbrous and difficult, but WALLIS cannot see the point—which becomes crucial, for example, in the attempt to interpolate by his method a sequence from which may be derived the general circle-segment—and refers in later letters to a valid proof for the general cissoid segment given by him in his (1670) *mechanica*[36]. But if we examine this proof closely, we find it based on a lemma proved geometrically by an exhaustion-method and not by any theorem in the arithmetick of infinites, *viz*:[37]

$$\int_0^{\vartheta}(1\pm\cos x)\cdot dx^{**}=\vartheta\pm\sin\vartheta,$$

around which indivisible considerations are inserted.

In fact, taking $AD\alpha$ the generating semi-circle of the cissoid ADC (defined such that, for the circle radius CD perpendicular to $A\alpha$ and AB a general chord meeting CD in H and the cissoid in b, $BH=Hb^{***}$, consider an equisection of the arc $AD\alpha$ into n parts, two of which are XB, BX': for n large we can take α, P, X; α, X', Y respectively colline sets of points, where AX, AX' cut BV in Y, P; bv in y, p; αE in v, π; so that $PY:py:\pi v=AV:Av:A\alpha=(1-\cos 2\vartheta):(1+\cos 2\vartheta):2$ (where $\widehat{A\alpha B}=\frac{1}{2}\times\widehat{ACB}=\vartheta$). Again $\Delta APY=\frac{1}{2}\times PY\times AV$; $\Delta Apy=\frac{1}{2}\times py\times Av\Big(=\frac{1}{2}\times\Big(PY\times\frac{Av}{AV}\Big)\times Av\Big)$; and $\Delta A\pi v=\frac{1}{2}\times\pi v\times A\alpha\Big(=\frac{1}{2}\times\Big(PY\times\frac{A\alpha}{AV}\Big)\times A\alpha\Big)$; or trapezium $pyv\pi=\Delta A\pi v-\Delta Apy=\frac{1}{2}\frac{PY}{AV}(A\alpha^2-Av^2)$, $=\frac{1}{2}\frac{PY}{AV}(2A\alpha-AV)$

* LEIBNIZ's idea is that by some transform $x\to y: x=\Psi(y)$ we can reduce the integral $\int_0^{\alpha}\Phi(x)\cdot dx$ into the form $\int_0^1\Phi(\Psi(y)).\ d(\Psi(y))$, to which conventional indivisible techniques may be applied.

** His "figura omnium sinuum versorum".

*** We easily show $bv^2=\frac{Av^4}{B'v^2}=\frac{Av^3}{v\alpha}$; $VC=Cv$ (or $AV=v\alpha$).

[36] That is, *mechanica, sive de motu* ..., Oxford, 1670: *pars secunda, quae est de centro gravitatis ejusque calculo*: ch. 5, prop. 29, *idem aliter* $\cdot\equiv\cdot$ *opera mathematica* **1** (1695): 904—910.

[37] *mechanica*: Book 2: prop. 17.

since $Av = A\alpha - AV$. Finally, as n increases indefinitely we can see ΔAPY as an area-element of the semi-circle, with the trapezium $pyv\pi$ as a corresponding

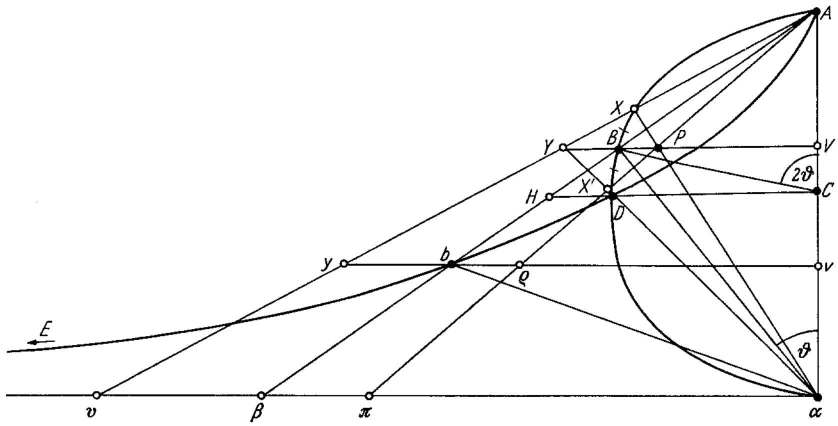


Fig. 72

area-element of the cissoid-space; so that, where E is $\infty_{\alpha\beta} = \infty_{Ab}$, area (circle segment $\widehat{ABA}$): area (cissoid space $b\beta\overset{\smile}{E}b$)

$$= \lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq\vartheta}[\Delta APY] : \sum_{0\leq\lambda\leq\vartheta}[\text{trapezium } pyv\pi]\right)$$

$$= \lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq\vartheta}\left[\frac{Av}{A\alpha}\right] : \sum_{0\leq\lambda\leq\vartheta}\left[\frac{2A\alpha - Av}{A\alpha}\right]\right)$$

$$= \int_0^{\vartheta}\frac{1-\cos 2x}{2}\cdot dx\left[= \frac{1}{4}\int_0^{2\vartheta}(1-\cos x)\cdot dx\right] : \int_0^{\vartheta}\left(2 - \frac{1-\cos 2x}{2}\right)\cdot dx$$

$$\left[= 2\int_0^{\vartheta}dx - \frac{1}{4}\int_0^{2\vartheta}(1-\cos x)\cdot dx\right]$$

$$= \frac{1}{4}(2\vartheta - \sin 2\vartheta) : \left[2\vartheta - \frac{1}{4}(2\vartheta - \sin 2\vartheta)\right]$$

$$= (2\vartheta - \sin 2\vartheta) : (6\vartheta + \sin 2\vartheta).\star$$

Here—and in his exhaustive treatment of the general cycloid segment along similar lines[38]—WALLIS is unconsciously on new ground. In effect, the proof proceeds by considering the "indivisible" $\widehat{X\alpha X'}$ (indefinitely small equisection) of the angle $\widehat{A\alpha D'}$, and it is on the basis that these indivisibles are all equal that he can, in his integral comparison, ignore them.

⋆ WALLIS gives this in a final tidier form: since the circle segment $\widehat{ABA}$ is $\frac{1}{2}AC \times (2\vartheta - \sin 2\vartheta)$, the cissoid-space $p\beta\overset{\smile}{E}b$ is $\frac{1}{2}AC \times (6\vartheta + \sin 2\vartheta)$: and so, adding to each the equal areas $\Delta\alpha AB = \Delta\alpha b\beta = AC \times \sin 2\vartheta$, circle segment $\widehat{AB}\alpha A = \frac{1}{2}AC \times (2\vartheta + \sin 2\vartheta) = \frac{1}{3}(\frac{1}{2}AC \times (6\vartheta + 3\sin 2\vartheta)) = \frac{1}{3} \times$ cissoid-space $\alpha b\overset{\smile}{E}\alpha$.

[38] In the preceding prop. 22.

The logical form of WALLIS' proof of his basic lemma, prop. 17, will be considered later (in the next chapter), but it is worth seeing exactly how the indivisible limit-sum consideration (obvious in a line-segment equisection) carries over to an angle equisection.[39] Consider, then, the semicircle $\widehat{AD\alpha A}$, where $Z\widehat{\alpha}X$ is the λ^{th} equisection of $A\widehat{\alpha}D$ in an n-part equisection; for simplicity, again, take diameter $A\alpha=2$ and $A\widehat{\alpha}D=\vartheta$, so that $A\widehat{\alpha}Z=\frac{\lambda}{n}\times\vartheta$ (where $X\widehat{\alpha}Z=\vartheta$). WALLIS' proof, in effect, assumes that segment $\widehat{AZD\alpha A}=\lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq n}[\Delta X\alpha P]\right)$ where P is the meet of αX with the half chord ZS: by geometrical considerations we easily show that segment $\widehat{AZD\alpha A}=\frac{1}{2}(2\vartheta+\sin 2\vartheta)$; further

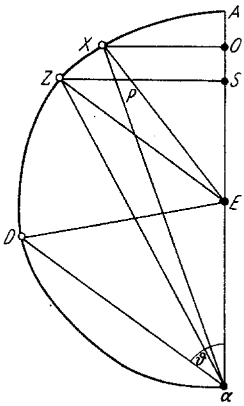


Fig. 73

$$\Delta Z\alpha P=\frac{1}{2}ZP\times S\alpha\approx\frac{1}{2}ZX\times S\alpha\approx\frac{1}{2}\widehat{ZX}\times S\alpha,$$

since

$$\widehat{ZPX}\left(=\widehat{OXP}=\frac{\pi}{2}-\frac{\lambda-1}{n}\times\vartheta\right):\widehat{ZXP}\left(=\widehat{\alpha ZS}=\frac{\pi}{2}-\frac{\lambda}{n}\times\vartheta\right)$$

tends to unity as n becomes indefinitely large so that

$$\Delta Z\alpha P\approx\frac{1}{2}\times\frac{2\vartheta}{n}\times\left(1+\cos\left(\lambda\times\frac{2\vartheta}{n}\right)\right);$$

or

$$\lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq n}(\Delta Z\alpha P)\right)=\lim_{n\to\infty}\left(\frac{1}{2}\times\sum_{0\leq\lambda\leq n}\left[\left(1+\cos\left(\lambda\times\frac{2\vartheta}{n}\right)\right)\times\frac{2\vartheta}{n}\right]\right)$$

$$=\frac{1}{2}\times\int_0^{2\vartheta}(1+\cos x)\cdot dx,$$

equivalently. We note that there is no comparison of integrals here*, but an absolute limit-process (given by WALLIS in a more discursive form) which defines

---

* Though by comparing two segments $\widehat{AZD_1\alpha A}$, $\widehat{AZD_2\alpha A}$ a suitable limit-sum comparison can be set up in more traditional indivisible theory manner.

[39] WALLIS, of course, was not the first to define an integral by limit-equisection of an angle. Perhaps that distinction belongs to ROBERVAL, particularly in his rectification of the general cycloid arc which is based on a lemma (his *propositio lemmatica prima*) which is equivalent to WALLIS' though not published till long after both were dead. See *Divers ouvrages de M. de Roberval*, Mémoires de l'ac. roy. des sciences, 1666—1699: **6** (Paris, 1730): 1—478, especially 247—359: *Traité des indivisibles*; and 361—427: *de trochoide ejusque spatio*, of which ROBERVAL's lemma is pp. 383—385). (ROBERVAL claims to have used the lemma to have rectified the cycloid in the period 1635—1640 (see p. 424), and from the crudities of his proof-structure and baldness of his concepts I see no reason to doubt the priority claim). But a fairly definitive treatment of angle indivisibles was given by PASCAL in his *lettres de A. Dettonville, contenant quelques unes de ses inventions en géométrie*, Paris, 1659: especially in the tract, *Un traitté des sinus et de leurs onglets*.

the integral $\Phi(x)$ in RIEMANNIAN form: that is,

$$\int_0^X \Phi(x)\cdot dx = \lim_{n\to\infty}\left(\sum_{0\leq\lambda\leq n}\Phi\left(\lambda\times\frac{X}{n}\right)\times\frac{X}{n}\right),$$

where the integration range $[0, X]$ is equisected. In the CAVALIERI integral-comparison indivisible theorems the equisection X/n could be ignored as an eliminable common factor, but WALLIS' example shows how natural it was to pass to an absolute concept of the integral in which the equisection X/n is reintroduced. It remains only to consider the extension which allows non-equisections of the integration-interval on the basis of some concept of maximum and minimum values of a function in a specified interval, and we have the CAUCHY-RIEMANN definition of an integral as a limit-sum. Some such consideration had already been introduced by PIETRO MENGOLI[40] but the general extension using a geometrical model was made by several mathematicians in the mid-century using exhaustion techniques.[41]

Perhaps we can say that indivisible and arithmetick of infinite treatments were a natural (and far from unrigorous) preliminary to a more exact theory of integration. Indeed, very widely in the mid-century indivisible techniques, while admittedly lacking the refinement of rigid proof, were seen as obvious and plausible, with the practical advantages of being easily and quickly applicable to a wide range of problems—but most important, as capable of a rigid (if long-winded) analogous proof by an exact exhaustion-method.

The great danger lurking in the standard indivisible proof was that no adequate notation had been devised to facilitate its use, and the existing universally accepted verbal treatment could, in its looseness of expression, lead to a "natural" but fallacious application of the method. A typical case is that of THOMAS HOBBES

[40] See his *geometria speciosa*, Bologna, 1659: ch. 6, *passim*.

[41] The general method is discussed in the next chapter, but consider the following particular examples:

EVANGELISTA TORRICELLI:

α. *opere* 1.2. Various manuscripts, but especially
- 227—274: *de infinitis hyperbolis*,
- 275—328: *de infinitis parabolis*,
- 335—347: *de hemhyperbola logarithmica*.

β. E. CARRUCCIO: *Evangelista Torricelli: de infinitis spiralibus*, Pisa, 1955.

R. DESCARTES:

letter to MERSENNE (on cycloid quadrature) of 27 July 1638, ·≡· (ed. ADAM & TANNERY) *Oeuvres* **2**: 260ff.; and compare 135ff.

FERMAT: *de linearum curvarum cum lineis rectis comparatione dissertatio geometrica*, published in appendix to A. LALOVERA: *tractatus de cycloide*, Toulouse, 1660 ·≡· (ed. P. TANNERY & C. HENRY) *Oeuvres* **1** (Paris 1891); 211—254.

C. HUYGENS:

α. *theoremata de quadratura hyperboles, ellipsis et circuli ex dato portionum gravitatis centro*, Leyden, 1651 ·≡· *Oeuvres* **11** (1908): 282—313.

β. Various manuscripts on quadratures and rectifications ≡ *Oeuvres* **14** (1920): 234ff.

ROBERVAL:

letter to TORRICELLI of 1646 *de solido acuto hyperbolico*. Méms. de l'ac. roy. des sciences **6** (Paris, 1730); 428—437.

who, in an enduring polemic against WALLIS[42], tried to show that the general parabola-arc is equal to a rational line-length, not allowing unfortunately for the modifying effect of changing gradient. A more subtle fallacy arising from loose indivisible thought was the belief—perhaps first sustained (and later retracted) by PAUL GULDIN[43] but widespread in the 1650's—that the arc-length of the first revolution of the Archimedean spiral was equal to half that of the circumscribing circle: a position likewise upheld by HOBBES[44] but which received an especially lengthy treatment at the hands of THOMAS WHITE of St. Albans[45], though ROBERVAL and TORRICELLI (in manuscript) had already asserted[46] that the spiral-arc is equal to that of a definable parabola-arc—rigid proof of which was given by PASCAL[47] shortly afterwards.

Despite, however, such theoretical disadvantages it remains historical fact that a large number of techniques later to become standard in integral calculus were introduced on an indivisible theory basis in simple examples which could in some way sustain the basic indivisible proof requirement of an interval-equisection. A fine example is that of WILLIAM NEIL's first rectification[48] (in 1657) of an algebraic curve (the semicubical parabola, $ky^2 = x^3$). Taking the curve by its geometrical definition as the point-set AfF, where for all f on $\widehat{AfF}$, $EF^2 : ef^2 = AE^3 : Ae^3$, NEIL equisects AE in n points e_λ, $\lambda = 1, 2, 3, \ldots, n$ ($e_n = E$) and through each section-point draws the normal ef. Clearly the rectification is reducible to finding the limit-sum

$$\lim_{n\to\infty}\Big(\sum_{0\leqq\lambda\leqq n}[f_\lambda f_{\lambda+1}]^2\Big) = \lim_{n\to\infty}\Big(\sum_{0\leqq\lambda\leqq n}[(e_{\lambda+1}f_{\lambda+1} - e_\lambda f_\lambda)^2 + (e_{\lambda+1}e_\lambda)^2]^{\frac{1}{2}}\Big).$$

[42] See HOBBES' *Six lessons to the Professors of Mathematics of the Institution of Sir Henry Savile*, Oxford, 1656: especially 41 ff. HOBBES' mistake was quickly pointed out by HUYGENS in his letter to WALLIS of 15 March 1656: *Oeuvres* **1**: 392ff. Significantly HUYGENS, a little later, —in the famous manuscript whose diagram is dated by an "ἕυρηκα, 27 October 1657"—was to show rigorously the logarithmic nature of the general parabola-arc.

[43] See his *centrobaryca*, Vienna, 1635: Book 2.

[44] Compare his *examinatio et emendatio mathematicae hodiernae*, London, 1660: Dialogue 5 ·≡· (ed. MOLESWORTH) *opera latina*: **4**: 189.

[45] In *exercitatio geometrica de geometria indivisibilium et proportione spiralis ad circulum*, London 1658.

[46] Probably on the basis of the length-preserving convolution transform to be developed later by WREN (see WALLIS: *tractatus duo de cycloide ..., de cissoide ...*, Oxford, 1659: especially 104—108) and given an exact treatment by exhaustion techniques in JAMES GREGORY's *GPU*: props 12—18.

[47] In his tract, *Lettre ... a Monsieur A.D.D.S. ... en luy envoyant la demonstration à la manière des anciens de l'égalité entre les lignes spirale et parabolique*, printed in his *Lettres de A. Dettonville ...*, Paris, 1658.

[48] Printed by WALLIS in his *tractatus duo de cycloide ..., de cissoide ...*: 90ff. As WALLIS observes, the method used—strictly equivalent to the modern integration formula, $s = \int_0^X \left(1+\left(\frac{dy}{dx}\right)^2\right)^{\frac{1}{2}} \cdot dx$—had been suggested in his *AI*: prop. 38 together with an outline of its possible application to the parabola (though he did not notice the parabola property of constant subnormal on which the first—and easiest—rectification proofs were based. (VAN HEURAET gives a similar but slightly differing method in a letter to VAN SCHOOTEN of 1659, printed in the Latin version of DESCARTES' *Géométrie* **1** (Amsterdam, 1659): 517—520.)

Now consider the simple parabola AbB defined, where EB is an arbitrary length, as the point-set b such that $EB^2:eb^2 = AE:Ae$,* where b is its meet with ef); and drawing SI perpendicular to EF, where S (in EF) is defined by $ES \times EF =$ parabola area (AEB) with s the meet of SI with ef, define the point-set h such that, where h is on ef, $(eh)^2 = (es)^2 [= (ES)^2] + (eb)^2$. It is then easy to show that the point-set h is a (simple) parabola IhH (with vertex F in AE) defined by $(eh)^2 : (EH)^2 = Fe : FE$**, and NEIL uses this to show that $ES \times \widehat{AfF}$ = parabola segment $(AEHI)$.*** Specifically, where n is taken indefinitely large

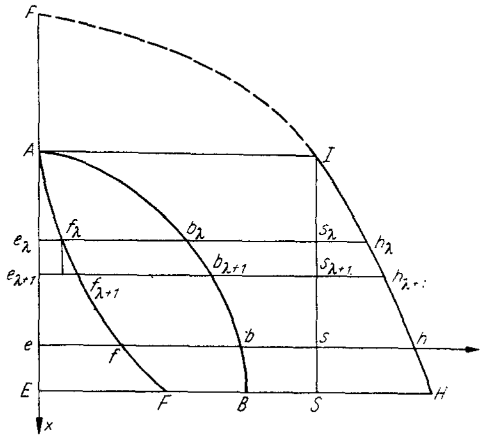


Fig. 74

$$e_\lambda b_\lambda \times e_\lambda e_{\lambda+1} = \text{area } (e_\lambda e_{\lambda+1} b_{\lambda+1} b_\lambda)$$
$$= \text{parabola area } (Ae_{\lambda+1} e_\lambda) - \text{parabola area } (Ae_\lambda b_\lambda)$$
$$= (e_{\lambda+1} f_{\lambda+1} - e_\lambda f_\lambda) \times \frac{\text{parabola area } (AEB)}{EF} [= ES],$$

or

$$(f_{\lambda+1} f_\lambda)^2 : (e_{\lambda+1} e_\lambda)^2 \big(= [(e_{\lambda+1} f_{\lambda+1} - e_\lambda f_\lambda)^2 + (e_\lambda e_{\lambda+1})^2] : (e_\lambda e_{\lambda+1})^2\big) = (e_\lambda h_\lambda)^2 : (ES)^2;$$

so that

$$\lim_{n\to\infty} \left[\sum_{0 \le \lambda \le n} (f_{\lambda+1} f_\lambda) \right] = \frac{1}{ES} \times \lim_{n\to\infty} \left[\sum_{0 \le \lambda \le n} (e_{\lambda+1} e_\lambda \times e_\lambda h_\lambda) \right]$$
$$= \frac{1}{ES} \times \text{area } (AEHI).$$

Finally, where $AE = a$, $EF = c$, we easily show

$$\widehat{AfF} = \frac{(4a^2 + 9c^2)^{\frac{3}{2}} - 8a^3}{27c^2}.\text{****}$$

---

* Analytically, with respect to origin A, $EB^2 \times x = AE \times y^2$. NEIL, in fact, defines the semicubical parabola by its then standard form, $EF:ef (= \frac{2}{3} AE \times EB : \frac{2}{3} Ae \times eb) =$ parabola area (AEB) : parabola area (Aeb).

** NEIL uses the awkward (but rigorous) result that $(e_{\lambda+1} h_{\lambda+1})^2 - (e_\lambda h_\lambda)^2 = (e_{\lambda+1} b_{\lambda+1})^2 - (e_\lambda b_\lambda)^2$, which increases with λ in arithmetical progression. More generally (an improvement virtually introduced by BROUNCKER and added in postscript), $(eh)^2 = (ES)^2 + (eb)^2 = EH^2 - EB^2 + EB^2 \times \frac{Ae}{AE}$, so that $(eh)^2 : (EH)^2 = (FA + Ae) : (FA + AE)$, where $FA = \left(\frac{ES}{EB}\right)^2 \times AE$.

*** Stated by NEIL in the proportion form,

area $(AEHI)$: area $(\square AESI)$: parabola area (AEB) $= \widehat{AfF} : AE : EF$.

**** For, taking the arbitrary length $EB = b$,

$$ES = \frac{1}{EF} \times \text{parabola area } (AEB) = \frac{1}{c} \times \frac{2}{3} ab,$$

It is unfortunate that the complexities of NEIL's argument—the introduction of the parabola AbB which so simplifies NEIL's treatment is, in fact, from a general viewpoint surplus—blurs the main outline of the approach. What we have, in effect, is a differential triangle technique applied to a rectification problem—and it was so emphasised, as we shall see, by JAMES GREGORY in his generalization of the NEIL method in *GPU* (prop. 6). The casually introduced proportion, $e_\lambda h_\lambda : ES = f_{\lambda+1} f_\lambda : e_{\lambda+1} e_\lambda$ [= element of arc $(\widehat{AfF})$: element of line (AE) as the number of equisections becomes unlimitedly large], is fundamental: specifically, where $fh' (= eh)$ is tangent to AfF at f and $e'h'$ is drawn parallel to AE (meeting ef in e'), then $e'h' = eh$; so that $fh' : e'h'$ $(= eh : ES)$ = element of arc $(\widehat{AfF})$: element of line (AE)—the classical "BARROW" differential triangle definition.

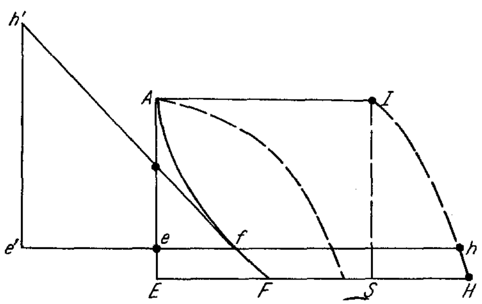


Fig. 75

Such an approach is, however, far more general than the indivisible (interval equisection) form in which NEIL gave it. Likewise, while an essential part of his proof is that for two neighbouring equisection points of AE, the difference of the squares of the two respective pairs of ordinates eb, eh be in arithmetical progression (a property unique to the parabola), the proof that IhH is a parabola can, as I have shown, be reformulated independently of an interval equisection (by limit considerations equivalent to the differential triangle approach).

The remark is general for all the calculus methods originally formulated in indivisible terms. When the concept of indivisible is found inadequate and more general concepts are found necessary—historically, about 1660—we find the indivisible methods embedded in a more general theory. The symbolic techniques of modern calculus owed much to the first rough indivisible formulation, but when outgrown it had to be discarded and even discredited: it is unfortunate that so many historians have not been able to see through that discrediting to the fact that indivisible theories had a real power and were not essentially unrigorous.

---

and
$$AF = \left(\frac{ES}{EB}\right)^2 \times AE = \left(\frac{2ab}{3c}\right)^2 \times \frac{1}{b^2} \times a = \frac{4a^3}{9c^2};$$

so that area $(AEHI)$ = parabola area (FEH) − parabola area (FAI)
$$= \frac{2}{3}\left(a + \frac{4a^3}{9c^2}\right) \times \left(b^2 + \left(\frac{2ab}{3c}\right)^2\right)^{\frac{1}{2}} - \frac{2}{3} \times \frac{4a^3}{9c^2} \times \frac{2ab}{3c}$$
$$= \frac{2ab}{81c^3} \times [(4a^2 + 9c^2)^{\frac{3}{2}} - 8a^3];$$

with finally $\widehat{AfF} = \frac{1}{ES}\left[= \frac{3c}{2ab}\right] \times \text{area}\,(AEHI)$. It is interesting to compare the treatment with the conventional modern treatment by $\widehat{AfF} = \int_0^a \left(1 + \left(\frac{dy}{dx}\right)^2\right)^{\frac{1}{2}} \cdot dx$, where $y = \frac{c}{a^{\frac{3}{2}}} x^{\frac{3}{2}}$ $\left(\text{and so } \frac{dy}{dx} = \frac{3c}{2a^{\frac{3}{2}}} x^{\frac{1}{2}}\right)$.

IX. Calculus

2. The method of proof by exhaustion

We must never, in developing the history of mathematics, accept some particular aspect of technique at its contemporary evaluation, but rather consider it in the light of modern knowledge and experience. The method of proof by exhaustion is a case in point.

A general 17th century attitude to the exhaustion-proof saw it as a rigorous but enormously prolix, particularised and even antiquated method. This, as I will show, is in large part an illusion. Rather, the prolixity of the method as it was used in 17th century mathematics came from a roughness and crudity of logical exposition, and it is possible, by developing the method in logical symbolism, to see its general power and its acceptability as a proof-form. Indeed, the method as it was generalized in the 17th century from the relatively simple classical Greek exhaustion technique becomes equivalent to a CAUCHY-RIEMANN definite integral defined on a convex point-set. What the 17th century mathematicians regarded as prolixity is merely the result of their unwillingness to detach the logical proof-form from each particular case, using it as a logical prenex. So, instead of stating the relatively simple conditions under which the general form could be applied and thus deriving the required result immediately, the 17th century mathematician felt that the whole complex procedure of setting up inequalities, and of using a *reductio ad absurdum* to prove the equivalence of upper and lower bounds to the integral (which shows it unique), had to be given *in extenso* over and over again in each particular application.

No one has, unfortunately, given an adequate analysis of the method as it was generalised in the 17th century, and even in considering the relatively simple examples to be found in Greek mathematics most analyses have been shoddy, usually remaining content to sketch the type-example of EUCLID: *Elements* Book 12: prop. 2 (which shows that circle-areas are as the squares of their diameters). E. J. DIJKSTERHUIS[1] is, however, the exception, and the approach which will be developed in the rest of this chapter derives essentially from his.

Historically, the method of proof by exhaustion seems to have been developed as a process for theoretically exhausting the area under given geometrical figures, and the Greeks themselves gave its invention to EUDOXUS[2]. In particular, inspiration seems to have come from the early Greek method of approximating to the area of a circle by considering the infinite sequences of circumscribed and inscribed regular polygons. Systematising this—by considering, in fact, the sequence of polygons of $\lambda = 2^n$ sides, $n = 1, 2, 3 \ldots$ successively—we have the basis for the EUDOXIAN proof of EUCLID **12**.2. Above all other Greek mathematicians ARCHIMEDES was the master of this method, using it elegantly and powerfully throughout his works in a variety of ways, but especially that which DIJKSTERHUIS has termed the "compression method". Since by far the largest

---

[1] See his *Archimedes* (English translation by C. DIKSHOORN) Copenhagen, 1956: especially 130—133.

[2] DIJKSTERHUIS (p. 130) emphasises that this exhaustion of the area, or more generally this passage to the infinite, is a limit-process which considers the bound to which the sequences considered converge, and on that ground prefers to name the technique the "indirect method for infinite processes".

number of 17th century exhaustion proofs follow this "ARCHIMEDEAN" model and since the generalisations which appear in the period use it as a spring-board for further development, it is very necessary (and strictly relevant) to consider this model in detail.

(Two forms of the ARCHIMEDEAN method exist, corresponding to the two basis number operations of $\pm$ and $\stackrel{\times}{\div}$,* but I will consider only the former, an analogous treatment of the latter being immediately derivable.)

Let us assume that, in some way supposed unique, we can assign real number measures $\lambda, \mu, \ldots$, $(\lambda, \mu, \ldots \in [-\infty, +\infty])$ to entities such that we can attach meaning to the operation of addition $\pm$ (that is, such that $\lambda \pm \mu$ is also a real-number measure) and that we are able, by considering their numerical values, to order the values $\lambda, \mu, \ldots$ in some way, say, $\lambda < \mu < \cdots$; and finally that we can by use of this ordering and the operation of $\pm$ bound suitable $\lambda, \mu, \ldots$ (in some specificable way) with upper and lower limits L, l; M, m; ... respectively to any required degree of accuracy—that is, such that $|L-\lambda|$, $|l-\lambda|$; $|M-\mu|$, $|m-\mu|$; ... can be made as small as we wish. We can then represent the proof-structure of the ARCHIMEDEAN model in the following way: If

1. $(i)\begin{pmatrix} A_i > \alpha > a_i \\ B_i > \beta > b_i \end{pmatrix}$.

2. $(i, j)\left(j > i \cdot \rightarrow \cdot \begin{bmatrix} A_i > A_j;\ B_i > B_j \\ a_i < a_j;\ b_i < b_j \end{bmatrix}\right)$.

3. For i sufficiently large (with ε indefinitely small) (EN) (i) $\big(i > N \cdot \rightarrow \cdot (A_i - a_i) < \varepsilon\big)$.

4. *a.* $(i)\,(a_i = b_i)$; *b.* $(i)\,(A_i = B_i)$.

5. $(i)\,(a_i > 0)$ $\big($and so immediately $\alpha, \beta > 0$ and $(i)\,(A_i, B_i > 0)\big)$, then $\alpha = \beta$.

ARCHIMEDES' standard proof shows that $\alpha \neq \beta$ is impossible by using the logical trick of *reductio ad absurdum.* Thus, supposing $|\alpha - \beta| = \lambda > 0$, by 3: (EN) (i) $\big(i > N \cdot \rightarrow \cdot (A_i - a_i) < \varepsilon\big)$ there exists some number N' such that, for all $i > N'$, $(A_i - a_i) < \lambda$, and the further argument proceeds by examining two cases:

Case 1: $\alpha > \beta$, or $|\alpha - \beta| = (\alpha - \beta) = \lambda > (A_i - a_i)$.

By 1. $(A_i - a_i) > (\alpha - a_i)$, or $(\alpha - \beta) > (\alpha - a_i)$, so that $\beta < a_i = b_i$, which contradicts *2.*

Case 2: $\alpha < \beta$, or $|\alpha - \beta| = (\beta - \alpha) = \lambda > (A_i - a_i)$.

By *1,* $(A_i - a_i)$ $\big(= (B_i - a_i)\big) > (\beta - a_i)$, or $(\beta - \alpha) > (\beta - a_i)$, so that $\alpha < a_i$, which contradicts *2.*

Finally, since $\alpha > \beta$, $\beta > \alpha$ exhaust the cases of $|\alpha - \beta| > 0$, $|\alpha - \beta| \ngtr 0$, or $|\alpha - \beta| = 0$, and $\alpha = \beta$.**

All this by itself is bare abstraction, and the richness of the model in its geometrical application lies in its use of lemmas (given by ARCHIMEDES apparently as axioms, though justifiable in an obvious way by limit-considerations) which

* Isomorphic under the mapping $a \pm b \leftrightarrow \alpha \stackrel{\times}{\div} \beta$, where $c \leftrightarrow \gamma$ by $c = \log(\gamma)$ (with $a = \log(\alpha)$, $b = \log(\beta)$, see chapter one). In particular, the mapping preserves the inequality $a < b$, since $\alpha < \beta$ follows from $\log \alpha < \log \beta$.

** The proof is in no way unique, though readily generalisable to more extended models—in particular, an important alternative proof developed by JAMES GREGORY is given below.

set up various measure-inequalities for convex curves.* A first set of lemmas introduce the application of exhaustion proof-methods to quadrature of the areas under (wholly) convex curves, and are both intuitively obvious and standard in classical Greek mathematics (particularly in the EUDOXIAN exhaustion-proofs of EUCLID: *Elements*: Book 12): if we are given any two convex curves ACB, $AC'B$ with the same two end-points A, B (where the curves may, in general, include any number of line-segments) such that one, say $AC'B$, is entirely contained within the other, $AC'B$, then the area $(ACBA)$ is greater than the area $(AC'BA)$. A second set of lemmas, far more subtle, were first stated by ARCHIMEDES[3] and these introduce the application to the rectification of arcs of convex curves: thus, given

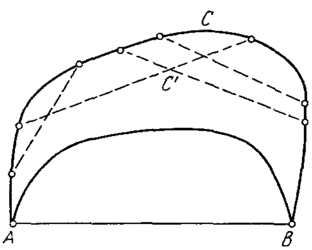


Fig. 76

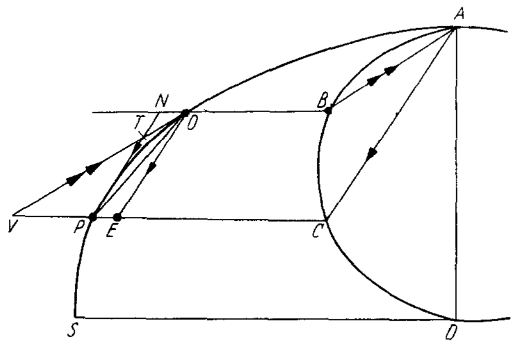


Fig. 77

the same two convex curves ACB, $AC'B$, arc-length $(\widehat{ACB})>$arc-length $(\widehat{AC'B})$ $(>$line-length $(AB))$. Analogous sets of lemmas, likewise first stated by ARCHIMEDES,[4] introduce the applications to cubature of convex volumes and rectification of convex surfaces.

WREN'S (1658) rectification of the general cycloid arc[5] is a pretty example of how the adaptation is made in a particular case. The proof depends on the standard result that the tangent to the cycloid arc APS at P is parallel to the chord AC of the generating circle $ACDA$ which joins the cycloid vertex A to C, the meet of the circle with PC drawn parallel to the cycloid base SD. Thus, considering the cycloid $AOPSD$, assumed (and provably) everywhere convex upward in the interval-arc $A\widehat{OP}S$, where $ABCD$ is the generating circle with OB, PC parallel to the cycloid base SD, the tangents PN at P, VO at O (which meet in T) are parallel to AB, AC respectively. The application of the ARCHIMEDEAN lemmas yields the result that $OV>\text{arc}\,\widehat{OP}>PN$.[6] For, taking OE parallel to PN (meeting PC in E), it follows immediately, since the cycloid is convex upwards, that the slope at P is greater than at O (where O is taken farther away from the base SD than P), so that $\widehat{NPC}>\widehat{OVE}$ and $VO=VT+TO>PT+$

* The convexity condition is defined: for any two points of the curve all points which lie on the line joining them lie within (or on) the curve.

[3] In the preface to *Sphere and cylinder*: Book *1*—compare DIJKSTERHUIS, *op. cit.*: 145—149.

[4] In the same preface to *Sphere and cylinder*: Book *1*.

[5] In JOHN WALLIS: *tractatus duo de cycloide, ... de cissoide ...*, Oxford 1659: 62—74.

[6] See *tractatus ...*: 62—63.

TO; and $PN=OE<OP$, and the result follows since by ARCHIMEDES' lemma the convex (line-segment) curve PTO has the same endpoints P, O and encloses the convex arc $\widehat{PO}$, so that $\widehat{PTO}$ $(=PT+TO)>\widehat{PO}>$ base-line PO.

Now set up the following division of the cycloid $I\widehat{S_k S_0}2_0 I$: for any 1_0 in $2_0 I$ (with $1_0 I>2_0 I$) define the sequence of section points k_0, $k=1, 2, 3, \ldots$, in $I2_0 1_0$ such that $I_{(k+1)0}:Ik_0=I2_0:I1_0$ (or the lines $I1_0$, $I2_0$, $I3_0$, $I4_0, \ldots$ are in decreasing continued proportion); and take parallels $S_{k-1}(2i)_0$ to the base through alternate section points $4_0, 6_0, 8_0, \ldots$ cutting the cycloid arc $I\widehat{S_k S_0}$ in the points S_{k-1}; further, at each S_k draw the cycloid tangent $S_k A_k$ (meeting $S_{k-1}(2k)_0$ in A_k) and

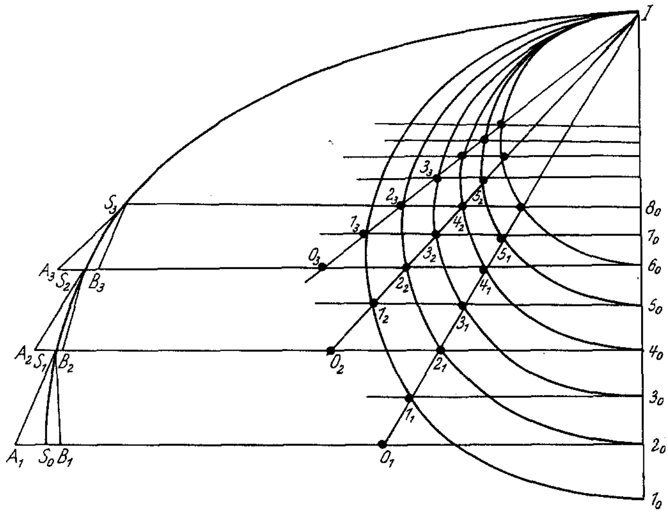


Fig. 78

$S_k B_k$ parallel to the tangent at S_{k-1} (meeting $S_{k-1}(2k)_0$ in B_k); and, finally, draw the semicircles on diameters Ik_0 (as in the figure), and the lines $I0_k$ (which meet the semicircle $I\widehat{1_k 1_0}$ in 1_k and the parallels $S_{k-1}(2i)_0$ in 0_k) such that, for each k successively, $0_k 1_k=1_{k-1}2_{k-1}$, $k=1, 2, 3, \ldots$. Then where the semicircle $I\widehat{k_j k_0}$ meets the line $I0_j$ in k_j we easily show that the sequence $I0_j$, $I1_j$, $I2_j$, $I3_j, \ldots$ is in decreasing geometrical proportion (and with the same ratio as that of the sequence $I1_0$, $I2_0$, $I3_0, \ldots$) and it follows that the line-segments $k_l(k-1)_l$, $(k-1)_{(l+1)}k_{(l+1)}$ are equal.

Further, the property of the cycloid tangent that $S_k A_k$ tangent at S_k is parallel to the chord $I0_k$* allows us, since for each cycloid arc $\widehat{S_{(k-1)}S_k}$, $S_k A_k>\widehat{S_{(k-1)}S_k}>S_k B_k$, to set up the inequality

$$\sum_{1\leq k\leq n}(S_k A_k)>\sum_{1\leq k\leq n}(\widehat{S_{k-1}S_k})\,[=\widehat{S_n S_0}]>\sum_{1\leq k\leq n}(S_k B_k),$$

or

$$\sum_{1\leq k\leq n}\big(k_0(k+1)_0+(k+1)_0(k+2)_0\big)>\widehat{S_n S_0}>$$
$$>\sum_{1\leq k\leq n}\big((k+1)_0(k+2)_0+(k+2)_0(k+3)_0\big),$$

* So that $S_k A_k=0_k 2_k=1_{(k-1)}3_{(k-1)}=\cdots=k_0(k+2)_0=k_0(k+1)_0+(k+1)_0(k+2)_0$ and similarly $S_k B_k=2_{(k-1)}4_{(k-1)}=(k+1)_0(k+3)_0=(k+1)_0(k+2)_0+(k+2)_0(k+3)_0$.

and, in particular, since $S_n \to I$ as n becomes infinite,

$$\lim_{n\to\infty}\Big(\sum_{1\leq k\leq n}(S_k A_k)\Big)\,[=1_0 I + 2_0 I] > \widehat{IS}_0 > \lim_{n\to\infty}\Big(\sum_{1\leq k\leq n}(S_k B_k)\Big)\,[=2_0 I + 3_0 I],$$

so that

$$1_0 I + 2_0 I\,(=2\times 2_0 I + 1_0 2_0) > \widehat{IS}_0 > 2_0 I + 3_0 I\,(=2\times 2_0 I - 2_0 3_0).$$

We can now satisfy the conditions *1* to *5* of the ARCHIMEDEAN exhaustion-model by taking

$$A_i \equiv \lim_{n\to\infty}\Big(\sum_{1\leq k\leq n}(S_k A_k)\Big) > \alpha \equiv \widehat{IS}_0 > a_i \equiv \lim_{n\to\infty}\Big(\sum_{1\leq k\leq n}(S_k B_k)\Big);$$

and $B_i \equiv 2\times I 2_0 + 1_0 2_0 > \beta \equiv 2\times I 2_0 > b_i \equiv 2\times I 2_0 - 2_0 3_0$, where the value of $1_0 2_0$ is arbitrary (but positive)*; so that, finally, cycloid arc $\widehat{IS}_0 = 2\times I 2_0$ (and by an obvious extension the general cycloid arc $\widehat{IS}_i = 2\times I 2_i,\ = 2\times I(i+2)_0$).

The standard ARCHIMEDEAN exhaustion-model received several extensions in the 17th century, and generalized models of proof (given in full for each particular application) were widely used in the period 1640–1670 before algorithmic calculus methods were developed which were apparently simpler and easier to handle if less rigorously based.

Thus, an immediate extension generalizes conditions *2* and *4* by:

2'. $$(i,j)\left(j > i \cdot \to \cdot \begin{bmatrix} A_i \geqq A_j;\ B_i \geqq B_j \\ a_i \leqq a_j;\ b_i \leqq b_j \end{bmatrix}\right).$$

4'. *a.* $(i)\,(a_i \leqq b_i)$; *b.* $(i)\,(A_i \geqq B_i)$.

Again, there is no unique proof-form, but a widely used approach[7] generalizes the ARCHIMEDEAN proof, twisting inequalities to show that $|\alpha-\beta|>0$ is impossible. Splitting the proof into two halves, as before, we have:

Case 1: $\alpha>\beta$. (By *1, 4' a*) $|\alpha-\beta| = (\alpha-\beta) < (A_i - b_i) \leqq (A_i - a_i)$,

and

Case 2: $\alpha<\beta$. (By *1, 4' a*) $|\alpha-\beta| = (\beta-\alpha) < (B_i - a_i) \leqq (A_i - a_i)$;

so that in either case $|\alpha-\beta| \leqq (A_i - a_i)$, which, however, (by *3*) for sufficiently large N we can make as small as we wish, and in particular less than λ. Immediately there arises the contradictory $(A_i - a_i) < \lambda = |\alpha-\beta| < (A_i - a_i)$.

A second form of proof is JAMES GREGORY's favourite "igitur quatuor quantitates" approach[8] which proceeds—roughly as before—to show that $(A_i - a_i) <$

---

* Condition *1* is immediate, and condition *4* ($A_i = B_i,\ a_i = b_i$) is proved above; condition *5* is immediate since curve-length can always be given positive measure; while $|A_i - a_i| = |B_i - b_i| = 1_0 2_0 + 2_0 3_0 = 1_0 3_0$, which by taking the arbitrary length $1_0 2_0$ small enough can likewise be made as small as desired; lastly, by considering two values for $1_0 2_0$, λ and λ', where $\lambda > \lambda'$, we easily show that, for $1_0 2_0$ decreasing, $A_i = B_i$ both increase, $a_i = b_i$ both decrease (which satisfies condition *2*).

[7] To be found, for example, in GREGORY ST. VINCENT's *opus geometricum*, Antwerp, 1647; *passim*.

[8] In his exhaustion proofs of *GPU* and *EG*.

$|\alpha-\beta|$ for suitably large i but finishes with the characteristic twist:

$$\text{(by } 1,\ 4'\,a,\ 4'\,b) \qquad (i)\begin{pmatrix} A_i \geqq B_i \cdot + \cdot B_i > \beta \cdot \rightarrow \cdot A_i > \beta \\ a_i \leqq b_i \cdot + \cdot b_i < \beta \cdot \rightarrow \cdot a_i < \beta \end{pmatrix},$$

so that, for all i, $A_i > \begin{bmatrix}\alpha\\ \beta\end{bmatrix} > a_i$, which involves the contradictory conclusion that $(A_i - a_i) > |\alpha - \beta|$ (true for all i).

In its application this generalized exhaustion proof-form is much more powerful than the simple ARCHIMEDEAN model in that it allows the use of transformations of area and curve-length in a far more general way. Specifically, considering the application to quadrature of the area under convex curves (which has an analogous treatment in the case of cubatures of convex volumes), we can give a general sketch which covers a wide variety of 17th century theorems on area-equivalences.[9]

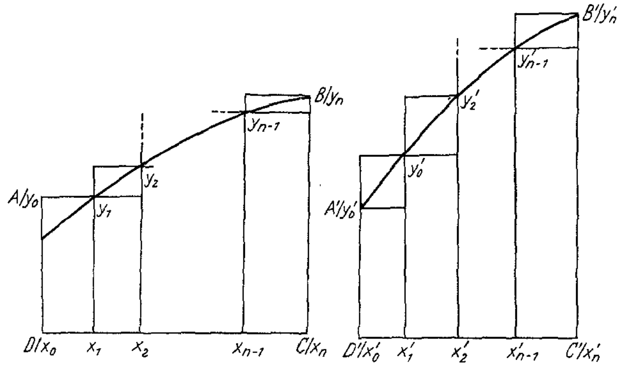


Fig. 79

Set up, then, two portions of convex curves* —for simplicity, we may take them both convex up—contained between ordinates AD, BC: $A'D'$, $B'C'$ perpendicular to the bases DC, $D'C'$ respectively. We can then define some n-division of the line-interval DC by the $(n-1)$ points $x_1, x_2, \ldots, x_{n-1}$ (where $x_0 = D$, $x_n = C$ and the x_i are ordered such that $x_{i+1} > x_i$ with respect to a real-number measure in the line DC), and a corresponding n-division of $D'C'$ by the points $x'_2, x'_3, \ldots, x'_{n-1}$ (where $x'_0 = D'$, $x'_n = C'$ and $x'_{i+1} > x'_i$), such that to every x_i there is a unique corresponding point x'_i and conversely (with the respective orderings preserved). Raising corresponding ordinates $x_i y_i$, $x'_i y'_i$ on the section points x_i, x'_i for each i, we can sketch in restrictions of a completely general type which are sufficient, using the extended exhaustion-model, to show the two areas $ABCD$, $A'B'C'D'$ equal.

Clearly we can satisfy condition *1* in the form

$$(n)\begin{pmatrix} A_n > \alpha > a_n \\ B_n > \beta > b_n \end{pmatrix}$$

by taking

$$A_n \equiv \sum_{0\leqq\lambda\leqq n-1} (\square\, x_\lambda\, y_{(\lambda+1)}), \qquad B_n \equiv \sum_{0\leqq\lambda\leqq n-1} (\square\, x'_\lambda\, y'_{(\lambda+1)});$$

$$\alpha_n \equiv \text{area } (ABCD), \qquad \beta \equiv \text{area } (A'B'C'D');$$

and

$$a_n \equiv \sum_{0\leqq\lambda\leqq n-1} (\square\, y_\lambda\, x_{(\lambda+1)}), \qquad b_n \equiv \sum_{0\leqq\lambda\leqq n-1} (\square\, y'_\lambda\, x'_{(\lambda+1)}).^{**}$$

---

* Equivalently, curves with monotonically increasing or decreasing slopes.

** These A_n, B_n; a_n, b_n are regularly called in the 17th century circumscribed and inscribed "mixtilinea".

[9] In particular much of the latter part of BARROW'S *LG*, and JAMES GREGORY'S *GPU*: props. 1—11: 1—29.

Condition *2'* is also immediate (if a little subtler) if we consider any further j-division of DC by points $x_{(n+\lambda)}$, $\lambda = 1, 2, 3, \ldots, j$ (no one of which coincides with any of the x_i, $i = 0, 1, 2, \ldots, n$), and a corresponding j-division of $D'C'$ by points $x'_{(n+\lambda)}$ (which preserves the ordering of x_i, x'_i, $i = 0, 1, 2, \ldots, n, n+1, \ldots, n+j$, in the extended interval sectioning): in particular any $x'_{(n+\lambda)}$ must come in some interval $[x_i, x_{(i+1)}]$ between a pair of adjacent points x_i, $x_{(i+1)}$ (with the corresponding $x'_{(n+\lambda)}$ in the interval $[x_i, x_{(i+1)}]$ between x_i, $x_{i(+1)}$); so that

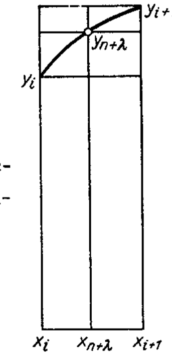


Fig. 80

$$\Box x_i \, y_{(i+1)} \geqq (\Box x_i \, y_{(n+\lambda)} + \Box x_{(n+\lambda)} \, y_{(i+1)}),$$

and

$$\Box x_{(i+1)} \, y_i \leqq (\Box x_{(n+\lambda)} \, y_i + \Box x_{(i+1)} \, y_{(n+\lambda)}),$$

with corresponding inequalities for $\Box x'_i \, y'_{(i+1)}$ and $\Box x'_{(i+1)} \, y'_i$ respectively—a general argument for all points in the extended sectioning of the interval, and so

$$\sum_{0 \leqq i \leqq n-1} (\Box x_i \, y_{(i+1)}) \geqq \sum_{0 \leqq i \leqq n+j-1} (\Box x_i \, y_{(i+1)})$$

and

$$\sum_{0 \leqq i \leqq n-1} (\Box x_{(i+1)} \, y_i) \leqq \sum_{0 \leqq i \leqq n+j-1} (\Box x_{(i+1)} \, y_i),$$

where x_i is the i^{th} point in the point-set $x_0, x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+j}$ ordered according to their position in the interval x_0, x_n.

Condition *3* is closely connected with the nature of the correspondence $x_i \leftrightarrow x'_i$. Usually we find it met in the 17^{th} century by defining corresponding n-sections such that the line-intervals $[x_i, x_{i+1}]$ are equal (which is the basis of all indivisible integration theory) or in geometrical progression[10]; but the sufficient condition is that in the correspondence for sufficiently large i we can make each interval-pair $[x_i, x_{(i+1)}]$, $[x'_i, x'_{(i+1)}]$ unlimitedly narrow, since

$$(A_n - a_n) = \sum_{0 \leqq \lambda \leqq n-1} (\Box x_\lambda \, y_{\lambda+1} - \Box y_\lambda \, x_{\lambda+1})$$
$$= \sum_{0 \leqq \lambda \leqq n-1} (\Box y_\lambda \, y_{\lambda+1}),$$

which is less than or equal to $x_n y_n \times \text{Max}\,(x_\lambda x_{\lambda+1})$, and this likewise may be made as small as desired. Finally, since condition *5* is immediate (by definition we give positive measure to all areas), it remains to satisfy condition *4'*. On the whole 17^{th} century geometers find this the hardest restriction but the most fruitful, since the comparison between the A_n, a_n; B_n, b_n usually contains within it the germ of the result which the exhaustion proof justifies. Regularly it is met by restricting corresponding n-sections such that, for all i,

$$\Box x_i \, y_{i+1} \geqq \Box x'_1 \, y'_{i+1}, \quad \Box x_{i+1} \, y_i \leqq \Box x'_{i+1} \, y'_i,$$

but the condition, while clearly sufficient, is not necessary.

---

[10] Both FERMAT (see *Oeuvres* **1**: 255—288, 1644 tract *de aequationum localium transmutatione ... in quadrandis infinitis parabolis et hyperbolis usus*) and TORRICELLI (in the unpublished tracts of the early 1640's printed in *opere 1* **1**: 227—274, 275—328) use this "continued proportion" section in squaring the general parabolas and hyperbolas, while, as we have seen above, it is fundamental in WREN's cycloid-arc rectification.

Such abstract considerations are, however, not found in 17^{th} century mathematics, and it will both illuminate our treatment and give a truer perspective if we consider an example in many ways typical of the actual use of the exhaustion form—prop. 11 of GREGORY'S *GPU*.[11] Here the comparison condition $(4'a, b)$ is satisfied by a lemma[12]: given any convex arc $\widehat{A\lambda D}$ and any line BN with the tangents at A, D (meeting in C) and the chord BC cutting BN in N, G, H respectively, then, where AB, DE are any suitably chosen (parallel) ordinates to the curve with CF, OG, KH, QN taken parallel to them (meeting as shown), trapezium $(ADEB) >$ mixtilineum $(A\lambda DLO)$ and rectilineum $(ACDEB) <$ mixtilineum $(A\lambda DSQ)$.* Now consider the convex curve $(BCGI)$ with abscissa IA and ordinates CE, GH, ... parallel to AB, and define a second curve $(IPSY)$ (easily

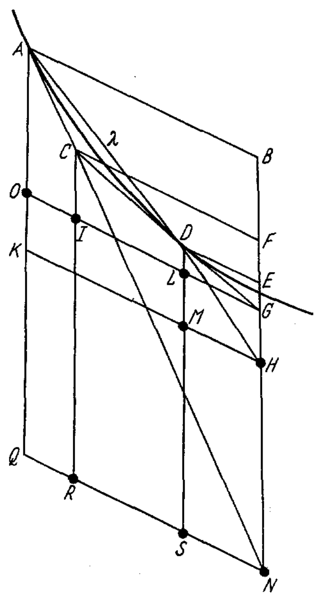


Fig. 81

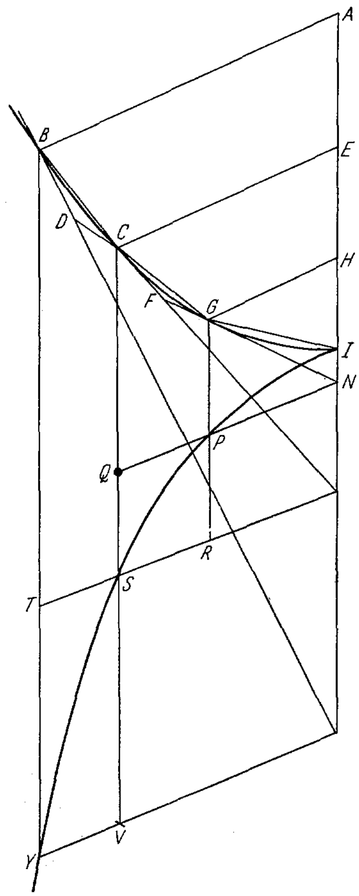


Fig. 82

* Clearly trapezium $(ADEB) =$ trapezium $(ADMK)$, which is greater than trapezium $(ADLO) >$ mixtilineum $(A\lambda DLO)$, since H (by the curve convexity) is between N and G: while trapezium $(ABFC) =$ trapezium $(ACRQ)$, and trapezium $(CFED) =$ trapezium $(CDLI) <$ trapezium $(CDSR)$, with rectilineum $(ACDSQ) <$ mixtilineum $(A\lambda DSQ)$.

[11] This is, in fact, a generalized form of a problem given by ROBERVAL to TORRICELLI, and proved with an exhaustion-method in TORRICELLI'S letter to ROBERVAL of 7 July 1646—see E. TORRICELLI: *opere*, Firenze, 1919: **3**: 389—391, and compare 361 ff. and 377 (where he uses the transform to add yet one more proof of quadrature of the simple parabola to those of *de dimensione parabolae*, published in his *opera geometrica*, Florence, 1644).

[12] *GPU*: prop. 10: 25—27.

shown convex) such that, where the tangent at general point G on $BCGI$ meets AI is N, NP drawn parallel to AB meets GP, parallel to AI, in a point P of the curve: then corresponding area-segments of the one are equal to corresponding area-segments of the other, or area $(A\widehat{BCG}HA) = \text{area } (Y\widehat{BCG}\widehat{PSY})$.

To apply the exhaustion proof as prenex we need, as above, only justify an interpretation of conditions *1* to *5*. If, then, we have an n-section of the line-interval AH (which in an obvious way induces a corresponding n-section of the arc PSY)*, we can satisfy condition *1* by taking

A_n = circumscribing mixtilineum $(Y\widehat{BGR}SVY)$,
α = mixtilineum $(Y\widehat{BG}\,\widehat{PSY})$,
a_n = inscribing mixtilineum $(T\widehat{BG}\,PQST)$,
B_n = circumscribing rectilineum $(ABDFGH)$,
β = mixtilineum $(A\widehat{BGH})$,
b_n = inscribing rectilineum $(ABCGH)$.

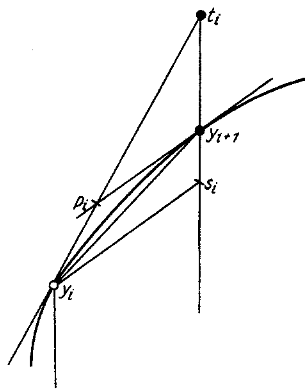


Fig. 83

Further GREGORY'S lemma shows condition *4'a*, *b*, since rectilineum $(ABDCE) <$ mixtilineum $(Y\widehat{BC}VY)$, trapezium $(ABCE) >$ mixtilineum $(T\widehat{BC}ST)$, and similarly for the other rectilinea and mixtilinea defined by the n-section: thus, rectilineum $(ECFGH) <$ mixtilineum $(S\widehat{CGR}S)$, trapezium $(ECGH) >$ mixtilineum $(Q\widehat{CG}\,PQ)$. Condition *3* is met by restricting the (arbitrary) n-section such that any interval between two adjacent section-points can be made as small as required (given a large enough number of section-points); condition *2'* follows by the curve convexity (as sketched above); while, as before, condition *5* is trivial. And so we have the proof.**

Application of the exhaustion-proof model to the general rectification problem with regard to convex arcs (and analogously to that of convex surfaces) is rather different[13]. Consider, then, the two convex arcs $\widehat{AB}$, $\widehat{A'B'}$ and, once again setting up corresponding n-sections of the base-intervals DC, $D'C'$, draw the tangents $y_i t_i$, $y'_i t'_i$ to the curves at y_i, y_i respectively, and $y_i s_i$, $y'_i s'_i$ parallel to $y_{i-1} t_{i-1}$, $y'_{i-1} t'_{i-1}$ respectively, where $y_i t_i$, $y_i s_i$; $y'_i t'_i$, $y'_i s'_i$ meet $y_{i+1} x_{i+1}$, $y'_{i+1} x'_{i+1}$ in t_i, s_i; t'_i, s'_i. Clearly this construction is modelled on the one invented by WREN in his cycloid-arc rectification, and the ARCHIMEDEAN convexity lemmas are modified in a similar fashion to give the inequality $y_i t_i > \text{arc}\, \widehat{y_i y_{i+1}} > y_i s_i$

* Only one pair of corresponding points E, S apart from the end-points A, N; P, Y is shown in the figure.

** GREGORY uses his "igitur quatuor quantitates" form of the exhaustion-proof.

[13] Such a general approach to rectification is, I think, original with FERMAT in his *de linearum curvarum cum lineis rectis comparatione dissertatio geometrica* (printed in appendix to LALOVERA'S *tractatus de cycloide* ..., Tolosae, 1660): prop. 2 (·≡· *Oeuvres* (ed. P. TANNERY & CH. HENRY); 1 (Paris, 1891): 211 ff.), though, clearly, FERMAT is generalizing the method used by WREN in his cycloid rectification. JAMES GREGORY (who quotes the LALOVERA work in the preface to *VCHQ*) virtually repeats the FERMAT exposition, though he gives a fuller discussion of the various cases of convexity, in *GPU*: prop. 1: 1–3: "sit curva quaecunque ... simplex et non sinuosa".

(with, correspondingly, $y'_i t'_i > \text{arc}\,\widehat{y'_i y'_{i+1}} > y'_i s'_i$). Thus, as before, take the tangents at y_i, y_{i+1} meeting in p_i; then, since the slope from y_i to y_{i+1} continuously decreases (or $\widehat{p_i y_{i+1} t_i}$ is obtuse), $\widehat{p_i t_i y_{i+1}} < \widehat{p_i y_{i+1} t_i}$, or $p_i y_{i+1} < p_i t_i$, so that $y_i t_i (= y_i p_i + p_i t_i) > y_i p_i + p_i y_{i+1}$, $>$ arc $\widehat{y_i y_{i+1}}$ (ARCHIMEDES); and similarly, since $y_i s_i$ is (by the convexity condition) contained within the curve and $\widehat{y_i s_i y_{i+1}}$ $(= \widehat{p_i y_{i+1} t_i})$ is obtuse, $y_i s_i < y_i y_{i+1}$, $< \text{arc}\,\widehat{y_i y_{i+1}}$ (by ARCHIMEDES' lemma).

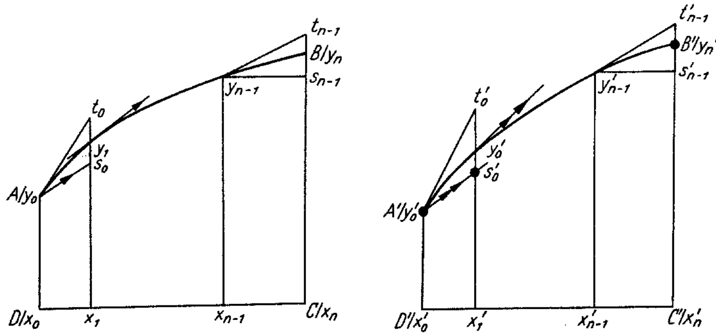


Fig. 84

We can now satisfy condition *1* of the extended exhaustion-proof by defining

$$A_n = \sum_{0 \leqq i \leqq n-1} (y_i t_i), \qquad B_n = \sum_{0 \leqq i \leqq n-1} (y'_i t'_i),$$

$$\alpha = \sum_{0 \leqq i \leqq n-1} (\widehat{y_i y_{i+1}}) = AB, \qquad \beta = \sum_{0 \leqq i \leqq n-1} (\widehat{y'_i y'_{i+1}}) = A'B',$$

$$a_n = \sum_{0 \leqq i \leqq n-1} (y_i s_i), \qquad b_n = \sum_{0 \leqq i \leqq n-1} (y'_i s'_i).$$

Further, since we always give positive measure to arc-length, condition *5* is trivial, and condition *3* is met by taking a suitable n-section of the base-intervals CD, $C'D'$. Condition *2'* will, however, in general be rather difficult to prove; and the whole power of the exhaustion-method (in showing $AB = A'B'$) will lie in specifying particular comparison techniques under which condition *4'* is satisfied—that is, by which we have for all i,

$$y_i t_i \geqq y'_i t'_i \quad \text{and} \quad y_i s_i \leqq y'_i s'_i.$$

In fact, the rectification application of the exhaustion-model is only very rarely made[14], and these mostly use the ARCHIMEDEAN model. Rectifications were in practice carried out in a more straightforward form by using a differential triangle approach[15], and when an exhaustion-proof was used it is often equivalent to a geometrical transform of tangent-length into curve-area and, as such, strictly

[14] Thus FERMAT in his *de linearum curvarum ... comparatione* ... gives only one specific example—that of the semicubical parabola (props. 3, 4 ·≡· *Oeuvres* **1**: 217 to 227), which virtually tightens up the HEURAET proof of its rectification (given a little later than but independently of NEIL's), while JAMES GREGORY is only a little more expansive, though prop. 58 of his *GPU*: 107—109 gives a general rectification procedure for the general parabolas (hinted at by FERMAT—see *Oeuvres* **1**: 227 ff.).

[15] Compare NEIL's rectification of the semicubical parabola (see previous chapter).

analogous to a differential triangle method.[16] FERMAT, however, has a neat example[17] of a curve-equivalence test which, while in fact using only the standard ARCHIMEDEAN exhaustion-model, has an obvious generalization which uses the full extended model. But perhaps the finest examples of such a rectification approach are TORRICELLI's (? 1646) rectification of the logarithmic spiral[18] —the first historical rectification of a (non-linear) curve—and PASCAL's proof[19] of the equivalence of the first revolution of an ARCHIMEDEAN spiral with the arc-length of a suitably defined parabola (which, however, uses a model slightly more general than even the extended ARCHIMEDEAN one[20]).

Application of the extended model was not restricted in the period to the comparison of line-intervals. While there are apparently no examples which, in other than a trivial way, compare two angle-intervals together, a very important part of 17^{th} century mathematics was devoted to the elaboration of what JAMES GREGORY in his definitive treatment[21] named the "involutio-evolutio" transform, which effectively sets up a correspondence between an angle-interval and a line-interval. Specifically, given fix-point A with emanating "radii" Al and the corresponding fix-line $\alpha_1\alpha_2$ with general "ordinate $\alpha\lambda$ (perpendicular to $\alpha_1\alpha_2$), the figures $A\widehat{L_1L_2}A$, $\alpha_1\widehat{\lambda_1\lambda_2}\alpha_2$ are defined to be in involute-evolute

[16] It is in this modified form, in fact, that both FERMAT's and GREGORY's treatments of note 14 are developed. In particular—and much as NEIL and HEURAET had done—FERMAT transforms, in his prop. 3, the tangent-lengths of a semicubical parabola into simple parabola-area.

[17] In the appendix to *de linearum curvarum ... comparatione ...* prop. 1 ·≡· *Oeuvres* **1**: 238—240.

[18] The full manuscript *(de infinitis spiralibus)* was published only in 1955 (at Pisa, edited by E. CARRUCCIO), though an incomplete form is given in *opere* **2** (1919): 349—399.

[19] See *Lettres de A. Dettonville*: tract *L'égalité entre les lignes spirale et parabolique, demonstrée a la manière des anciens.*

[20] What is new in PASCAL's proof-technique is his subtle use of the modulus form ("difference" ·≡· GREGORY's "differentia"). Briefly, PASCAL showed that where $(i)\begin{pmatrix}A_i>\alpha>a_i\\ B_i>\beta>b_i\end{pmatrix}$ and $(A_i-a_i)<Z$, the "differences" $|A_i-B_i|$, $|a_i-b_i|$ are both less than Z, where Z may be indefinitely small, and tried to show that $|\alpha-\beta|$ can be made indefinitely small. His proof, as given, contains a lacuna, but FERMAT (see CARCAVY's letter to HUYGENS of 22 September 1659) and, more naturally, HUYGENS soon filled it. As HUYGENS emends PASCAL's proof (in his letter to CARCAVY of 26 February 1660—see HUYGENS: *Oeuvres* **3**: 27ff.), $|B_i-A_i|<Z$, $|a_i-b_i|<Z$, $(A_i-a_i)<Z$ imply $|B_i-b_i|<3Z$ and, *a fortiori* $|\beta-b_i|<3Z$; again $(A_i-a_i)<Z$ implies $|\alpha-a_i|<Z$, and, since $|a_i-b_i|<Z$, $|\alpha-b_i|<2z$; so that $|\alpha-\beta|<5Z$, which can be made indefinitely small.

[21] In *GPU*: props. 12—18: 29—41. The approach developed historically from the ARCHIMEDEAN proof that the area of the first revolution of the ARCHIMEDEAN spiral was half that of a suitably defined parabola (which—as TORRICELLI and ROBERVAL guessed and PASCAL proved rigorously—in fact has the same arc-length as that of the spiral); but in the 17^{th} century received an increasingly abstract and generalized treatment in the hands of the Italian CAVALIERI school—CAVALIERI himself, TORRICELLI and GREGORY's teacher STEFANO DEGLI ANGELI. Later BARROW in his *LG*: Books 8ff. widely uses the two forms, involuted and evoluted, stating a wide variety of theorems in dual form. An interesting modern account of GREGORY's systematisation is that of A. PRAG in his *On James Gregory's "geometriae pars universalis"*. ≡ GREGORY (*TV*): 487—505, especially 493—497.

correspondence if corresponding points l, λ in the respective arcs $\overset{\frown}{L_1 L_2}$, $\overset{\frown}{\lambda_1 \lambda_2}$ (with $L_1 \leftrightarrow \lambda_1$, $L_2 \leftrightarrow \lambda_2$) are such that $Al = c\lambda$ and $\overset{\frown}{L_1 l} = \overset{\frown}{\lambda_1 \lambda}$ (or equivalently $\overset{\frown}{l L_2} = \overset{\frown}{\lambda \lambda_2}$).* Giving some inequalities[22] GREGORY shows[23] that the area of the evolute figure is twice that of the involute figure, using inscribed and circumscribed mixtilinea to establish the inequalities on which he can apply his "igitur quatuor quantitates" exhaustion form; more outstandingly, he proves[24] the conformality of the transform—a property used by WREN[25] in his treatment of the logarithmic spiral and general contracted and protracted cycloids (denoting the transform as a "convolution").

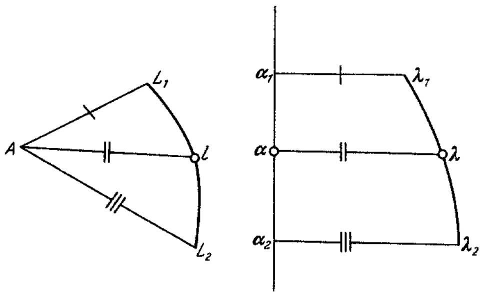


Fig. 85

A not unsimilar line-angle comparison is the "coordinate" transform which WALLIS develops in proof of a lemma basic to his quadrature of the general cissoid segment[26]. Consider the semi-circle $\overset{\frown}{AD\bar{A}A}$ and the angle-interval $\widehat{A\bar{A}D}$, together with the "coordinate" rectangle $A'\bar{A}'\alpha\bar{A}_1$ and line-interval $A'\bar{A}'$,

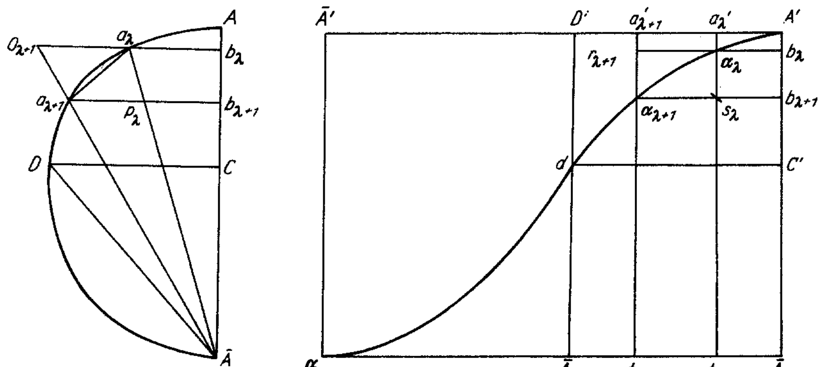


Fig. 86

where, for all points D in arc $\overset{\frown}{A\bar{A}}$, D' in $A'\bar{A}'$, arc $\overset{\frown}{AD} = A'D'$ (and in particular $\overset{\frown}{A\bar{A}} = A'\bar{A}'$) and $A'\bar{A}_1 = A\bar{A}$: then, where C' in $A'\bar{A}_1$ corresponds to C in $A\bar{A}$

* GREGORY calls $A\overset{\frown}{L_1 L_2}A$ the "involuta" of $\alpha_1 \overset{\frown}{\lambda_1 \lambda_2} \alpha_2$, and $\alpha_1 \overset{\frown}{\lambda_1 \lambda_2} \alpha_2$ the "evoluta" of $A\overset{\frown}{L_1 L_2}A$; the points l, λ "mutuo relativa puncta"; fix-point A and angle $\widehat{L_1 A L_2}$ "centre" and "angle" of involution; fix-line $\alpha_1 \alpha_2$ the "evolute axis".

22 *GPU*: props. 12, 13.

23 *GPU*: props. 15, 16.

24 *GPU*: props. 17, 18.

25 WALLIS: *tractatus ... de cycloide ...*, Oxford, 1659: 69—72, especially 70—71 (on contracted and protracted cycloids); 104—108 (on convoluted triangle, *sc.* logarithmic spirals, and convoluted pyramids).

26 In his *mechanica, sive de motu ...*, London, 1670: Book 2: prop. 17A, *figura sinuum versorum ... est semicirculi correspondentis dupla, et partes partium (respective sumptarum) duplae.* Compare previous chapter.

by $A'C' = AC$, define the curve $A'd\alpha$ as the point-set of d, the meet of the normal to $A'\bar{A}'$ at D' with the normal at C' to $A'\bar{A}_1$. WALLIS' proof shows that $2\times$ area $(\widehat{AD}\bar{A}A) = \text{area}\,(\widehat{A'd}\bar{A}_2\bar{A}_1A')$* in a typical application of the extended exhaustion-model: for, taking two corresponding n-sections of $\widehat{AD}$, $A'D'$ by $Aa_\lambda = A'A'_\lambda$ and defining o_λ, p_λ, r_λ, s_λ, t_λ as the respective meets of $\bar{A}a_\lambda$, $a_{\lambda-1}b_{\lambda-1}$; $\bar{A}a_\lambda$, $a_{\lambda+1}b_{\lambda+1}$; $a'_\lambda\alpha_\lambda$, $\alpha_{\lambda-1}b'_{\lambda-1}$; $a'_\lambda\alpha_\lambda$, $\alpha_{\lambda+1}b'_{\lambda+1}$; and $a'_\lambda\alpha_\lambda$, $\alpha\bar{A}_1$ (where α_λ in arc $\widehat{A'd\alpha}$ and b'_λ in $A'\bar{A}_1$ correspond to a'_λ in $A'A'$), we can use the ARCHIMEDEAN convexity lemmas[27] to show

$$o_{\lambda+1}a_\lambda > \widehat{a_{\lambda+1}a_\lambda} > (a_{\lambda+1}a_\lambda >)\, a_{\lambda+1}p_\lambda;$$

so that, for each λ,

$$2\times\Delta\, o_{\lambda+1}a_\lambda\bar{A}\,(= o_{\lambda+1}a_\lambda \times b_\lambda\bar{A}) > \widehat{a_{\lambda+1}a_\lambda}\,(= a'_{\lambda+1}a'_\lambda)\times b_\lambda\bar{A}\,(= b'_\lambda\bar{A}_1),$$

and

$$2\times\Delta\, a_{\lambda+1}p_\lambda\bar{A}\,(= a_{\lambda+1}p_\lambda\times b_{\lambda+1}\bar{A}) < \widehat{a_{\lambda+1}a_\lambda}\,(= a'_{\lambda+1}a'_\lambda)\times b_{\lambda+1}\bar{A}\,(= b'_{\lambda+1}\bar{A}_1)$$

or

$$2\times\Delta\, o_{\lambda+1}a_\lambda\bar{A} > \Box\, r_{\lambda+1}t_{\lambda+1}t_\lambda\alpha_\lambda, \quad\text{and}\quad 2\times\Delta\, a_{\lambda+1}p_\lambda\bar{A} < \Box\,\alpha_{\lambda+1}t_{\lambda+1}t_\lambda s_\lambda;$$

and finally we can satisfy the exhaustion-proof conditions in an obvious way by taking

$$A_n = \sum_{0\leqq\lambda\leqq n}(2\times\Delta\, o_{\lambda+1}a_\lambda\bar{A}), \qquad B_n = \sum_{0\leqq\lambda\leqq n}(\Box\, r_{\lambda+1}t_{\lambda+1}t_\lambda\alpha_\lambda),$$

$$\alpha = 2\times\text{area}\,(\widehat{AD}\bar{A}A), \qquad \beta = \text{area}\,(\widehat{A'd}\bar{A}_2\bar{A}_1),$$

$$a_n = \sum_{0\leqq\lambda\leqq n}(2\times\Delta\, a_{\lambda+1}p_\lambda\bar{A}), \qquad b_n = \sum_{0\leqq\lambda\leqq n}(\Box\,\alpha_{\lambda+1}t_{\lambda+1}t_\lambda s_\lambda).^{**}$$

On occasion even the extended exhaustion-proof model proved inadequate, and was further generalized. Thus, GREGORY in a straightforward quadrature theorem[28] finds that some revision is necessary, and introduces a proof which we can symbolize in the following way: Given the conditions

1. $(i)\begin{pmatrix}A_i > \alpha > a_i\\ B_i > \alpha > b_i\end{pmatrix}$

2. $(i,j)\,(j > i \cdot\rightarrow\cdot \begin{pmatrix}A_i \geqq A_j;\ B_i \geqq B_j\\ a_i \leqq a_j;\ b_i \leqq b_j\end{pmatrix}$

3''. For i, j sufficiently large and ε, ε' as small as we desire,

$$(EN)\,(i)\,(i > N \cdot\rightarrow\cdot (A_i - a_i) < \varepsilon),$$

and

$$(EN')\,(j)\,(j > N' \cdot\rightarrow\cdot (B_i - b_i) < \varepsilon'),$$

---

* Or, where $A\bar{A} = 2$, $\widehat{A\bar{A}D} = \vartheta$ (and so $A'D' = 2\vartheta$), $2(\vartheta + \frac{1}{2}\sin 2\vartheta) = \int_0^{2\vartheta}(1 + \cos x)\cdot dx$, where dx is the element of the arc AD.

** WALLIS, *in fact*, bound up in indivisible considerations, restricts the n-section unnecessarily to an equisection.

[27] See previous chapter.

[28] *GPU*: prop. 3.

4''. *a*. $(i, j)\,(A_i \geqq b_j)$; *b*. $(i, j)\,(B_i \geqq a_j)$,

5''. $(i)\,(a_i > 0,\, b_i > 0)$ (and so $\alpha, \beta > 0$, $(i)\,(A_i > 0,\, B_i > 0)$), then $\alpha = \beta$. In proof, he shows that $|\alpha - \beta| = \lambda > 0$ is impossible.

Case 1. $\alpha > \beta$. By *1.* $A_i > a$, $b_j < \beta$ and so

$$|\alpha - \beta| = (\alpha - \beta) < (A_i - b_j) = (A_i - a_i) + (B_j - b_j) - (B_j - a_i),$$

with (by 4''b) $B_j \geqq a_i$, so that $|\alpha - \beta| < (A_i - a_i) + (B_j - b_j)$.

Case 2. $\alpha < \beta$. By *1.* $B_j > \beta$, $a_i < \alpha$, and so

$$|\alpha - \beta| = (\beta - \alpha) < (A_i - a_i) + (B_j - b_j) - (A_i - b_j),$$

with (by 4''a) $A_i \geqq b_j$, so that, again, $|\alpha - \beta| < (A_i - a_i) + (B_j - b_j)$.

In either case, then, for sufficiently large i, j, $|\alpha - \beta|$ can be made less than $(\varepsilon + \varepsilon')$, and so as small as we wish, and in particular less than $\lambda = |\alpha - \beta|$ which proves contradiction (though GREGORY finishes with an "igitur quatuor quantitates" twist).

The far more general comparison techniques which can be introduced under the revised conditions *4''a, b* make this form extremely powerful—allowing, in particular, comparisons between convex curves and separating out particular cases according as both, one or neither are convex up[29] (though the GREGORY example is, apparently, unique*); but it is even more important to notice the tendency away from the logical trick of reversal of inequalities to the more fundamental concept hidden away in that *reductio ad absurdum*, that of two bounding sequences $[A_i]$, $[a_i]$ to α (with $(i)\,(A_i > \alpha > a_i)$) such that in the limit the magnitude difference $(A_i - a_i)$ $(= |A_i - a_i|)$ vanishes. The concept lies deep in the theory of convergent sequences (and, in particular, in the justification of the convergence of the CAUCHY-RIEMANN integral), and was introduced by PASCAL in his later geometrical work[30] with even less pretension to the logical device of reversing inequalities (which, inevitably I think, appears less convincing the more one ponders it).

Clearly, the 17th century exhaustion-proof is no simple thing but rather of the highest degree of complexity. In all applications, however, the convexity lemmas of area and length are basic in defining the circumscribed and inscribed mixtilinea which yield ever more sharpened bounds to the quantities compared. Of these, the convexity lemmas for curve-area seem obvious at an intuitive level —the concept of "area" has implicit in it the assumption that an area which contains wholly a second area be greater than it (without exception). In contrast,

* Nor do further extensions of the basic ARCHIMEDEAN proof-model seem to exist, though it is tempting to generalize condition *1* (to be covered by suitable comparison inequalities between the individual $A_i, A_j, \ldots$) to the n-set,

$$(i)\begin{pmatrix} {}_1A_i > {}_1\alpha > {}_1a_i \\ {}_2A_i > {}_2\alpha > {}_2a_i \\ \cdots \quad \cdots \quad \cdots \\ {}_nA_i > {}_n\alpha > {}_na_i \end{pmatrix}.$$

[29] Which, significantly, adds appreciably to the lengths of his proofs of *GPU*: props. 1, 2, in particular.

[30] See note 20.

the second (ARCHIMEDEAN) set of convexity lemmas on curve-length seem not at all obvious, and especially in ARCHIMEDES' own generalization of them to include line-lengths[31]. Further (and almost certainly through the copyists' incomprehension) the Greek and Latin texts as they existed in the late 16th century were full of misreadings and illucid alterations from the original, and we find even BARROW[32] in his standard university text of ARCHIMEDES (in a modernised form) admitting his inability to understand the significance of the convexity lemma 2 (that is, of two convex curves that which completely contains the other has the greater arc-length). Nor, significantly, have modern editors been more forthcoming in explanation, and in general—as HEATH and DIJKSTERHUIS, for example—are content merely to state the lemmas as axioms, devoting their attention to the more immediately attractive "ARCHIMEDES'" axiom which accompanies them (as axiom *5* in most editions). J. HJELMSLEV has, however, gone more deeply into the matter[33], emphasising that with ARCHIMEDES' lemmas essentially new magnitudes which go far beyond those envisaged in early Greek mathematics are introduced, and that, as a result, an extension of EUDOXUS' axioms which define equality between ratios had to be made to define the corresponding axioms of inequality[34]. Thus, to point some of the logical difficulties which can arise, HJELMSLEV considers a geometry which has a model in an algebraic PYTHAGOREAN 2-space of "points" $[p, q]$, where p, q are rational numbers and a distance function is defined by

$$D([p, q], [p', q']) = [(p - p')^2 + (q - q')^2]^{\frac{1}{2}}:$$

all rectilinea circumscribed and inscribed to the circle arc α say, A_i, a_i (where $A_i > \alpha > a_i$) are defined in the geometry, but not the arc-length α of the circle-arc itself (which is transcendental), and so we cannot assume uncritically the existence of the limit α even where *all* particular upper and lower bounds A_i, a_i can be shown to exist.

An obvious (and historical) way out of such difficulties is to introduce the concept of limit (upper and lower) bounds to the A_i, a_i, where $\lim (A_i) = \lim (a_i)$ are both defined and $|\lim (a_i) - \alpha|$ (or, equivalently, $|\lim (A_i) - \alpha|$) can be made less than any assignable finite quantity. We have then a concept basic to all modern treatments of curve-length (and, of course, of curve-area similarly), which define it as the (unique) upper bound to the set of the perimeter-lengths

---

[31] In *On the sphere and cylinder*, Book 1: *Lambanomena* $\cdot\equiv\cdot$ DIJKSTERHUIS (*op. cit.* note 1): 145.

[32] See his *Archimedis opera*: *Apollonii Pergaei conicorum libri iiii* ..., London 1675: 4: "hoc pronunciatum ab editoribus hactenus acceptum est pessime: in duo quippe discerpunt, unum veritate, alterum et sensu cassum, vide Rivaltum et stupe".

[33] See *Eudoxus' axiom und Archimedes' Lemma*, Centaurus **1** (1950—1951): 2—11; and *Über Archimedes' Größenlehre*, Det. Kgl. Danske. Videnskabernes Selskab. Matem.-Fysiske Meddelelser **25** (Kopenhagen, 1950): 4ff.

[34] In particular ARCHIMEDES uses $\frac{a}{b} > \frac{c}{d} \cdot\rightarrow\cdot \frac{a+b}{b} > \frac{c+d}{d}$ without proof. Proofs given by PAPPUS and EUTOCIUS assume the existence of a fourth proportional to b, a, d, though HJELMSLEV neatly avoids this by adapting the EUDOXIAN inequality definition $\frac{a}{b} > \frac{c}{d} \cdot\equiv\cdot (Em, n)$ $(ma > mb$ and $mc \leqq nd)$ (which has, equivalently, (Em, n) $(m(a+b) > (m+n)\,b$ and $m(c+d) \leqq (m+n)\,d)$.

of all inscribed rectilinea* (or, more rarely, as the lower bound of the set of perimeter-lengths of all circumscribed rectilinea). In fact, exactly this concept is implicit in the convexity lemmas which apply the exhaustion-proof to geometrical models.

Consider, for example, the ARCHIMEDEAN convexity lemma 2, that of two convex curves with the same end-points the outer-one has the greater perimeter-length (away from the line-segment joining the two end-points). Taking any convex arc-length AB where the (unique) tangents at A, B meet in D, we have by the lemma $(AD+DB) > \text{arc } AB > AB$. Alternatively, considering any point α_1, in the arc AB (where the tangent at α_1, meets AD, BD in a_1, b_1),

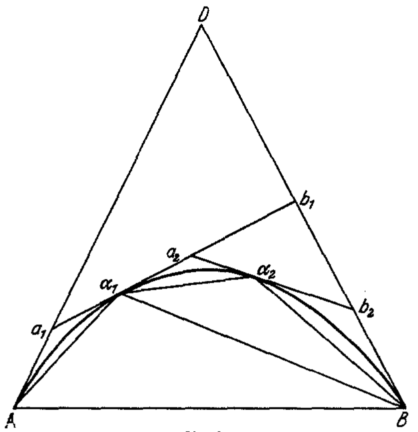


Fig. 87

$$(AD+DB) = (Aa_1 + a_1D) + + (Db_1 + b_1B) > (Aa_1 + a_1b_1 + b_1B)$$

(since $(a_1D+Db_i) > a_i b_i$), and $AB < A\alpha_1 + \alpha_1 B$; and, again, considering a second point α_2 in $\widehat{AB}$ (say in $\alpha_1 B$), where the tangent at α_2 meets $\alpha_1 b_1$, $b_1 B$ in a_2, b_2,

$$\alpha_1 b_1 + b_1 B = (\alpha_1 a_2 + a_2 b_1) + + (b b_2 + b_2 B) > \alpha_1 a_2 + a_2 b_2 + b_2 B$$

(since $(a_2b_1 + b_1b_2) > a_2b_2$), and similarly $\alpha_1\alpha_2 + \alpha_2 B \; \alpha_1 B$, so that $Aa_1 + a_1a_2 + a_2b_2 + b_2B < Aa_1 + a_1b_1 + b_1B < AD + DB$, and $A\alpha_1 + \alpha_1\alpha_2 + \alpha_2 B > A\alpha_1 + \alpha_1 B > AB$; and, in general, where $(A\alpha_1\alpha_2 \ldots \alpha_n B)$ is the ordering of the sequence of points (α_i) in $\widehat{AB}$, we can show that, for each n successively, the perimeter-length of the circumscribing rectilineum $(Aa_1a_2 \ldots a_n b_n B)$ continually decreases, while the perimeter-length of the inscribing rectilineum $(A\alpha_1\alpha_2 \ldots \alpha_n B)$ continually increases. However, at all stages

$$\text{perimeter-length } (Aa_1a_2 \ldots a_nb_nB) \left[= (A + a_1a_1\alpha_1) + (\alpha_1a_2 + a_2\alpha_2) + \cdots + (\alpha_nb_n + b_nB)\right]$$
$$> \text{perimeter-length } (A\alpha_1\alpha_2 \ldots \alpha_n B) \quad \left[= A\alpha_1 + \alpha_1\alpha_2 + \cdots + \alpha_n B\right].$$

By choosing a suitably dense set of points α_i (indefinitely close to both α_{i-1}, α_{i+1} for each i) in the arc AB we can, finally, make the difference between the perimeter-lengths of the circumscribing rectilinea (Rn) and inscribing rectilinea (rn) as small as we wish, and so we have the full CAUCHY definition of the common limit of two sequences (R_n), (r_n) (R_n, r_n monotonically decreasing, increasing respectively with increasing n) so defined that, for all n, $R_n > r_n$, with $(EN)\,(n)\,(n > N \cdot\!\rightarrow\!\cdot |R_n - r_n| < \text{arbitrary } \varepsilon)$. We now see that the ARCHIMEDEAN lemma assumes equivalently the concept of monotonic increase, decrease (in the convexity concept) and defines the curve-length—as in a modern exact treatment as the respective (unique) upper, lower bound of (R_n), (r_n).

These ideas, all implicit in ARCHIMEDES and the work of many of the greater 17th century mathematicians, contain necessary and sufficient conditions for

* The uniqueness is immediate "visually" in the case of the geometrical model, but analytical justification will have to show it in a more elaborate deductive way—for example, by considering chains of inscribing rectilinea, the upper bound of whose perimeter length we show unique, and then quantify the argument.

formulating the concept of definite integral on a rigorous analytical base in the restricted case where the function shall be convex in the integration interval* but the abstraction of logical form which was necessary to formulate these ideas was not, in fact, more than hinted at by those mathematicians who were masters of exhaustion-techniques—TORRICELLI, DESCARTES, FERMAT, MENGOLI, HUYGENS, PASCAL, ROBERVAL and (in England) JAMES GREGORY, BARROW, WREN and even NEWTON[35] were all unwilling to make the conceptual effort required to establish the general types of exhaustion proof which they used so readily as a logical prenex form with regard to which algorithmic forms could be worked out; and, indeed, while admitting its power and rigour, were in favour of suppressing it for the apparently simpler (if less rigorous) indivisible methods (especially in the CAVALIERI-theorem form, which lent itself to the development of generally applicable geometrical transforms).

The exhaustion proof-form in its many ramifications is the most rigorous deductive theory developed before the 19th century axiomatic developments, and far more so than the model 17th century mathematical theorists professed to admire: the proof-structure of EUCLID's *Elements*. Perhaps, indeed, the exhaustion-technique was viewed as rigorous less through an understanding of the method than because it was classically Greek, a "methodus veterum"—certainly an untoward amount of attention was given (and still is today) to the logical trick of *reductio ad absurdum* by which the conventional proof is rounded off (and ignoring the growing practice of substituting the idea of absolute magnitude of the "differentia", $|\alpha-\beta|$). Whether, however, because it was seen as essentially a classical theory (and so as something which should be improved on if modern mathematics was with any dignity to assert its independence) or because of a wrong idea of the range of the new analytical techniques of infinite series and differential algorithms, by 1670 the exhaustion-method was largely discarded. This rejection had very little basis in fact. TORRICELLI and those other mathematicians who used indivisible methods not only to produce a result but also to give it in a form of proof readily transformable into an exact exhaustion procedure had an intuition of the truth. But the weight of mathematical opinion was with HUYGENS[36]

* Or, rather more generally, to being continuous with unique tangents at every point in the integration-interval, since—by some such procedure as FERMAT's minimax method—we easily isolate inflexion points, and can then divide the curve into sections, each of which is convex (up or down, as the case may be).

[35] See *CUL Add.* 4000: 135Rff. (manuscript on "crooked lines").

[36] Compare the interesting manuscript passage (to be dated 1657) in HUYGENS *Oeuvres*, **14** (1920): 337: "... Sometimes by indivisibles. But they are deceived if they claim it for a proof, though to convince the knowledgeable it matters little whether a rigorous proof is given or just the basis of a proof whose sight resolves any doubt that a rigorous proof could be given. And yet I admit that in elaborating this ritualistic form of proof with clarity, consistency and the greatest possible precision great learning and native genius shine out, as in all ARCHIMEDES' works. But what matters first and above all is the process *(ratio)* of invention, and it is this which delights us especially and which we demand of the masters. It seems better, therefore, to follow this method which can more shortly and more clearly be understood and be exposed naked to the eye. Then indeed we spare ourselves the labour of writing it out and save others the toil of reading it, who will at length have no time to peruse the huge mass of geometrical findings ... if writers continue to use this prolix, though rigorous, method of the ancients."

who could see, in his middle years, the rigorous complexity of exhaustion-proof only as a hindrance to expression of the underlying thought, attacking the —in fact, eliminable—tiresome repetitions of the full proof form in each particular case, and not seeing that we need give only conditions which justify the application of the logical proof-model proved in general form once for all. It seems a pity that LEIBNIZ, perhaps the man above all others who had the logical power to create an abstract method, should have been distracted by having HUYGENS as his mathematical teacher in the 1670's.

X. Calculus

3. The concept of tangent

It is accepted fact that the general concepts of curve-tangent were ultimately subsumed, in one way or another (and with respect to several differing types of coordinate system) under the general theory of fluxional or differential calculus, and it is a plausible, if tentative, hypothesis that at a very early period the inverse nature of differential and integral techniques were suggested on this geometrical tangent model, the general integration problem being viewed as an "inverse method of tangents".[1] To a considerable extent the main outlines have been fairly conclusively drawn, but here yet once again historians tend to oversimplify the process as it crystalized into a symbolic differential technique, and in simplifying it introduced considerable distortion. While in the conventional account it is suggested that the tangent problem was solved analytically by DESCARTES and FERMAT in the 1630's, that contribution was, in fact, only one part of a much wider development whose extent is reflected in the immense 17th century literature which related to it.[2]. It will be illuminating, therefore, to discuss the particular methods invented to resolve the tangent-problem, and this will yield a truer perspective on the elegant general treatments which were later abstracted from the particularised methods of the mid-century.

Conceptually the most elementary (and yet till the 1670's the most subtle and widely applicable) of these tangent-methods were generalizations of the traditional Greek approach developed with respect to conics[3], which extend this synthetic method to treat of general smoothly continuous, convex curves[4]. Typically in the classical approach the tangent was defined as a line meeting a

[1] This idea was current at least as early as the late 1630's, occurring in DESCARTES' letter to DEBEAUNE of 20 February 1639. See P. TANNERY: *Pour l'histoire du problème inverse des tangentes* · ≡ · *Mémoires scientifiques* **6** (Paris, 1926): 457—477.

[2] There are, for example, hundreds of pages of NEWTON manuscript in the Portsmouth Collection which discuss tangents variously by geometrical, analytical and fluxional methods.

[3] Compare EUCLID: *Elements* **3**: prop. 16; APOLLONIUS: *Conics* **1**: props. 17, 32.

[4] Perhaps the first 17th century adaptation was WREN's derivation of the tangent at a general point on the cycloid arc (see note 6 below), but the generalized treatment with respect to general convex curve arcs was immediately given by FERMAT along with several analogous treatments of similarly defined curves (see his *de linearum curvarum cum lineis rectis comparatione dissertatio geometrica*, in appendix to LALOVERA's tract on the cycloid, Toulouse, 1660—and compare J. ITARD: *Fermat, précurseur du calcul différentiel*, Archives internationales d'histoire des sciences **1** (1947 to 1948): 589—610, especially 598—605). Further (increasingly rich) treatments are to be found in GREGORY's *GPU*: props. 6—9, and BARROW's *LG*: especially lectio **10**.

curve in a unique point, and a line was shown tangent by proving that it could not meet the curve again in a second point (at least, in a reasonably close interval of the curve arc). Thus, where $\widehat{AOB}$ is some conic-arc with OC the diameter conjugate to the ordinate ACB (so that $AC=CB$), the tangent at O is constructed by drawing the parallel to ACB through O and showing that, for $TT'C$ any parallel to OC (meeting the tangent at O in T, and $\widehat{AOB}$, $\widehat{ACB}$ in T', C'), $TC'>T'C'$ (except where $TT'C'$ coincides with OC): uniqueness of the tangent is proved by showing that any other line through O (not parallel to ACB) must meet the conic again in a second point S (distinct from O). In the 17th century generalization

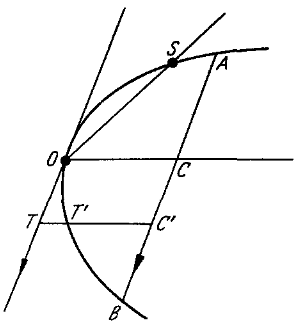


Fig. 88

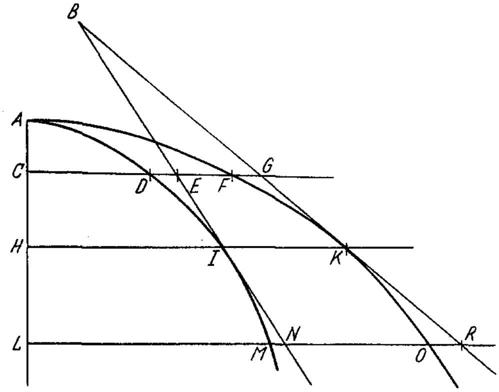


Fig. 89

an implicit condition of smooth continuity is made—which makes justification of tangent uniqueness trivial, since a continuous curve has a (unique) tangent at every non-singular point—and the general synthetic tangent-problem is reduced merely to showing the existence of a tangent by constructing a line which has a unique common point with the given curve. Further, to avoid such difficulties as inflexion-points, in this general treatment the curve-arcs are, for the most part, restricted to being convex.

A neat example which shows how this general idea is applied in a particular case is that of JAMES GREGORY's generalization[5] of WREN's proof[6] of a construction for the tangent at a general point on the cycloid arc: Given any (convex) curve $ADIM$ (with axis AL) and defining a second curve $AFKO$ such that, for any HIK at fixed angle AHI to AL (I in $\widehat{ADIM}$, K in $\widehat{AFKO}$), always $\widehat{AI}:IK=P:Q=$ constant, we construct the tangent BK at general point K on $\widehat{AFKO}$ from given tangent BI at I to $\widehat{ADIM}$ by showing that it passes through B on BI such that $BI=\widehat{AI}$. GREGORY's proof (which has for corollary that $\widehat{AFKO}$ is convex in

[5] GREGORY: *GPU*: prop. 8: 22—24. His prop. 9: 24—25 (repeated by BARROW: *LG*: lectio **10**; 5, 6: 76) tidies up a FERMAT generalization (in *de linearum curvarum ... comparatione ...*: prop. 6. · ≡ · *Oeuvres* **1**: 228—233) which considers the similar problem of constructing the tangent at general point k, where the curve $AFKR$ is defined from curve $\widehat{ADIM}$ by $\widehat{AI}:HK=P:Q$ for general parallel HIK. As we shall see, both are derivable in a simple way from motion considerations (and were probably first so found).

[6] In WALLIS' *tractatus de cycloide ..., de cissoide ...*, Oxford 1659: 63—64.

the same direction as $\widehat{ADIM}$) depends essentially on establishing the inequalities which show existence of a tangent. Thus, taking parallels above and below HIK ($CDEFG$, $LMNOR$ respectively), by ARCHIMEDES' convex curve-length lemma it follows that $IE < \widehat{ID}$ and $IN > \widehat{IM}$: so that

$$IK:(NR - IK) = IB:IN < IB\,(= \widehat{AI}):\widehat{IM} = IK:(MO - IK),$$

and

$$IK:(IK - EG) = IB:IE > IB\,(= \widehat{AI}):\widehat{DI} = IK;\ (IK - DF);$$

or $MO < NR$, $< MR$ and $DF < EG$, $< DG$, which shows that all points of BK (except K itself) lie outside the curve $\widehat{AFKO}$.*

Implicitly in this approach the assumption—to be justified in an immediate way by the smooth convexity of the curves considered—is made that in an arbitrarily small neighbourhood of the point of tangency the distance between corresponding points becomes indefinitely small. Clearly this assumption is equivalent to that of differential-triangle and limit-motion methods: that in this same small neighbourhood the element of tangent-length may be taken for the element of curve-length (in both magnitude and direction); and we find, in fact, that the rigorous Greek tangent-conception is widely introduced in 17th century treatments which try to give rigorous justification to the more intuitively-defined methods of the latter. Thus, where NEIL in his rectification of the general semi-cubical parabola had introduced a differential-triangle treatment on indivisible considerations, JAMES GREGORY justifies a generalized approach as follows:[7] where AO is a common axis, if from given (convex) curve $\widehat{BS}$ we define the curve $\widehat{AFLP}$ such that, for all IN normal to AO (with BR parallel to AO), $IL^2 = IN^2 - IM^2$, and a second curve $\widehat{AEKQ}$ such that $IK = \frac{\text{area }(\widehat{ALIA})}{IM}$, then, where C is taken in AI such that $IC:IK = IM:IL$, KC is tangent at K to $\widehat{AEK}$. In effect, what GREGORY has to prove is that ΔCIK is a differential triangle of $\widehat{AEK}$, and his treatment depends on showing that the triangle with sides IL, IM, IN is a second differential triangle—which follows easily since $IC^2 + IK^2 = CK^2$, and $IM^2 + IL^2 = IN^2$ with $IC:IM = IK:IL$ (so that $CK:CI = IN:IM$, the basis

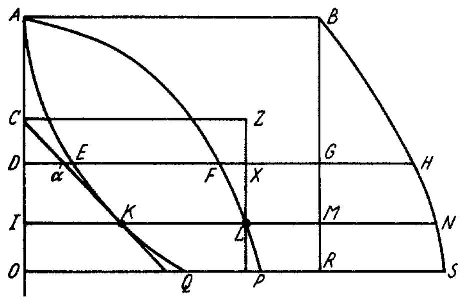


Fig. 90

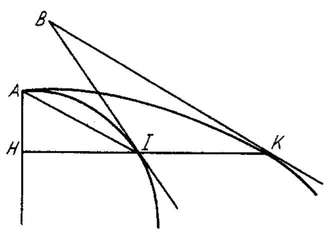


Fig. 91

* In the original WREN proof the curve $ADIM$ is a semicircle and $P = Q$ (or $\widehat{AI} = IK$, for all I). Clearly, since $IB = \widehat{AI}$, $\widehat{IBK} = \widehat{IKB}$; and, again, AI bisects $\widehat{BIH}$ $(= \widehat{IBK} + \widehat{IKB})$; so that $\widehat{AIH} = \widehat{BKI}$ and cycloid tangent BK is parallel to AI.

[7] See chapter eight. GREGORY's generalization is *GPU*: prop. 6, 17—19.

of NEIL'S differential-triangle transform which reduces elements of arc-length to corresponding elements of the axis AO, with finally area $(ABSO) = IM \times$ arc-length $\widehat{AQ}$). Thus it follows that $IK \times IM = \text{area}\,(\widehat{ALIA})$, $= IC \times IL =$ rectangle (IZ); and similarly, where $D\alpha EFGH$ is a general parallel to $IKLMN$ (meeting CK, $\widehat{AK}$ in α, E), $DE \times DG = \text{area}\ (\widehat{AFDA})$; so that, assuming $\widehat{AFL}$ is convex right (and therefore rectangle $(IX) > \text{area}\,(\widehat{FLID})$,

$$\text{rectangle}\ (IZ) : \text{rectangle}\ (DZ)\ (= IC : DC,\ = IK : D\alpha)$$

$$> \text{area}\ (ALIA) : \text{area}\,(\widehat{AFDA})\ (= IK : DE),$$

or $D\alpha < DE$ for all parallels $DEFGH$ above $IKLMN$ (with similar argument for the case of parallels taken below).

The complexities of GREGORY'S example rather confuse the basic outlines of the disguised limit-approach, and a neat proof of NEWTON[8] of the pole-polar relation in conics shows the method more clearly. Specifically, where DK is a general chord of a conic and AOB the conjugate diameter (whose meet C with DK therefore bisects it) through conic centre O, then the tangent at D meets AB in H such that $OB^2 = OH \times OC$. NEWTON'S proof draws general parallel $GFfe$ (meeting HD in F): then, since DH is tangent $Fe > fe$ (with equality only when F, f are at D), APOLLONIUS 3, 17 shows that $\dfrac{DC \times CK\,(= DC^2)}{BC \times CA} = \dfrac{De \times eK}{fe \times eG}$; so that when F, f pass into $D\left(\text{and so } \dfrac{De}{fe} = \dfrac{De}{Fe} \to \dfrac{DC}{CH}\right.$, while $eK \to DK$, $eG \to DG'$ or $\left.\dfrac{eK}{eG} \to \dfrac{DK}{DG'} = \dfrac{DC}{CO}\right)$, $\dfrac{DC^2}{BC \times CA} = \dfrac{DC^2}{CH \times CO}$; or $CH \times CO(= (OC - OH) \times OC) = BC \times CA = (OC - OB) \times (OC + OA\ (= OB)) = OC^2 - OB^2$.

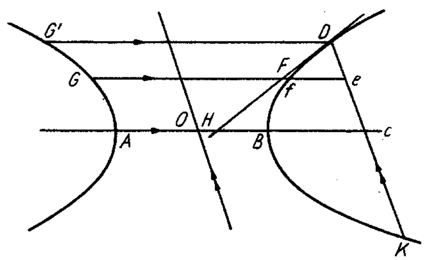


Fig. 92

Here the crucial point in the proof is the assumption that in the limit as the differential triangle $DFe\,(Dfe)$ becomes indefinitely small the element of the general parallel Fe intercepted between chord DK and tangent DH may be taken equal to the corresponding ordinate length fe; and similarly we might have used the equivalent limit-equality $DF = \widehat{Df}$ (which is the crucial part of NEIL'S rectification-method and GREGORY'S extension of it). Reformulating this slightly, we can see the limit-length of the curve $\widehat{Df}$ as having the same length and direction at the point D as the corresponding limit-length of the tangent DF (and so the same direction as the tangent DH), and this insight was to prove the basis for the more analytical investigations of the tangent concept which were to be built into the basis of the differential calculus, especially in its NEWTONIAN fluxional form.*

---

* The early manuscripts of NEWTON (especially *CUL Add.* 4004: *passim*) show that the fluxional calculus developed as a study in limit-motions.

[8] *CUL Add.* 3963: 107 R.

In particular, many of the results proved by developing inequalities in the Greek manner seem to have been suggested by such a viewpoint, associating the tangent-direction at a point on a curve with the instantaneous direction of a point which by its (smoothly continuous) "movement" generates the curve. Thus FERMAT[9] develops a classical proof by inequalities of the standard result that, where ABC is an arbitrary (convex) curve with respect to which a second curve is defined such that, for a general normal $B'Bb$ to fix-line axis Ab, always the ratio $B'b:Bb$ is constant, then the tangents of corresponding points C, C'—that is, such that CC' is perpendicular to Ab—meet in a point T on Ab. Clearly, however, the "instantaneous direction" of points C, C' is made up of a constant downward component CF, $C'F'$ and horizontal components which are respectively

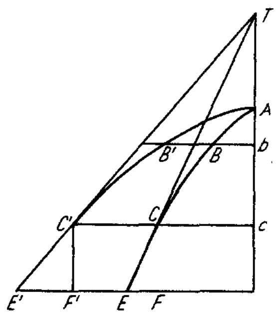


Fig. 93

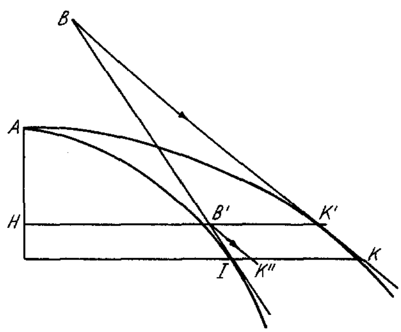


Fig. 94

proportional to cC, cC' (where CC' meets Ab in c), and the result is immediate. GREGORY's generalization of WREN's cycloid-tangent construction follows equally naturally by limit motion considerations: here, since curve AHK is defined from curve $AB'I$ by $\widehat{AB'I}:IK$ is constant where I, K are corresponding points cut off by a general parallel, the limit-motion of K is compounded out of motions in the instantaneous direction of the curve AI at I and again, parallel to HIK whose magnitudes are in the ratio $AI\ (=BI):IK$; and so, since BI, IK are drawn in these respective directions with an equal ratio of length, the triangle BIK is a differential triangle at the point K, and so BK is tangent at K.

Perhaps the most elaborate 17th century treatment of the tangent-problem through the concept that the tangent-direction at a point is that of the direction of the limit-motion of the generating point of the curve at that point was that written up by ROBERVAL sometime in the mid-century in his treatise *Sur la composition des mouvemens*[10], whose "Axiome ou principe d'invention"[11] enunciates

[9] See *de linearum curvarum ... comparatione ...*: prop. 3. A wide selection of similar examples are to be found in BARROW, *LG*: lectiones **9**, **10**—**10**: 10 is a completely typical example which slightly generalizes FERMAT's proof (given a similar proof by inequalities).

[10] *Observations sur la composition des mouvemens, et sur le moyen de trouver les touchantes des lignes courbes*, first published in 1730 in Méms. de l'ac. roy. des sc. (1666—1699): **6**: 1—89. Perhaps the first publication of such limit-motion ideas is contained in a minor work of DESARGUES (found only in 1951) printed apparently with his *Brouillon project* of 1639, *Atteinte aux evenements des contrarietez d'entre les actions des puissances ou forces* (see R. TATON: "*L'œuvre mathématique de G. Desargues*, Paris, 1951: 181—184).

[11] *op. cit.* 24.

explicitly: "la direction du mouvement d'un point qui décrit une ligne courbe est la touchante de la ligne courbe en chaque position de ce point-là", which is applied[12] in his "Règle générale": "Par les proprietez spécifiques de la ligne courbe (qui vous seront données) examinez les divers mouvements qu'a le point qui la décrit à l'endroit où vous voulez mener la touchante: de tous les mouvemens composez en un seul, tirez la ligne de direction du mouvement composé, vous avez la touchante de la ligne courbe."

These definitions were accepted in more or less equivalent form by those —BARROW, GREGORY and especially NEWTON in England[13]—who used the limit-motion definition of the tangent. In particular, definitions 9, 10 of NEWTON's manuscript *geometria curvilinea*[14] state axiomatically:

9. "The locus of a moving point is the ... curve which the point describes by its motion",

and

10. "The determination of the motion of a point is the position of the line touching that curve at the moving point"*.

Using such definitions (and the limit-process implicit in them), it becomes possible to resolve the general tangent-problem in a wide variety of "mechanically" defined curves—ROBERVAL in his tract discusses the conics, cycloid, cissoid, various conchoids, PASCAL's limaçon, quadratix, ARCHIMEDEAN spiral, curves given, for the most part, improved treatment by NEWTON in his manuscript studies of 1664 to 1665—but from a general viewpoint it is important to emphasise the complete generality of treatment afforded by composition of motions. Where the classical treatment by line-inequalities is restricted to considering those curves which are defined in some way as the point-set meet of coordinate line-lengths (which may be, as in the case of the cycloid, curve arc-lengths in general), the limit-motion method considers general coordinate systems indifferently, and in particular extends to polar coordinate systems. Perhaps the simplest of all curves defined in a full bipolar coordinate system are the central conics referred to their foci: in fact, the condition that, where S_1, S_2 are fix-points, the condition that the sum or difference of S_1P and S_2P be constant defines P to be on an ellipse or hyperbola, and it is immediate that the compound motion of P is directed in the directions Pt_2, Pt_1 respectively which bisect $\widehat{S_1'PS_2}$, $\widehat{S_1PS_2}$ (being made up of equal increments or decrements of S_1P, S_2P).

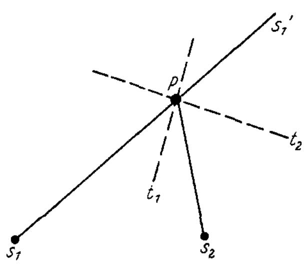


Fig. 95

* Specifically, these are part of an axiom-scheme on which NEWTON erects a geometrical theory of fluxions in full EUCLIDEAN manner.

[12] *op. cit.* 28.

[13] Compare BARROW *LG*: lectiones 3ff.; GREGORY *GPU*; and NEWTON *CUL Add.* 4004 (dated 1664—1665): 4Vff., 50Vff., *Add.* 3958. 3 (dated October 1666): *passim*, *Add.* 3960. 14 (to be dated 1671): prob. 4: 43—56 (*curvarum tangentes ducere*, especially 45ff.), *Add.* 3963. 7: 46R—60V (*geometria curvilinea*, to be dated about 1680).

[14] *Add.* 3963: 47R.

A more interesting generalization was sketched by FATIO DE DUILLIER[15] at the end of the century, and his proof is a curious combination of limit-motion considerations and the Greek conception of a tangent as lying wholly outside the curve it touches: Defining the point-set C with respect to fix-poles a, d by $\lambda \times ac + \mu \times cd = v$, constant, or all positions of c, then where m, p are taken in ac, cd with $mc = pc$ and n is taken in np such that $mn:np = \mu:\lambda$, then cn is normal to the curve $[c]$ (and so ce, drawn perpendicular to cn, is tangent at c). In proof FATIO considers an arbitrary length ce, drawing geh, ef parallel to cd, ca respectively (where cg, dh; af, eb are perpendicular to cd, ca) with mo, pq perpendicular to cn: then $cm:om = ec:cb$ and $cp\ (=cm):pq = ec:cg$, or $\mu:\lambda = mn:np = mo:pq = cb:eg$ (so that $\lambda \times cb = \mu \times eg$); further,

$$\begin{aligned} v &= \lambda \times ac + \mu \times cd \\ &= \lambda \times ab - \lambda \times bc\ (= \mu \times eg) + \mu \times gh \\ &= \lambda \times fe + \mu \times eh < \lambda \times ae + \mu \times ed, \end{aligned}$$

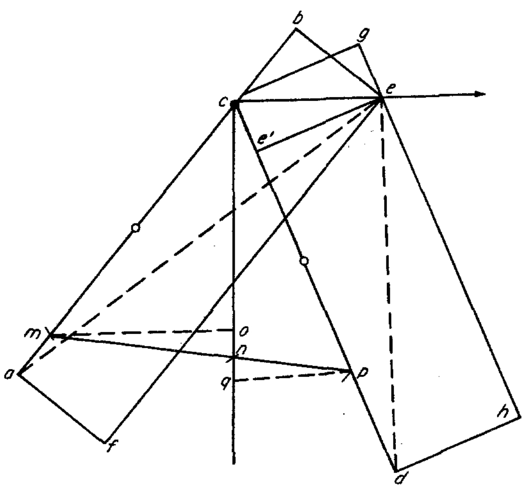


Fig. 96

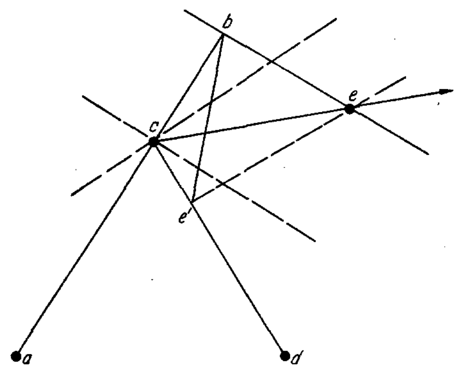


Fig. 97

or e lies outside the curve (and a similar proof holds for e taken on the further side of c).* Basic in this proof is the transform which is derived from $cb:eg$ $(=ce') = \mu:\lambda$—in fact the triangle cbe' is virtually a differential triangle of the instantaneous increment and decrement of c, where cb, ce' are taken in the same proportion $\big(\text{since } \lambda\alpha + \mu\beta = \lambda(\alpha + \mu k) + \mu(\beta - \lambda k)\big)$ and the point e on the tangent at c is found immediately as the point set of the meets of the normals be, $e'e$ to cb, ce' at b, e' respectively.

The general concept is no more difficult to set up on a general coordinate model[16]. Essentially we have a curve PP' defined as a point-set, each point P

* FATIO sketches in the generalization to n fix-poles $a, d, \ldots$ and develops a model of weighted means to determine the instantaneous normal cn. The "pressure centre" form in which his result is stated makes it certain that limit-motion ideas are basic in the concept.

15 The generalization had been suggested by TSCHIRNHAUS in his *medicina mentis*, Amsterdam, 1687 but his method of solution contained a conceptual error (and a wrong general result), and FATIO DE DUILLIER evolved his method in correction (see *Réflexions de Mr. N. Fatio de Duillier sur une méthode de trouver les tangentes de certaines lignes courbes* · ≡ · *Bibliothèque universelle et historique* 5 (Amsterdam 1681): 25–33.

16 Such a general treatment is given by NEWTON in the 1671 tract on analysis—see prob. 4: *curvarum tangentes ducere* (*CUL Add.* 3960. 14: 43–56 · ≡ · HORSLEY's *de tangentibus curvarum ducendis* in his printed text, *geometria analytica*: cap. 6: 1: 430–443).

of which is determined by the cut of, say, two lines XP, YP, themselves definable in position in some (unique) way for each point P—that is, where XP, YP form a set of coordinate line-lengths. Defining the tangent-direction at P as the instantaneous direction of the curve of P and considering (suitably expanded) momentary motions of $XP \to X'P'$, $YP \to Y'P'$ (which may in the limit, by the postulate of continuity, be taken as coparallel translations of XP, YP respectively), we can see these increases ($X''P'$, $Y''P'$ say) when the arc PP' becomes indefinitely small as the components of the limit motion of P' at P, and the problem of constructing the tangent at a point on the curve reduces to using the particular defining coordinate system and the determining "relatio" between XP, YP which exists for general P to calculate the value of the limit-ratio $X''P$: $Y''P$ (or the ratio of the limit-normals Px, Py which define the limit-direction

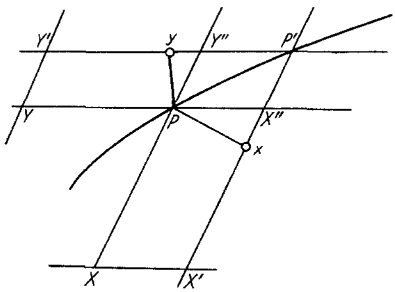


Fig. 98

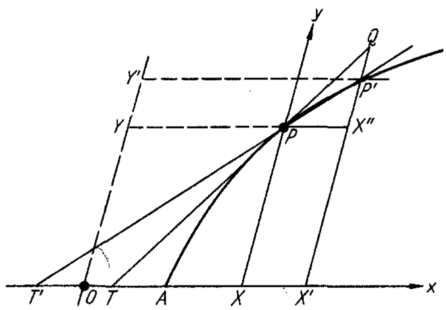


Fig. 99

PP' of the curve at P in a similar way). Clearly the way is then open for application of the concept of fluxional (or differential) increase, and for that inevitable process of abstraction by which tangent-methods came finally, in the later 17^{th} century, to be subsumed into the calculus as but a single (and in no way unique) model of differential procedures. In this growing alliance of geometry and calculus, though particular curves received polar definitions, and NEWTON in manuscript developed still other systems, especially the bipolar[17], only the Cartesian coordinate system (with its two sets of co-parallel line-lengths)[18] was applied generally, and to that we restrict our attention.

Taking general oblique axis-directions Ox, Xy, and assuming that the curve $\widehat{APP'}$ is defined as the point-set of P which satisfies some "relatio" f between the abscissa OX $(=YP)$ and ordinate XP, consider a second point P' on the curve which we can in the limit take indefinitely near to P (with corresponding ordinate $X'P'$ meeting the tangent at P in Q so that, as both $X'P'$, $X'Q \to XP$, $X'P' \approx X'Q$). Analytically, taking $OX = x$, XP $(=OY) = y$, we can define the curve APP' by $y = f(x)$, where x, y are generally real, and f is some definable relation connecting x and y; and similarly, where $OX' = x'$, $X'P' = y'$, point P' corresponds to x', y' where $y' = f(x')$. Then, since $PX'' = x' - x$, $X''P' = y' - y$ and $\lim_{x' \to x} (QX'') = \lim_{x' \to x} (P'X'')$, we can find the length of the subtangent by considering either the limit-equality $P'X'' = QX''$, or the limit-position of T' (the meet

[17] See previous note, and compare *Add.* 4004: 50Vf.

[18] Compare chapter seven for a more detailed analysis of the system.

of chord PP' with OX), which produce respectively

$$\text{subtangent } XT \left[\begin{array}{l} = PX \times \lim\limits_{x' \to x} \left(\dfrac{PX''}{X''P'} \right), \\ = PX \times \lim\limits_{x' \to x} \left(\dfrac{PX''}{X''Q} \right). \end{array} \right.$$

Historically both approaches were used in the first analytical derivations of subtangent-length in the 1630's, the first by DESCARTES and the second by FERMAT[19]. Both demonstrated their methods on the simple parabola, $y = (kx)^{\frac{1}{2}}$, and it is interesting to point the slight differences of treatment required in the two approaches by sketching their proofs.

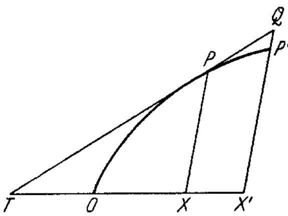


Fig. 100

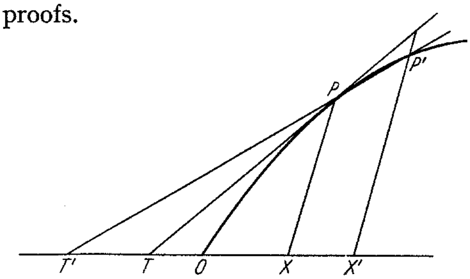


Fig. 101

In the FERMATIAN approach we consider the limit-meet of the general ordinate $X'P'$ with the tangent PT at P: specifically, taking the parabola OPP' defined by $OX:OX' = XP^2:X'P'^2$ or $x:x' = y^2:y'^2$, we have immediately, since the parabola is convex up, that $X'P' < X'Q$, and, again, by similar triangles, $XP:X'Q = TX:TX'$; so that, taking $XT; t$, $x:x' \geqq t^2:(t+x'-x)^2$, with equality only in the limit as $x' \to x$; and finally, $\frac{x'-x}{x} \leqq \frac{(t+x'-x)^2 - t^2}{t^2}$, or $t \leqq 2x + \frac{x'-x}{t}$, so that $XT = \lim\limits_{x' \to x} \left(2x + \frac{x'-x}{t}\right) = 2x$.

In contrast, DESCARTES' approach considers the limit-meet T' of the chord PP' with OX as $P' \to P$: taking $XT' = t'$, it follows that $OX':OX = X'P'^2:XP^2$, or $x':x = y'^2:y^2$, with $y':y = T'X':T'X$ by similar triangles; or

$$\frac{x'-x}{x} = \frac{(t'+x'-x)^2 - t'^2}{t'^2}, \quad \text{or} \quad t' = 2x + \frac{x'-x}{t'};$$

and finally subtangent

$$XT = \lim_{x' \to x} (t') = \lim_{x' \to x} \left(2x + \frac{x'-x}{t'}\right) = 2x.$$

[19] G. MILHAUD in his *Descartes savant*, Paris 1921: 149—175: *La querelle de Descartes et de Fermat au sujet des tangentes*, especially 149—162, 164—165 takes care to separate the two viewpoints. Compare, too, MERSENNE'S letter to DESCARTES of 28 April 1632 (*Oeuvres*, ed. ADAM & TANNERY: 2: 119—120) and DESCARTES' answering letter of 27 July 1632 (*Oeuvres*, 2: 252). It is interesting to note that DESCARTES does not use the subtangent method in his (1637) *Géométrie*: rather he finds the subnormal length first by calculating the position of the centre of the circle which touches the curve (that is, wh ich has two coincident meets with it)—see *Géométrie*, Leyden 1637: 342—351 (ed. SMITH & LATHAM) 94—113; and compare MILHAUD: *Descartes, savant*: 128.

The FERMAT approach is not very generally applicable without troublesome justification of the limit-equality $X'P' = X'Q$.[20] WALLIS, however, gives a fine example of the method applied to finding the subtangent at a general point on the cissoid[21], defined with regard to rectangular coordinate axes by $PX^2(2AC - AX) = AX^3$, or $y = \left(\frac{x^3}{2r-x}\right)^{\frac{1}{2}}$, where $AX = x$, $XP = y$, $AC = r$: from the known convexity of the cissoid, where PX'' is tangent and $P'X''X'$ an arbitrary ordinate distinct from PX, $P'X' > X''X'$; or taking as before $XT = t$, $AX' = x'$, $X'P' = y'$, then $\left(\frac{x'^3}{2r-x'}\right)^{\frac{1}{2}} > \frac{t+x'-x}{t} \times \left(\frac{x^3}{2r-x}\right)^{\frac{1}{2}}$, with equality in the limit as $x' \to x$; and finally (squaring, multiplying out, cancelling) in the limit as $P' \to P$ (and so $x' \to x$) we find $t = \frac{x(2r-x)}{3r-x}$.

The DESCARTES approach,* however, which takes the meet of the limit chord PP' with the axis AX as $P' \to P$ has a far richer application, and it is in this form that the extensions of the simple analytical treatments were made in the final, completely general methods evolved. And it had the practical advantage of substituting a limit position of the subtangent XT for the rather clumsy introduction of the FERMATIAN inequality which tends to equality in the limit.

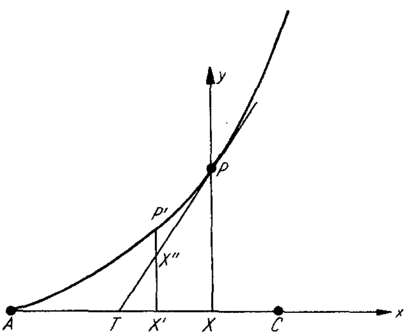


Fig. 102

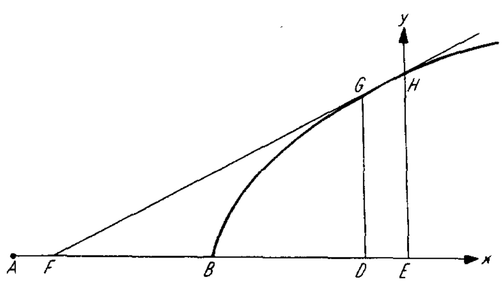


Fig. 103

A fine example of the approach is given by JAMES GREGORY[22] in constructing the subtangent to the general hyperbola $\alpha y^m = \beta x^m(a+x)$. In particular, he gives a type-solution for the case $a^3y^3 = c^3x^2(a+x)$ which, with respect to axis AK and general ordinate HE, he gives in geometrical form by $BE^2 \times AE : EH^3 = c^3 : a^3$ (constant). Then, where $AB = a$ (or B is the vertex), $BE = x$, $EH = y$, $EF = t'$, consider a second point G on the curve indefinitely near to H, denoting the corresponding axis-segment DE by o, a "nothing or lately so" (nihil seu serum o); by similar triangles $FD:FE = DG:EH$, or $\frac{t'-o}{t'} = \left(\frac{(x-o)^2 \times (a+x-o)}{x^2 \times (a+x)}\right)^{\frac{1}{3}}$; or, reducing and dividing by o, in the limit as o becomes indefinitely small (and $t' \to$ the subtangent t), we have finally $t = \frac{3x(a+x)}{2a+3x}$.

---

* Many historians—surely unfairly—call the approach "FERMATIAN".

[20] Though it is applied to the ellipse and hyperbola in a similar way by WALLIS in *de sectionibus conicis ...*, Oxford, 1656: props. 30, 36; and repeated in his article *binae methodi tangentium ...*, *PT* **7** (1672): 4010—4016.

[21] In *binae methodi tangentium ...* (*op. cit.* previous note): 4012ff.

[22] In his *GPU*: prop. 7: 20—22: *rectam ducere datam curvam tangentem in ejus puncto dato, si modo curva sit ex earum numero quas Cartesius appellat geometricas.*

The DESCARTES approach has an immediate analytical generalization in the form $\frac{y}{t} = \lim_{x' \to x}\left(\frac{y'-y}{x'-x}\right)$, where $y=f(x)$ is some *relatio* which connects x and y*—indeed from the mid-17th century the simple analytical tangent problem became reduced to that of finding dy/dx in some equivalent form, where the relatio $f:[y=f(x)]$ is given, and it is in no way surprising that in a suitably interpreted form the major standard differential forms should be found, for example, in BARROW'S *LG* in subtangent form.[23] Clearly some modification of the above approaches was sufficient to construct the subtangent to any algebraic curve whose representing equation can be put in the form $y=f(x)$, and by the late 1660's all the important standard examples of that form had been "differentiated" in a wide variety of ways.** In general, where $y = \sum_{0 \leq i \leq n} (a_i \cdot x^i)$ is the general polynomial, the constructions derive, in one shape or another,

$$\frac{y}{t} = \lim_{x' \to x}\left(\frac{y'-y}{x'-x}\right) = \lim_{x' \to x}\left(\frac{\sum_{0 \leq i \leq n} a_i (x'^i - x^i)}{x'-x}\right)$$
$$= \sum_{0 \leq i \leq n}\left(a_i \times \lim_{x' \to x}\left(\frac{x'^i - x^i}{x'-x}\right)\right),$$

but especially favoured reductions are the FERMATIAN which substitutes $x' = x + (x'-x)$, and so derives

$$\lim_{x' \to x} \frac{x'^i - x^i}{x'-x} = \lim_{x' \to x}\left(\frac{i\, x^{i-1}(x'-x) + O\,[(x'-x)^2]}{x'-x}\right) = i\, x^{i-1};$$

and the more straightforward development—used by TORRICELLI in manuscript and, as we shall see, by SLUSIUS in his generalization—which yields

$$\lim_{x' \to x} \frac{x'^i - x^i}{x'-x} = \lim_{x' \to x}\left(\sum_{1 \leq j \leq i} [x'^{i-j} x^{j-1}]\right) = i\, x^{i-1}.$$

* This is an exact analytical equivalent of the differential triangle approach—clearly, in the usual symbolism $\frac{y}{t} = \left(\frac{dy}{dx}\right)$, or $t = y \cdot \frac{dx}{dy}$.

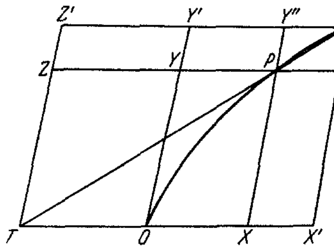


Fig. 104

** Notably TORRICELLI[24] had given a remarkable reduction of the subtangent problem to one of quadratures. Thus where $OX = x$, $XP = y$ and the point-set P has representing equation $y^m = \lambda^{m-1} \cdot x$, subtangent TX: abscissa OX = parallelogram $(ZZ'Y''P)$: parallelogram $(YY'Y''P)$ = area $(ZZ'P'P)$: area $(YY'P'P) \approx$ area $(XX'P'P)$: area $(YY'P'P) = m:1$, since area OXP : area OYP = area $OX'P$: area $OY'P' = m:1$ by a known quadrature.

[23] J.M. CHILD in *The geometrical lectures of Isaac Barrow*, Chicago 1916 insists far too much on this interpretation, without emphasising the conditioning factor of a geometrical model.

[24] See *Delle tangenti alle parabole infiniti per lineas supplementares* ·≡· *opere* **3**: 320ff.—compare E. BORTOLOTTI: *L'opera geometrica di Evangelista Torricelli*, Monatsschr. für Math. und Physik **48** (1939): 457—486.

(The former approach is that tidied up in a pleasing way by BARROW in a passage in LG[25] quoted in all the standard histories: specifically he introduces the notation, derived almost certainly from "FERMAT's" A, E symbolism, of $a=y'-y$, $e=x'-x$ and considers $\lim_{a,e\to 0}(a/e)=y/t$, which gives him the basis for his substitution rule $\begin{bmatrix} a\to y \\ e\to t \end{bmatrix}$.

A generalization which soon suggests itself is to the case where the representing equation of the curve is given in the implicit form $g(x, y)=0$. Particular cases afford no difficulty—we merely use the substitution $\begin{bmatrix} y'=y+(y'-y) \\ x'=x+(x'-x) \end{bmatrix}$ to reduce between the given equations $g(x, y)=0$, $g(x', y')=0$, and derive $\left(\frac{y}{t}=\right)\lim_{x'\to x}\left(\frac{y'-y}{x'-x}\right)$.

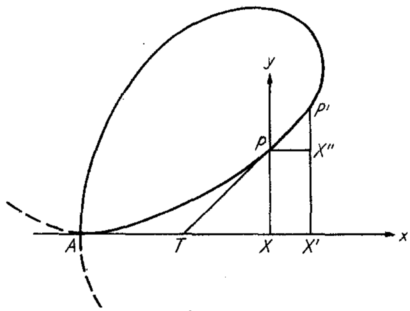


Fig. 105

Thus BARROW[26] constructs the subtangent to DESCARTES' folium, $x^3+y^3=\lambda xy$, by substituting $\begin{bmatrix} y\to y+a \\ x\to x+e \end{bmatrix}$; so that

$$O=(x+e)^3+(y+a)^3-x^3-y^3-\lambda(x+e)(y+a)+\lambda xy$$

$$=3x^2e+3xe^2+e^3+3y^2a+3ya^2+a^3-\lambda(ey+ax+ae)$$

or

$$\lim_{a,e\to 0}\left(\frac{a}{e}\right)=\lim_{a,e\to 0}\left(\frac{\lambda y-3x^2+\lambda e-3xe-e^2}{3y^2-\lambda x+3ya+a^2}\right)$$

and finally

$$\frac{y}{t}=\frac{\lambda y-3x^2}{3y^2-\lambda x}.$$

From the late 1650's general rules began to appear which removed the necessity for brute-force calculation afresh in each new curve or representing equation. JOHANN HUDDE[27] in fact, was the first to evolve a workable rule—derived apparently from a numerical induction over particular instances—but its complexity and cumbrousness made it little appreciated[28]. More widely known

---

[25] BARROW LG: lectio **10**: 80—84. His rule 3 defines essentially the basic equality (his "differential triangle") $\frac{t}{y}=\frac{a}{e}=\lim_{x'\to x}\left(\frac{x'-x}{y'-y}\right)$.

[26] LG: lectio **10**: 82 ·≡· example 3 of previous note (BARROW calls the curve "la galande").

[27] Hinted at in his tract *de maximis et minimis*—which is his letter of 27 January 1659 to VAN SCHOOTEN as printed in Book **1** of the 1659 Latin translation of DESCARTES' *Géométrie*, Amsterdam, 1659: 507—516. The application of his rule to tangents (exactly as NEWTON was to make it again in 1665) occurs in a second letter to VAN SCHOOTEN of 21 November 1659, printed only in 1713 in the Journal Literaire de La Haye (reprinted in GERHARDT (B): 234—237).

[28] NEWTON, however, was powerfully influenced by reading HUDDE during his formative years 1664—1665. The HUDDE tract is quoted several times in the *Waste Book* (*CUL Add.* 4004) and specifically several times on 47V; while HUDDE's rule is the basis for his own second-order partial difference forms—see next chapter.

in the period was SLUSIUS' restatement of the HUDDE rule, whose publication[29] led to a minor priority dispute with NEWTON in the 1670's (who had independently made the same generalization of HUDDE's rule but did not publish it). Extent manuscripts fill the lack of any proof in the published text, and show that SLUSIUS took the general implicit representing equation of a curve by $g(x, y) = 0 = \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq m} (a_{ij} \cdot x^i \cdot y^j)$ from which (and the second equation $g(x', y') = 0$) he derives the subtangent length t in the standard form $\frac{y}{t} = \lim_{x' \to x} \left(\frac{y' - y}{x' - x}\right)$ by using a "TORRICELLI" type reduction of $\lim_{x' \to x} \left(\frac{x'^i - x^i}{x' - x}\right) = i \cdot x^{i-1}$. Thus, to sketch his rather lengthy treatment,

$$O = g(x', y') - g(x, y)$$

$$= \begin{cases} \frac{g(x', y') - g(x, y')}{x' - x} \times (x' - x) + \frac{g(x, y') - g(x, y)}{y' - y} \times (y' - y) \\ \frac{g(x', y') - g(x', y)}{y' - y} \times (y' - y) + \frac{g(x', y) - g(x, y)}{x' - x} \times (x' - x), \end{cases}$$

or

$$\frac{y}{t} = \lim_{x' \to x} \left(\frac{y' - y}{x' - x}\right) = -\frac{\partial g}{\partial x} \Big/ \frac{\partial g}{\partial y},$$

where

$$\frac{\partial g}{\partial x} = \lim_{x' \to x} \left(\frac{g(x', y') - g(x, y')}{x' - x}\right) = \lim_{x' \to x} \left(\frac{g(x', y) - g(x, y)}{x' - x}\right) = \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (a_{ij} i x^{i-1} y^j),$$

and

$$\frac{\partial g}{\partial y} = \lim_{\substack{x' \to x \\ y' \to y}} \left(\frac{g(x', y') - g(x', y)}{y' - y}\right) = \lim_{x' \to x} \left(\frac{g(x', y) - g(x, y)}{y' - y}\right) = \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (a_{ij} x^i j y^{j-1});$$

so that finally

$$t = \frac{\sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (j a_{ij} x^i y^j) \times x}{\sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (i a_{ij} x^i y^j)}$$

(which is exactly HUDDE's rule)*.

This proof holds only for rational algebraic functions $g(x, y) = 0$, and SLUSIUS does not seem to have extended by pattern-analogy to a general non-algebraic function g—such a generalization was made by NEWTON in 1665[30] (though he did

* This, in the form $t \frac{\partial g}{\partial x} + y \frac{\partial g}{\partial y} = 0$, is the more familiar $\frac{\partial g}{\partial x} \frac{dx}{dy} + \frac{\partial g}{\partial y} = 0$ $\left(\text{since } t = y \frac{dx}{dy}\right)$.

[29] In *PT* 7 (1672): 5143—5147: ... *Slusius ... his short and casie method of drawing tangents to all geometrical curves without any labour of calculation.* SLUSIUS, perhaps influenced by TORRICELLI during his stay in Italy (in 1642—1651), seems to have come upon the rule in the late 1650's—compare L. ROSENFELD: *René-François de Sluse et le problème des tangentes,* Isis **10** (1928): 416—434, who gives a detailed analysis of the manuscripts now in the Bibliothèque Nationale.

[30] In *CUL Add.* 4004: 48Rff. (dated May 21st 1665), and tidied up a year and a half later in the October 1666 manuscript *On resolving problems by motion, Add.* 3958. 3: 48—76, especially Problem 2d: 55ff.

not publish any hint of it till 1713[31], and that heavily disguised). NEWTON apparently derived the restricted SLUSIUS rule in the FERMATIAN way, and then generalized by analogy. Taking again

$$g(x, y) = 0, \ = \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (a_{ij}\, x^i\, y^j),$$

he substitutes $\begin{bmatrix} x' = x + (x' - x) \\ y' = y + (y' - y) \end{bmatrix}$ in

$$O = g(x', y') - g(x, y) = \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} \left(a_{ij}(x'^i\, y'^j - x^i\, y^j)\right),$$

ignoring powers of $(x' - x)$, $(y' - y)$ higher than the first (in accordance with the BARROW rule), so that

$$x'^i\, y'^j - x^i\, y^j \approx i\, x^{i-1}\, y^i (x' - x) + x^i j\, y^{j-1} (y' - y);$$

and finally

$$\frac{y}{t} = \lim_{x' \to x} \left(\frac{y' - y}{x' - x}\right) = - \frac{\sum\limits_{0 \leqq i \leqq m} \sum\limits_{0 \leqq j \leqq n} (i\, a_{ij}\, x^{i-1}\, y^j)}{\sum\limits_{0 \leqq i \leqq m} \sum\limits_{0 \leqq j \leqq n} (j\, a_{ij}\, x^i\, y^{j-1})},$$

which is HUDDE'S rule again. NEWTON symbolizes these operations[32]: where)(is the general function $g(x, y)$, or

$$)(= \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (a_{ij}\, x^i\, y^j),$$

he defines

$$\cdot)(= \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (i\, a_{ij}\, x^i\, y^j) \left(= x \frac{\partial g}{\partial x}\right)$$

and

$$)(\cdot = \sum_{0 \leqq i \leqq m} \sum_{0 \leqq j \leqq n} (j\, a_{ij}\, x^i\, y^j) \left(= y \frac{\partial g}{\partial y}\right),$$

---

[31] In *commercium epistolicum*: 29–30 (which prints his letter to COLLINS of 10 December 1672). A mangled account is inserted in the corrected second edition of *PM* (*CUL Adv.* b. 39.2: attached between pp. 226–227) where he states, having given an improved version of the HUDDE rule: "This is a very small part or rather a corollary of a general method which extends without any laborious calculation not only to drawing tangents to any curves, whether geometrical or mechanical, relating in whatever manner right lines or other curves (defined with respect to a suitable coordinate reference framework), but also to resolving other more abstruse types of problems or curvatures, areas, rectifications, centres of gravity of curves *etc.*, nor (as Hudde's method for maxima and minima) is it restricted to those equations which are free from surds ..." All this is expounded at great length in the October 1666 manuscript of the previous note, and will be examined in more detail in the next chapter. The general method, of course, is what later came to be called his fluxion theory, but in the 1666 tract is merely given as theorems on limit-motion.

[32] See *Add.* 4004: 48V–49R; *Add.* 3958.3: 55ff. (of which a full account is given in the (? WM. JONES) tract on history of fluxions at *Add.* 3960. 2, especially 11ff.). NEWTON intended to publish the rule at least twice in the early 1700's but never did: *Add.* 3968: 245R is a note on his method apparently meant to be added to RAPHSON'S *History of fluxions*, while *Add.* 3965: 377R is a note which was meant to be added in a scholium to $_3PM$, 1726.

and can now write the general rule as

$$\left[y\frac{dx}{dy}=\right]t=-\frac{)(\cdot}{\cdot)(}\times x\left[=-\frac{\partial g}{\partial y}\Big/\frac{\partial g}{\partial x}\times y\right].\star$$

It is an easy assumption to suppose that this rule is true for all real i (and, even, that an analogous rule holds for functions $)(=g(x,y)=0$ which are not representable by simple algebraic polynomials of finite degree).

In this most general form (given, admittedly, without rigorous proof) the tangent-problem was virtually solved for general algebraic functions—the extension to $g(x_1, x_2, \ldots, x_n)=0$ is sketched by NEWTON—and, taken in conjunction with contemporary advances in using (convergent) polynomials to approximate to given functions, effectively solved the general tangent problem at a practical level. But no corresponding method could be applied to the case of the non-algebraic function—DESCARTES' "mechanical" curves, LEIBNIZ' "transcendental" equations—till the limit-operation $\frac{y}{t}=\lim_{x'\to x}\left(\frac{y'-y}{x'-x}\right)$ could be adapted to consider such functions by a general method. This, in turn, depended on having a general concept of analytical function (corresponding to the geometrical model of "curva quaevis"), but progress towards that was slow and little more than begun in the period. We find in the later 17^{th} century a curious mixture of geometrical and analytical ideas in considering the tangent problem for non-algebraic curves.

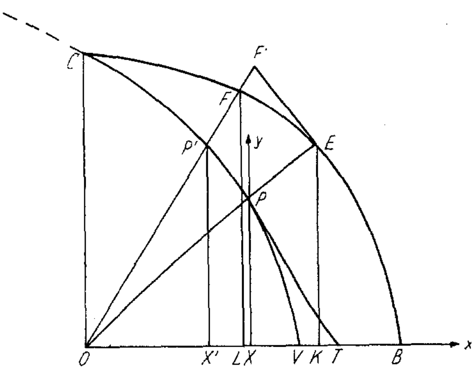


Fig. 106

BARROW'S construction of the subtangent to the quadratrix neatly makes the point[33]. Where the quadratrix CPV is defined from the circle quadrant OBC such that, for PX parallel to OC, $PX : CO = \text{arc}\,\widehat{BE} : \text{arc}\,\widehat{BC}$ (or $\widehat{BE} : PX = \widehat{BC} : CO = \frac{1}{2}\pi$), consider a second point P' on the curve which we will take indefinitely near to P in the limit: taking $OX=x$, $OX'=x'$: $PX=y$, $P'X'=y'$; OC $(=OB)=r$, we wish to find the subtangent length $TX=t$. Immediately, by definition,

$$FE=\tfrac{1}{2}\pi(P'X'-PX)=\tfrac{1}{2}\pi(y'-y),$$

$\star$ Clearly the side-dots are equivalent to partial-differential operators: $x\frac{\partial g}{\partial x}\equiv\cdot)($, $y\frac{\partial g}{\partial y}\equiv\cdot)($, and NEWTON saw as much. Further he states $\cdot(\cdot)()\,[\equiv\cdot\cdot)(]=x^2\frac{\partial^2 g}{\partial x^2}$, $()(\cdot)\cdot[\equiv)(\cdot\cdot]=y^2\frac{\partial^2 g}{\partial y^2}$ and $(\cdot)()\cdot=\cdot()(\cdot)\,[\equiv\cdot)(\cdot]=xy\frac{\partial^2 g}{\partial x\,\partial y}$ (for the elementary functions, at least, considered in NEWTON'S day), using them to derive a formula for the radius of curvature of a general point on $)(\equiv g(x,y)=0$. (See next chapter.)

[33] BARROW *LG*: lectio **10**: example 4: 82–83. The following example 5 (which yields, in equivalent form, $dy/dx=\sec^2 x$, where $y=\tan x$) has been analyzed by J.M. CHILD in some detail—see his *Geometrical lectures of Isaac Barrow*, Chicago, 1916: 121–123, which perhaps insists over strongly on the supremacy of the analytical equivalents in BARROW'S mind.

and
$$EK = PX \times \frac{OE}{OP} = x \times \frac{r}{(x^2+y^2)^{\frac{1}{2}}};$$

further, taking $EF' = m$ tangent at E (meeting OF in F'), we easily show that
$$LK = EF' \times \frac{PX}{OP} = m \times \frac{y}{(x^2+y^2)^{\frac{1}{2}}}$$

and
$$OK = OX \times \frac{OE}{OP} = x \times \frac{r}{(x^2+y^2)^{\frac{1}{2}}} \quad \text{or} \quad OL = \frac{rx - my}{(x^2+y^2)^{\frac{1}{2}}},$$

and
$$FL^2 : OL^2 = \left(r^2 - \frac{(rx-my)^2}{x^2+y^2}\right) : \frac{(rx-my)^2}{x^2+y^2}$$
$$= \left(r^2(x^2+y^2) - (rx-my)^2\right) : (rx-my)^2;$$

and, finally, in the limit at $P' \to P$ $\left(\text{or } EF' \to EF,\ m \to \frac{1}{2}\pi(y'-y)\right)$, $FL^2 : OL^2 = PX'^2 : OX'^2$, or
$$\frac{r^2(x^2+y^2) - (rx - \frac{1}{2}\pi y(y'-y))^2}{(rx - \frac{1}{2}\pi y(y'-y))^2} = \frac{(y+(y'-y))^2}{(x+(x'-x))^2},$$

so that by reducing, cancelling and dividing out $(x'-x)$ we find
$$\frac{y}{t} = \lim_{x' \to x}\left(\frac{y'-y}{x'-x}\right) = \frac{r^2 y}{r^2 x - \frac{1}{2}\pi r(x^2+y^2)},$$

or
$$-t = TX = \frac{x^2+y^2}{2r/\pi} - x = \frac{OP^2}{OV} - OX,\star$$

and
$$OT\ (= OX + TX) = \frac{OP^2}{OV}.\star\star$$

More straightforwardly analytical procedures were introduced over the next few decades, particularly (in Britain) by DAVID GREGORY[34] and JOHN CRAIG[35], whose (1693) *tractatus mathematicus* was perhaps the last work which could omit

---

⋆ Using the classical result that $OV \left(= \lim_{y \to 0}\left[y \cot \frac{\pi y}{2r}\right]\right) = \frac{2r}{\pi}$.

⋆⋆ This, of course, has a strict analytical equivalent in finding
$$\frac{t}{y}\left(= \frac{dx}{dy}\right) = \lim_{y' \leftarrow y}\left(\frac{y' \cot \frac{\pi y'}{2r} - y \cot \frac{\pi y}{2r}}{y'-y}\right) = -\frac{\pi y}{2r} \operatorname{cosec}^2 \frac{\pi y}{2r} + \cot \frac{\pi y}{2r},$$
since $x = y \cot \frac{\pi y}{2r}$.

34 In his *exercitatio geometrica de dimensione figurarum* ..., Edinburgh, 1684, which largely gives an analytical treatment—often by indivisible methods—of his uncle JAMES' *GPU* and *EG*.

35 In *methodus figurarum ... quadraturas determinandi*, London, 1685, which (p. 28ff.) introduced into England LEIBNIZ' $\frac{dy}{dx}$ notation for $\lim_{x' \to x}\left(\frac{y'-y}{x'-x}\right)$, published the previous year by LEIBNIZ in *AE* (1684): 467—473: *nova methodus pro maximis et minimis, itemque tangentibus* ...; and in his *tractatus mathematicus de figurarum ... quadraturis* ..., London, 1693: *pars posterior*: 44—48: *methodus determinandi tangentes linearum transcendentium*.

an account of NEWTON'S fluxional methods—first published, in a bare sketch, the same year by WALLIS.[36]

CRAIG'S early mathematical work was strongly influenced by LEIBNIZ' various published articles (in the periodical Acta Eruditorum)—though he had met NEWTON at Cambridge in the middle 1680's and been shown his fluxion manuscripts[37]—and, accepting the LEIBNIZIAN classification of curves into algebraic and "transcendental", he tried to generalize LEIBNIZ' new differential algorithm (applied in *AE* (1684) only to the tangent-problem for algebraic curves) to deal with general types of transcendental curves. In particular[38], he obtains a generalized rule for deriving the subtangent at general point P on the curve OPP', defined as the point-set of x, y such that $y = f(v, x)$, where $OX = x$, $PX = y$, and $v = \text{arc}\,\widehat{OQ}$ of the "quadratrix" curve OQQ', defined by some relation

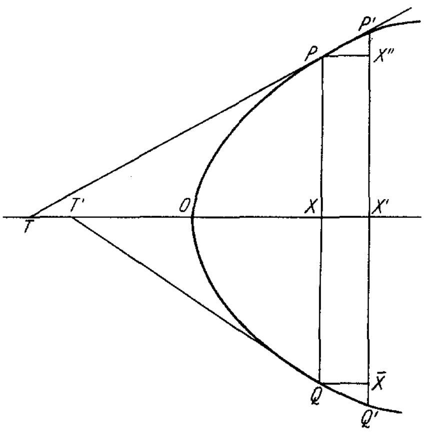


Fig. 107

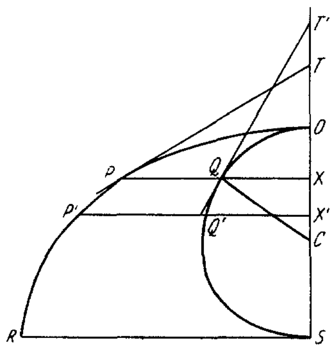


Fig. 108

$z = g(x)$ connecting OX and QX. Supposing $XT = t$ and $T'X = t'$ the subtangents to the two curves we can set up a modified "BARROW" rule which gives, equivalently,

$$\frac{y}{t}\left(= \frac{dy}{dx}\right) = F\left(x, \frac{z}{t'}\left[= \frac{dz}{dx}\right]\right)$$
$$= F'\left(x, \frac{dv}{dx}\right),$$

using $z = g(x)$ straightforwardly to derive the value of dv/dx. His example 3[39] will clarify this rather unwieldy general exposition: where $\widehat{OPR}$ is the cycloid, we can take the semicircle $\widehat{OQS}$ as the "quadratrix" curve: where C is the circle

[36] See his *opera mathematica* **2**: 391—396, where WALLIS inserts an outline of letters from NEWTON of the previous year (1692) in the Latin edition of his *Algebra*.

[37] Especially parts of the 1671 tract on analysis—compare *CUL Add.* 3962.1: 29V, marginal note: "annis abhinc quinque vel sex [sc. about 1686] cum D. Joannes Craige, Cantabrigiae diutius commoratus, seriem hic positam vidisset et quadraturam curvae quae (si recte memini) in exemplo hoc secundo habetur, cum Scotis suis per literas hinc datas communicasset, D. David Gregorius, Matheseos Professor Edinburgensis, acuto vir ingenio, eandem seriem sed minus concinnam alia methodo sane non ineleganti invenit".

[38] *tractatus mathematicus*: 44.

[39] *tractatus mathematicus*: 46.

centre (on OS) and taking $OX=x$, $XP=y$, $XQ=z$, $OS=2a$, $\widehat{OQ}=v$, we can define the cycloid by the representing equation $y=v+z$, where $z^2=x(2a-x)$; so that

$$\frac{y}{t}=\lim_{x'\to x}\left(\frac{y'-y}{x'-x}\right)=\lim_{x'\to x}\left(\frac{v'-v}{x'-x}\right)+\lim_{x'\to x}\left(\frac{z'-z}{x'-x}\right),$$

and, since LEIBNIZ' rules for differentiating yield immediately

$$\lim_{x'\to x}\left(\frac{z'-z}{x'-x}\right)\left[=\frac{dz}{dx}\right]=\frac{a-x}{(x(2a-x))^{\frac{1}{2}}},$$

where $z=(x(2a-x))^{\frac{1}{2}}$, therefore

$$\frac{y}{t}=\lim_{x'\to x}\left(\frac{v'-v}{x'-x}\right)\left[=\frac{QT'}{XT'}\right]+\frac{a-x}{z},$$

or

$$t=XT=\frac{yz}{(a-x)+\frac{QT'}{XT'}\times z},$$

where QT'/XT' is known.*

By the late 17th century the tangent-problem became entirely embedded in the general differentiation algorithms which were developed—NEWTON'S fluxional methods in England, and the LEIBNIZIAN differential methods on the continent—and further approaches to the construction of tangents which depended on an examination of a representing equation were discarded. It remains merely to give an outline of how this (and other aspects considered in previous chapters) became part of a general calculus algorithm, and this we will now consider in a final chapter.

XI. Calculus

4. Differentiation and integration as inverse procedures: the calculus as an algorithm

We cannot pinpoint a particular moment at which (or a particular person with whom) the concept of differential or integral calculus was born. Largely, of course, it depends on what we wish to allow into our definition, what standards we wish to introduce. If we include merely the notion of a limit-quotient, then the concept of infinitesimal calculus so defined is at least as early as ARCHIMEDES, and a large number of more or less general applications had been worked out by 1670: if, however,[1] we measure each particular advance made against standards of rigour and abstraction from particular models introduced in the 19th century, then no invention had yet been made in the period of a rigorously deductive structure of calculus theory. It is very tempting, nevertheless, to admit two criteria into a working definition (without excluding others); first, that differentiation and

* Indeed, substituting their respective line-lengths,

$$XT=\frac{PX\times QX}{CX+\frac{QT'}{XT'}\times QX[=CQ,=CS]}=PX\times\frac{QX}{SX}=PX\times\frac{OX}{QX},$$

which shows that PT is parallel to circle chord QO.

[1] As, for example, C. B. BOYER in his *Concepts of the calculus* (*op. cit.*, note 1 to chapter 8).

integration be seen as inverse procedures; and, secondly, that both be defined with respect to an adequate algorithmic technique. This last chapter will trace how these two criteria developed in some aspects of English fluxional calculus in the 17th century.

Historians[2] have given BARROW[3] credit for the first proof of the inverse nature of differentiation and integration in the geometrical model in which the operation of constructing the subtangent to one curve is seen as strictly inverse to that of finding the area beneath a second curve (whose ordinates are connected by an analytical function which is the derivative of that relating the ordinates of the first. BARROW, in fact, supposes[4] the curve $\widehat{SYY'}$ so defined from the given curve $\widehat{OPP'}$ —say convex up—that the ordinate XP of $\widehat{OPP'}$ is $1/R \times \text{area}(OXYS)$: then the tangent at P to $\widehat{OPP'}$ cuts off subtangent XT such that $PX:TX = XY:R$.*

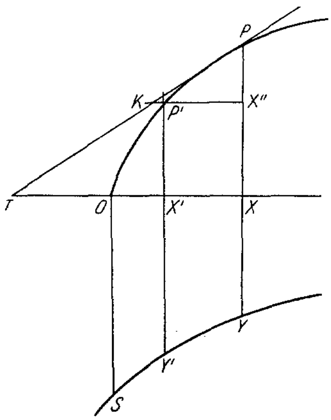


Fig. 109

Really, this is a neat amendment of GREGORY's generalization of NEIL's rectification method[5]. Essentially, what GREGORY had shown (by an equivalent use of the Greek tangent-concept) was that, where $\widehat{AKQ}$ is a given (convex) curve, and a second curve BNS is defined such that, where $IKMN$ is a general parallel, the tangent at K to $\widehat{AKQ}$ cuts off subtangent CI from AI, with $CI:CK = IM\ (=AB):IN$ then $AD \times \text{arc-length}\ \widehat{AKQ} = \text{area}\,(ABNSO)$—with the immediate converse that if the curves AKQ, BNS are defined such that $AB \times \widehat{AKQ} =$

* In proof, BARROW shows that, for $PX:TX = XY:R$, TP is tangent at P (as GREGORY, below); for consider any other ordinates $X'P'$, $X'Y'$ with, say, $X'P' < XP$ (and so $X'Y' > XY$, since curve OPP' is convex *up*), taking $P'X''$ parallel to X meeting TP in K: then $PX:TX = XY:R, = PX'':KX''$, or

$$KX'' \times XY = R \times PX''(= PX - P'X') = \text{area}\,(OXYS) - \text{area}\,(OX'Y'S)$$
$$= \text{area}\,(X'XYY'),\ > XY \times X'X\,(= P'X'');$$

so that, except when $P' \to P$, $KX'' > P'X''$. Similar proofs hold for the other cases.

[2] Especially J.M. CHILD in his *Geometrical lectures of Isaac Barrow*, Chicago, 1916, whose hypothesis is now widely accepted as "factual".

[3] Specifically, in his *LG*: lectio **10**: § 11: 78; and conversely in lectio **11**: § 19: 90.

[4] *LG* **10**: § 11: 78.

[5] GREGORY *GPU*: prop. 6—compare chapters 8, 10. A. PRAG, who first pointed out that this is a geometrical proof of inverseness (in *On James Gregory's "geometriae pars universalis"* GREGORY *TV*: 487—505)—without noting the connection with BARROW or that it is a generalization of the NEIL proof—adds (pp. 491—492): "... we have here a first proof of the Fundamental Theorem of the Integral Calculus. But tangent and area problems are still strictly geometrical. As long as differentiation and integration are not conceived of as operations, as calculating processes, we are hardly justified in describing or explaining old results by applying our modern abstract conceptions, which are based precisely on the idea of "operation" ... though Gregory is, indeed, ... very near to abstract thoughts."

area $(ABSO)$, then the subtangent CI corresponding to any point K has $CI:CK = AB:IN$. BARROW's treatment merely replaces the element of arc $\widehat{AKQ}$ by the elements of ordinate IK (and so derives the modified $CI:IK = AB:IN$), placing —for clarity—the curve BNS on the opposite side of AO.

This connection with a rectification method is no accident—rather, where the curve OP has corresponding analytical equivalents $OX = x$, $PX = y$, $PT = a'$,

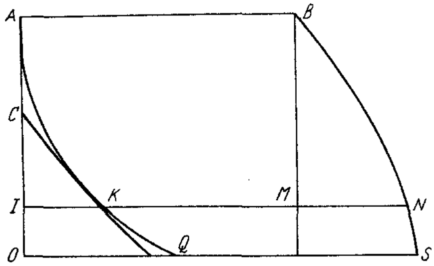


Fig. 110

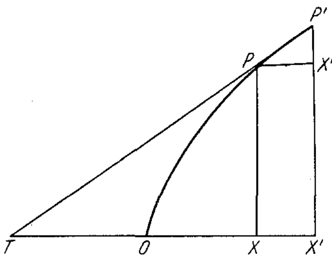


Fig. 111

$XT = t$, $\widehat{OP} = s$, and triangle $PP'X''$ is the differential triangle of the curve at P (with sides $PX:P'X:PP' = dx:dy:ds$), any rectification proof using the property

$$s'\left(= y\frac{ds}{dy}\right):t\left(= y\frac{dx}{dy}\right) = \frac{ds}{dx} = \frac{d}{dx}\left(\int_0^s ds\right)$$

is a proof of the inverse nature of the processes of integration and differentiation where, by suitable definitions, these are applied to a geometrical model. Since then the broad basis for NEIL's rectification of the semicubical parabola is to be found in WALLIS' *AI*,[6] there seems no reason why some credit for glimmerings of the concept of inverseness should not be given to WALLIS.*

But even more generally we can say that any geometrical theorem which in some way ties the subtangent at a point on a curve to the area of a simply defined corresponding curve has in it the germ of an inverseness proof with respect to a particular integral-differential form**. It may, however, be misleading to read concepts into a particular known proof-structure which were not present

* Why not, indeed, go back to the first historical example of a rectification theorem—that of ARCHIMEDES' *On the sphere and cylinder*, Book 1—which equates the surface of a sphere with the curved surface of the circumscribing cylinder?

** TORRICELLI's (? 1645) proof, using an exhaustion-approach, of the constancy of the subtangent to the logarithmic curve, $y = \log(x)$, by tying it to the defined portion of the area under the curve is a fine case in point.[7]

[6] *AI*: prop. 38, scholium: 28—31, where he states that the rectification problem is solved if we can find some way of summing the infinitesimal arc-length elements ds. This basic assumption that $ds = ((dx)^2 + (dy)^2)^{\frac{1}{2}}$ in the limit is restated clearly in his *epitome binae methodi tangentium* ..., PT **7** (1672): 4010—4016: (p. 4013): "A second method of tangents (following the exposition of my *de angulo contactus* and *arithmetica infinitorum*) views the curve as conflated out of particles infinitely slight but having a known position, and as equivalent (since the angle of contact has no measure or is infinitely small) with the tangent to the curve at the point, and so having an equal slope ..."

[7] See E. TORRICELLI: *opera*, Faenza, 1919: 1.2: 337—347: *de hemhyperbola logarithmica*; and compare G. LORIA: *Le ricerche inedite di Evangelista Torricelli sopra la curva logarithmica*. Bibliotheca mathematica$_3$ **1** (1900): 75—81.

consciously in the author's mind. Certainly the historical influence of an idea depends largely on its conscious recognition, and it remains bluntly factual that no one in the mid-17th century seems to have seen in even the very general BARROW theorem more than a subtle theorem on the relation of properties of two curves[8]. Only many years later, in the first priority disputes over the new calculus, do we find some acknowledgement of the generality of BARROW's work.[9]

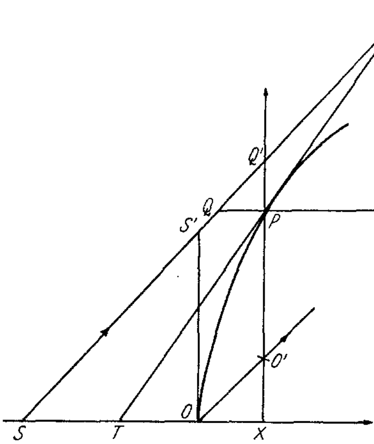


Fig. 112

Perception of the dual nature of integration and differentiation processes had, however, arisen in a slightly different form in DESCARTES' correspondence with DEBEAUNE in the 1630's, culminating in the famous letter of DESCARTES of 20 February 1639[10]. Apparently DESCARTES had in previous letters proposed several problems on curves defined by subtangent properties, and in particular the following: what is the curve OP such that, where TX is the subtangent of the point P and $O'X$ is taken on PX equal to OX, $PX:TX=\alpha:PO'$ for some given magnitude α?[11] Analytically this condition is, where $OX=x$, $PX=y$, $TX=t$, $\frac{y}{t}\left[=\frac{dy}{dx}\right]=\frac{\alpha}{y-x}$, and DESCARTES, introducing transforms which are, in effect $\sqrt{2}y=y'$, $z=\alpha-(y-x)$, derives the equivalent of $\frac{dz}{dy'}=-\frac{z}{\sqrt{2}\alpha}$ (which shows that the curve has a constant subtangent $t'=z\frac{dy'}{dz}=-\sqrt{2}\alpha$, and so is logarithmic).*

* Or so we assume. There is an easy proof: by taking $SO=\alpha$ in OX, and SV parallel to OO', with QPR parallel to OX (meeting VR, parallel to PX, in R), then $PX:TX=SO:PO'(=PX-OX)$, $=VR(=QR):PR$ $(=QR-QP)$; but $SO+OX=QP(=Q'P)+PX$, or $SO:(PX-OX)=(QP+PX-OX):(PX-OX)$; so that $QP:(PX-OX)=QP:(VR-QP)$, or $VR=QP+PX-OX=SX-OX=SO$, and so $QV=SS'$, where OS' is tangent at the vertex O $(=\sqrt{2}\times SO$ where $\widehat{SOS'}$ is right).

[8] There is, indeed, disappointingly little evidence either way relating to the acceptance of BARROW's *LG* by his contemporaries. We know that *LG* sold very badly—remaining copies were apparently later put on the market at a nominal price when the publisher went bankrupt, and apparently many were pulped (see COLLINS' letters to BAKER of 10 February 1676/7 ≡ RIGAUD (*C*), **2**: 14—15; and of 24 April 1677 ≡ RIGAUD (*C*), **2**: 20—22).

[9] For example, JAMES BERNOULLI (in AE (Jan. 1691): 91) accused LEIBNIZ of deriving his fundamental ideas on the calculus from BARROW's *LG*. LEIBNIZ had, indeed, bought a copy in 1673—it still exists in Hanover Royal Library—but the claim is satisfactorily refuted by J.E. HOFMANN: *Entwicklungsgeschichte der Leibnizschen Mathematik* ..., München, 1949, who points out that, anyway, all the important general theorems of the *LG* are modifications and adaptations of equivalent ones in GREGORY's *GPU*.

[10] Printed in *Oeuvres* (ed. ADAM & TANNERY) **2** (Paris, 1898) 510—519. Compare P. TANNERY: *Pour l'histoire du problème inverse des tangentes* ≡ *Mémoires scientifiques* **6** (Paris, 1926): 457—477; and G. MILHAUD: *Descartes savant*, Paris, 1921: 169—175.

[11] See *Oeuvres* **4**: 229, which prints a Latin letter of DESCARTES to an unknown correspondent in June 1645.

The problem remains to identify the relation between abscissa $SQ(=-y')$ and ordinate $QP\,(=z)$ which determines the curve, and in resolving it DESCARTES introduces the idea of defining a curve as the point-set of the meets of the tangents at two indefinitely near points, say K, K', of the curve[12]: taking tangents KLM, $K'L'M'$ (cutting as shown), we can show, since $HL'<HK$, $H'L<H'K'$ (by the provable convexity of the curve), that

$$\frac{HK-H'K'}{HK}<\left(\frac{HK-H'L}{HK}=\right)\frac{HH'}{HM}=\frac{HH'}{H'M'}\left(=\frac{HL'-H'K'}{H'K'}\right)<\frac{HK-H'K'}{H'K'};$$

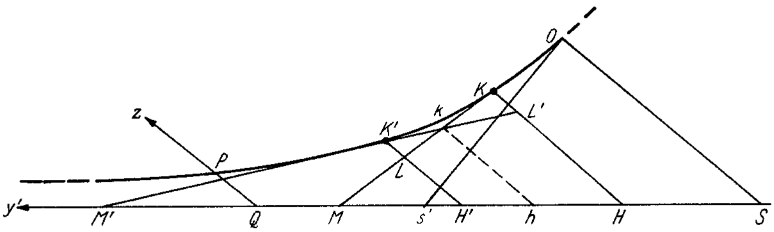


Fig. 113

so that, supposing $OS{:}HK{:}H'K'{:}PQ=mr{:}n{:}(n-1){:}ms$, $(r>s)$, with $SS'=HM=H'M'=\lambda\left(=-\sqrt{2}\,\alpha\right)$, it follows that $\frac{1}{n}<\frac{HH'}{\lambda}<\frac{1}{n-1}$; and finally

$$\sum_{ms+1\leqq n\leqq mr}\left(\frac{1}{n}\right)<\sum_{ms+1\leqq n\leqq mr}\left(\frac{HH'}{\lambda}\right)<\sum_{ms+1\leqq n\leqq mr}\left(\frac{1}{n-1}\right),$$

or, since

$$SQ=\sum_{ms+1\leqq n\leqq mr}(HH'),\quad \sum_{ms+1\leqq n\leqq mr}\left(\frac{1}{n}\right)<\frac{SQ}{\lambda}<\sum_{ms\leqq n\leqq mr-1}\left(\frac{1}{n}\right).$$ [13]

From a conceptual viewpoint what is significant in this sketch-proof of DESCARTES is that, under given conditions (furnishable by MENGOLI'S theory of the logarithm), a rigorous proof is suggested, for the particular curve whose representing equation is $y=k\cdot\log(x)$, that $\int_{y=0}^{y=y}(dy/dx)\cdot dx=y$. Historically, however, DEBEAUNE seems to have mislaid the letter—or perhaps chose to keep it secret—and it was not made public till the 1660's[14], when both NEWTON and LEIBNIZ

[12] DESCARTES writes, specifically (511 ff.): "... en considerant quelle doit estre cette courbe afin que cette intersection se fasse tousiours entre ces deux points, et non au decà ny au de là, on en peut trouver la construction; mais il y a tant de divers chemins à tenir, et ie les ay si peu pratiquez, que ie n'en sçaurois encore faire un bon conte."

[13] Which, by MENGOLI'S definition of logarithm proves $\frac{SQ}{\lambda}=\log\left(\frac{r}{s}\right)=\log\left(\frac{OS}{PQ}\right)$; though DESCARTES leaves the matter there, contenting himself with a motion-construction of the curve which follows directly from the constant subtangent property, $z\frac{dy'}{dz}=\lambda$.

[14] CLERSELIER printed it for the first time in his 1667 edition of DESCARTES' correspondence (3: 409—416). It is interesting, however, that it was almost immediately seized upon by LEIBNIZ as an example of integration by a straightforward process of inverse differentiation (and not by setting up a limit-sum), and, as such, he sent it to NEWTON for comment (see LEIBNIZ' letter to OLDENBURG of 27 August 1676 · ≡ ·

had come (or were beginning) to accept differentiation and integration as inverse analytical operations, so that it could seem already no more than a historical curiosity.

NEWTON and LEIBNIZ, in fact, derived the inverse property from their new (but respectively equivalent) definitions of the basic concepts. At least as early as 1673 LEIBNIZ, viewing the integral as the limit-sum $\int_{x_1}^{x_2} y \cdot dx = \lim_{n\to\infty} \sum_{0\leq i\leq n} (y_i(x_{i+1}-x_i))$, where the points x_i are taken in some interval X_1, X_2 sufficiently densely (that is, such that for all i any arbitrarily chosen neighbouring pair x_i, x_{i+1} are indefinitely close to each other), and the derivative as the limit-quotient $\frac{dy}{dx} = \lim_{x'\to x}\left(\frac{y'-y}{x'-x}\right)$, where x, y are related by some *relatio*), saw that for

$$z_i = \int_k^{x_i} y \cdot dx = \lim_{x_{j+1}\to x_j}\left(\sum_{k=x_j\leq X_i}[y_j(x_{j+1}-x_j)]\right),$$

$$\frac{dz_i}{dx} = \lim_{x_i'\to x_i}\left(\frac{z_i'-z_i}{x_i'-x_i}\right) = \lim_{x_{i+1}\to x_i}\left(\frac{y_i(x_{i+1}-x_i)}{x_{i+1}-x_i}\right) = y_i.$$ [15]

NEWTON, however, had derived an equivalent approach—far more geometrically slanted—by 1666[16] and it is absorbing to follow his concept through successive manuscript drafts[17]. Reformulating his approach[18] slightly—it is stated in the typical limit-motion model—, we can say that NEWTON considers a curve OP defined by some function $y=f(x)$ connecting $OX=x$ and $XP=y$, and a second function z (his "area" $(OXPO')$), where $z=\int_0^x y\cdot dx$, $=g(x)$ say: then the derivative of $g(x)$ is $\lim_{x'\to x}\left(\frac{z'-z}{x'-x}\right)$, where $OX'=x$, $X'P'=y'$ and $z'=\text{area}\,(OX'P'O)$

GERHARDT: *Briefwechsel* ... 1: 193—200, especially 200: GERHARDT indeed, quotes, pp. 201—203, a LEIBNIZ manuscript of July 1676, *methodus tangentium inversa*, which shows how deeply LEIBNIZ pondered DESCARTES' sketch); NEWTON replied in his famous letter of 24 October 1676 to OLDENBURG (see GERHARDT: *Briefwechsel* 1: 203—225, especially 224).

[15] Compare the various tracts of C.I. GERHARDT translated in J.M. CHILD: *The early mathematical manuscripts of Leibniz*, Chicago, 1920; and J.E. HOFMANN: *Entwicklungsgeschichte der Leibnizschen Mathematik ... in Paris 1672—76*, München, 1949; *passim*.

[16] In a manuscript written during the fluxion dispute (about 1714) NEWTON wrote: "I found the method (of fluxions) by degrees in the year 1665 and 1666. In the beginning of the year 1665 I found the method of approximating series and the rule for reducing any dignity of a binomial into such a series. The same year in May I found the method of tangents of Gregory and Slusius, and in November had the direct method of fluxions, and the next year in ... May ... I had entrance into y^e ... inverse method of fluxions ..." (see *CUL Add.* 3968. 41: 86).

[17] In *CUL Add.* 4004: 7V—57V, especially 57R—57V (dated 13 November 1665): *Add.* 4000 (undated summaries); and *Add.* 3958: Section 3, especially 47V—63V (which is the final draft of October 1666 *On resolving problems by motion*). The *de analysi* of 1669 (now in the Library of the Royal Society) is a mere fragment of these manuscripts, suitably systematised.

[18] Compare, for example, *Add.* 3958. 3: 57R "Problem 5^t. To find y^e nature of y^e crooked line whose area is expressed by any given equation. That is, y^e nature of y^e area being given, to find y^e nature of y^e crooked line" (reworked, of course, —in outline—in *de analysi*, where it is the basic proposition).

correspondingly. But, as $P'X' \to PX$ area $(XX'P'P) \to PX \times XX'$ $\big(= y\,(x'-x)\big)$, and so

$$\lim_{x' \to x}\left(\frac{g(x')-g(x)}{x'-x}\right) = \lim_{x' \to x}\left(\frac{\text{area }(XX'P'P)}{XX'}\right) = PX \; (=y).$$

But before such considerations were introduced, the inverse nature of the two processes must have long been accepted in wide classes of known results. Thus, by indivisible and exhaustion techniques, the area under the curve $y=f(x)=x^m$ (defined in a CARTESIAN coordinate system) had been shown to be

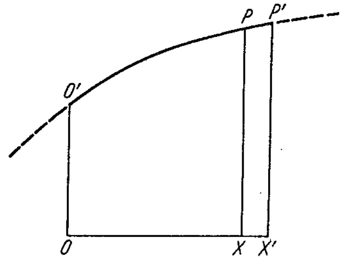


Fig. 114

$$\frac{1}{m+1}\,x^{m+1}\left(=\int_0^x x^m \cdot dx\right),$$

while from the tangent methods developed $y/t\,(=dy/dx)=m\cdot x^{m-1}$, where t is the corresponding subtangent. This simple result, derived without appeal to an abstract general argument, was in fact sufficient to justify almost all the practical uses made of calculus procedures, since it is sufficient to derive both the integral and derivative of the polynomial $\sum\limits_{0\leq i\leq n}(a_i x^i)$.

Clarification, however, on a general basis of the inverse nature of the two procedures opened the way to the derivation of algorithmic processes on a vast scale, and the crystallization from such processes of a rigorous basic calculus structure was, in a strong sense, inevitable (if not over rapidly forthcoming). Indeed, if indivisible methods and exhaustion techniques were typical of the mid-century, the rather rambling but useable compilations of algorithms and calculus methods[19] represent the rapidly widening exact knowledge of the 1700's.

Inevitably at first much time was spent in arranging in more convenient shape results obtained before uniform, standardized treatment was possible, and in this reformulation there was felt urgently the need for a suitable notation in which to systematize and generalize the comparatively cumbersome form of the early geometrical results. The operation of integration ("quadration") demands a distinguishing mark in some way equivalent to "Integrate!" together with indication of the variable (or variables) over which the integration is performed and of the integration range; and similarly for the operation of differentiation. Not surprisingly the first calculus notations are clumsy*, but what is lacking in the notation itself is to be found in the accompanying verbal text or to be filled in by common-sense assumption (much as later, with EULER, we assume that the lower bound of the range of his integrals is 0, or 1, or some other obvious number).

* Even the modern forms are criticized—for example, by KARL MENGER—as not being very satisfactory.

[19] Especially L'HOSPITAL'S *Analyse des infiniment petits*, Paris, 1696 (strongly based on then unpublished work of JOHN BERNOULLI), the first textbook of LEIBNIZ' differential calculus; L. CARRÉ'S *Méthode pour la mesure des surfaces, la dimension des solides ... par application du calcul intégral*, Paris, 1700; and—the first comprehensive printed fluxion text—CHARLES HAYES' *A treatise of fluxions*, London, 1704. An even greater volume of material existed in manuscript (to be printed only fitfully at a later date).

A symbol for integration, "O" (for "omnes") had already been introduced by MENGOLI[20] as the limit-form of the same symbol used for finite summation, and this approach was followed by LEIBNIZ in his "$\int$" (for "summa") integration symbol—later improved by indicating the bounded variable, as "$\int : [\;]\, dx$"[21]—which is the direct ancestor of the modern form. NEWTON himself in his early manuscripts[22] widely uses the ideograph "$\square$" (for "area (under)"), though in his later printed works prefers almost exclusively to indicate the operation of integration by suitable verbal phrasings. Both the NEWTON and LEIBNIZ forms for differentiation—"$\dot{y}/\dot{x}$",* "dy/dx" respectively—arise naturally from the geometrical model of the differential triangle, which determines the slope of a curve at some given point; that is, as the tangent of the gradient angle $P'PX''$, $P'X''/PX''$, where PX'', $P'X''$ are the limit-increments of $OX = x$ and $PX = y$. (Indication of the particular point in the variable range at which the differential operation is carried out remains, however, a 19^{th} century innovation.) The great advantage of using explicit symbolism was, of course, that a great deal of automatic thinking could be off-loaded on to the notation, while there are thereby created patterns of plausible truth (usually at a visual level) which offer exciting possibilities for further advance, even if in fact misleading.

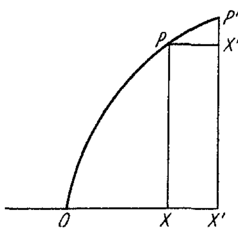


Fig. 115

Largely, the period 1665—1690 is one of consolidation, rather than innovation, in calculus procedures, and a great deal of attention was paid to the derivation of basic algorithms, especially $D(x+y) = D(x) + D(y)$** and the analogous $D(xy) = xD(y) + yD(x)$ $\left(\text{or } \frac{D(xy)}{xy} = \frac{D(x)}{x} + \frac{D(y)}{y} \text{ in its logarithmic form}\right)$. Curiously, it is in this latter form that we can derive the second from BARROW's LG[24]: thus, where the curves $P\alpha$, $P\gamma$ and line-direction PX are given, define the curves $P\beta$, $P\delta$ such that $L\beta^2 = L\alpha \times L\gamma$, $= LL' \times L\delta$ for any $L\alpha\beta\gamma\delta$ parallel to PX and consider an arbitrary fix-line PL' through P with $XLABCD$ a general parallel to it; it follows easily, since $\frac{\alpha\beta}{\beta\gamma} = \frac{L\beta - L\alpha}{L\gamma - L\beta} = \frac{L\alpha}{L\beta} = \frac{L\beta}{L\gamma}$, that as $LL' \to XP$ the limit-ratio of $\frac{\alpha\beta}{\beta\gamma}$ is unity, or where LL' meets the tangents to $P\alpha$, $P\beta$, $P\gamma$, $P\delta$

* This seems, in fact, a comparatively late NEWTONIAN form: in the middle 1660's NEWTON was playing tentatively with the form "$\ddot{p}$" $= x\frac{dp}{dx}$, but more commonly takes dy/dx by p/q, where p, q are the respective "limit-speeds" of x, y.[23]

** Where D is the differential operator.

[20] In his *geometria speciosa*, Bologna, 1659: especially Book 6.

[21] Compare the works quoted in note 15 above, especially J.M. CHILD.

[22] *CUL Add.* 4004: 7V—57V: *Add.* 4000: *passim*; *Add.* 3958: Section 3.

[23] Though at *Add.* 3958. 2: 30V (dated in verso "October 30th 1665") he appears to use "$\ddot{p}$" $= dp/dx$. The p, q notation is used exclusively in the 1671 tract on fluxion applications (*Add.* 3960: 14 which is HORSLEY's *geometria analytica*, and COLSON's *Method of fluxions* except that they are rewritten in dottage notation), and the matter is complicated by a third notation which he seems to have been experimenting with at the time of *PM*, "$\boxed{A}$" (or dA)—thus $\boxed{AB} = Ab + Ba$, where a, b are the fluxions of A, B (see fragments in *Add.* 3960).

[24] *LG*: lectio 9: §§ 10, 12; 73—see J.M. CHILD: *The geometrical lectures of Isaac Barrow*, Chicago, 1916: 111.

in $a, b, c, d, ab = bc$, so that, projecting $(\infty_{LL'} abc)$ into $(XABC)$, $\frac{2}{XB} = \frac{1}{XA} + \frac{1}{XC}$[25]; further, BARROW shows[26] that $LL' \times L\delta = L\beta^2$ implies $XB = 2 \times XD$; or, finally, $\frac{1}{XD} = \frac{1}{XA} + \frac{1}{XC}$, from which the logarithmic form of the product-derivative follows by taking the subtangents XA, XC, XD in analytical form.*

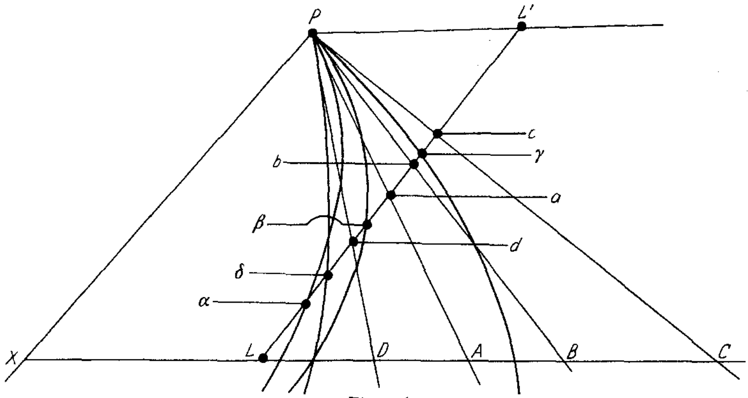


Fig. 116

This is not a very pretty derivation—if, indeed, it was at all consciously present in BARROW's mind—when we compare it with the standard analytical proof developed by LEIBNIZ: where $z = xy$,

$$\frac{dz}{dt} = \lim_{t' \to t}\left(\frac{x'y' - xy}{t' - t}\right) = \lim_{t' \to t}\left(x'\left(\frac{y' - y}{t' - t}\right) + y\left(\frac{x' - x}{t' - t}\right)\right) = x\frac{dy}{dt} + y\frac{dx}{dt}.$$

In a balder form, which is virtually the FERMAT tangent rule, the argument goes:

$$\left.\begin{array}{l} z = xy \\ z' = x'y' \end{array}\right\} \text{ implies } \begin{array}{l} z + (z' - z) = \big(x + (x' - x)\big)\big(y + (y' - y)\big) \\ \qquad = xy + y(x' - x) + x(y' - y) + (x' - x)(y' - y), \end{array}$$

and in the limit we can "blot out" the last term. It is in trying to avoid this difficult concept of discardable infinitesimals that NEWTON develops his "cheat" proof of *PM*[27]: if $z = xy$, and x, y, z are increased by the limit-increments dx, dy, dz in the same indefinitely small period of time, then

$$\left.\begin{array}{l} z + \frac{1}{2}dz = (x + \frac{1}{2}dx)(y + \frac{1}{2}dy) \\ z - \frac{1}{2}dz = (x - \frac{1}{2}dx)(y - \frac{1}{2}dy) \end{array}\right\}, \quad \text{or, subtracting,} \quad dz = x\,dy + y\,dx.^{**}$$

* Specifically, with respect to some common abscissa t,

$$\frac{1}{z\frac{dt}{dz}} = \frac{1}{x\frac{dt}{dx}} + \frac{1}{y\frac{dt}{dy}}, \quad \text{where } z = xy.$$

** The basic fallacy is that, if $u \leftrightarrow v$ by $u = \Phi(v)$ and $u + du \leftrightarrow v + dv$, in general it is not true that $u - du \leftrightarrow v - dv$. Specifically $\Phi(v - dv) = u - du$ implies $\Phi(v - dv) = \Phi(v) - (\Phi(v + dv) - \Phi(v))$, or $2 \cdot \Phi(v) = \Phi(v - dv) + \Phi(v - dv)$, which restricts $\Phi(v)$ to being linear, $\Phi(v) = \alpha v + \beta$, for α, β constants—and in that case the product-derivative theorem is trivial anyway.

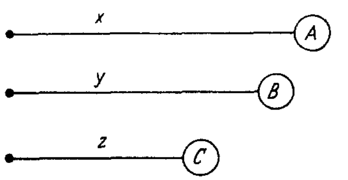
NEWTON's proof gives a line of slope *parallel* to the tangent at a point of the parabola in the particular case of $y = x$ (when the objection is not valid), but NEWTON still has to *prove* this and certainly cannot assume it.

[25] This is the BARROW cross-ratio theorem of chapter 6 with $n = 2$, $m = 1$.

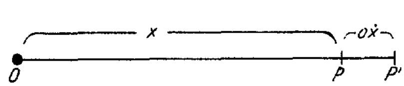
[26] *LG*: lectio 9: 1–70; and compare chapter 10.

[27] NEWTON, *PM*: Book 2: Section 2, lemma 2.

NEWTON, as so many other 17th century mathematicians, was on firmer ground when arguing from a geometrical model, and, in fact, his earliest ideas on fluxions are developed with respect to the basic idea of the motion of a generating element (usually a point moving in a line). Thus, in a draft of 13 November 1665 [28], the problem is posed: "An equation being given expressing y^e^ relation of two or more lines, x, y, z etc. described in y^e^ same time by two or more moving bodies A, B, C etc. to find the relation of their velocities p, q, r etc." In the resolution he assumes that in some time-unit O (which in the limit is taken indefinitely small as his "moment of time") $x, y, z, \ldots$ are increased by $p\times o, q\times o, r\times o, \ldots$ where $p, q, r, \ldots$ are the instantaneous speeds of $A, B, C, \ldots$ at $x, y, z, \ldots$ respectively; so that, if $f(x, y, z, \ldots)=0$ holds, then $f(x+op,\ y+oq,\ z+or, \ldots)=0$ also holds, and by some appropriate reduction technique from the form $f(x+op,\ y+oq,\ z+or, \ldots) - f(x, y, z)=0$ we will be able to derive the "differential" relation which holds between $p, q, r, \ldots$ Specifically NEWTON uses a FERMATIAN reduction: "Hence may be observed first, y^t^ those terms ever vanish in w^ch^ o is not because they are y^e^ propounded equation. Secondly y^e^ remaining equation being divided by o those termes also vanish in w^ch^ o still remaines because they are infinitely little. Thirdly y^t^ y^e^ still remaining termes will ever have y^t^ forme w^ch^ by y^e^ first preceding rule* they should have (as may partly appear by OUGHTRED's Analytical Rule)**..."



In later more systematized form [29] he introduces the now familiar concept of "fluxion"—that is, the instantaneous speed of a point P which moves along a



line. Where the length of OP is represented by the analytical measure x, the limit-segment PP' as $P'\to P$ is representable by $\lim_{o\to \text{zero}} (o\,\dot{x})$, where $\dot{x}$ is the instantaneous speed of the point at P. From this the definition of the fluxion quotient $\dot{y}/\dot{x}$, where some relation f connects x, y by $y=f(x)$, is immediate: for $y+o\dot{y}=f(x+o\dot{x})$, or $\frac{\dot{y}}{\dot{x}}=\frac{o\dot{y}}{o\dot{x}}=\lim_{0\to\text{zero}}\frac{f(x+o\dot{x})-f(x)}{(x+o\dot{x})-x}\left[=\frac{dy}{dx}\right]$. Often, too, he introduces a simplification in which x is seen as the time-continuum (and so $\dot{x}$ is constant and may be taken as the unit, since time "flows" uniformly), or $\dot{y}\,(=\dot{y}/\dot{x})$ can then be taken to represent the fluxional rate of increase of y. As before, translation to the geometrical model is immediate: using a basic time-scale measured by t (where t is taken as unit-measure) define $g(y)$ as the fluxion

* This rule is merely the generalized HUDDE-SLUSIUS rule for the function $f(x_1, x_2, \ldots, x_n)=0$—that is $\sum_{i\leq i\leq n}\left(\frac{\partial f}{\partial x_i}\times\frac{dx_i}{dt}\right)=0$.

** That is, the "PASCAL" triangle of binomial coefficients, which is OUGHTRED's "Tabula Posterior" (see *clavis mathematicae*, $_3$1652: 37).

[28] *CUL Add.* 4004 (the *Waste Book*): 57R, revised in the October 1666 manuscript (*Add.* 3958.3) ≡ prob. 7: 49R.

[29] Apparently, however, not before the middle 1670's, but established by the 1680's.

of the relation $f(y)$, or, inversely, $f(y)$ as the "fluent" of $g(y)$; and we can then[30] represent "the fluents of quantities of any kind by the areas under curves, the fluxions by the ordinates, the time-interval by the abscissa, the limit-moment of time by the limit-moment of the abscissa, and the moments of other fluents by ordinates to the corresponding limit-moment of the abscissa"—that is, where $OX=t$, $PX=y=\varphi(t)$, and $XX'=\dot{t}o$, $=o$ (since $\dot{t}$ is taken $=1$) and $P'X''$ $(=P'X'-PX)=\dot{y}o$, the fluent is the area $oPX=z=\square\, y$ under $y=\varphi(t)$, and the fluxion of the fluent is the abscissa $PX=y=\dot{z}\left(=\frac{\dot{z}o}{\dot{t}o}\right)$.

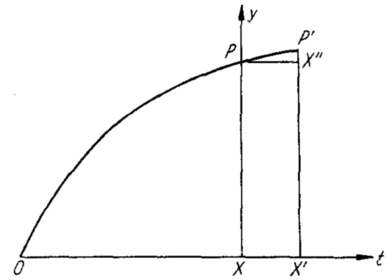


Fig. 117

NEWTON returned to such considerations very late in life when, sometime after 1690, his interest in pure geometry revived, and—apparently still dissatisfied with the doubtful rigour of a purely analytical treatment—he tried to give fluxion theory a strict geometrical basis in his projected *geometria curvilinea*[31]. Here, after a lengthy introduction in which he stresses the supremacy of a synthetic (geometrical) method in deriving a rigorous structure of mathematical theory, he gives[32] two important definitions:

1. "fluens est quod continua mutatione augetur vel diminuitur",

and

2. "fluxio est celeritas mutationis illius";

and two fundamental postulates:[33]

1. "lineas quasvis quacunque ratione geometrice moveri. per rationem geometricam intelligo talem rationem movendi in qua quaevis positio lineae motae potest geometrice designari",

2. "alias lineas per puncta vel intersectiones priorum discribere".

On this basis and using simple geometrical models* he is able to derive quite powerful results—for example, his proposition 2 shows that, where $A:B=B:C$ with A constant, B, C variable, then fluxio B: fluxio $C=A:2B$ (which is immediate in the analytical theory if we parametrize by $A=1$, $B=x$, and so $C=x^2$, since $\frac{\text{fluxio}(B)}{\text{fluxio}(C)}=\frac{d(x)}{d(x^2)}=\frac{1}{2x}=\frac{A}{2B}$).

The whole development is intimately connected with the general concept of point-correspondences which filled his mind at that time[34], and it is a pity that

---

* In particular, that of a right angle whose vertex is centred on a fix-point round which it moves, and in its motion its two arms intersect a given fix-line.

[30] Compare *CUL Add.* 3968: 246R.

[31] He never seems to have written more than its first part (and that in an incomplete fashion) but, together with several minor redrafts and predrafts, a representative section of Book 1 exists in *CUL Add.* 3963: 46R—60V.

[32] *Add.* 3963: 47R.

[33] *op. cit*; 48R.

[34] See chapters 6, 7.

we do not have NEWTON's work in complete form. It is beyond doubt, however, that his contemporaries had no knowledge of such work, all publications tending more and more to the analytical. However interesting, inspiring and provocative NEWTON's geometrical theories of fluxions might be, they quickly passed unnoticed into oblivion, except for the few traces which crept into his *PM*[35]—and even they were discounted later, rather unjustly, as trivial geometrical applications of a basic analytical theory of fluxions.

On the whole later 17th century mathematicians were content to accept the logical basis, however unrigorously defined, and devoted their energies increasingly to exploring and expanding particular aspects of the calculus ("fluxion theory")*. Above all, with the introduction of analytical methods the way was clear for generalized treatments, and in particular for extension of the primitive concept of derivative differential to that of n^{th} order derivative, full and partial. In Germany, France and Switzerland, LEIBNIZ and the BERNOULLIS, discarding the geometrical model of curve-area and curve-tangent completely (except as a particular application of the general method), developed a purely analytical approach. NEWTON, however, —the only English contemporary of equal creative mathematical power—still preferred to keep more or less closely to the geometrical model, using it as a basis for his definitions and concepts. Using such a model n^{th} order derivatives are introduced by defining a succession of line-lengths, curve-areas and other elements in terms of previously defined ones—an approach that makes for clumsiness, and it is to NEWTON's credit that many times it does not appear so.

Thus, let us consider his development of the concept of curvature in detail. We can see NEWTON's ideas on the subject developing in a series of manuscript drafts from late 1664[38] to their systematisation in the important manuscript of October 1666 on resolving problems by motion,[39] where the problem is resolved:

"Resolution. ffind that point fixed in y^e crooked lines perpendicular w^ch is y^n least motion, for it is y^e center of a circle w^ch passing through y^e given point is of equal crookednesse with y^e line at y^t given point. Now since y^e crooked lines tangent and perpendicular etc. (at y^t moment) circulate about y^t center, I observe y^t every point fixed in y^e tangent or perpendicular, or whose position to y^m is determined, doth describe a curve line to w^ch y^e right line drawne from y^t center is perpendicular, and is also y^e radius of a circle of equall crookednesse with it: 2dly, y^t y^e motion of every such point is as its distance from y^t center: and so are y^e motions of y^e intersection points in w^ch any radius drawn from y^t center intersects two parallel lines."

* It is revealing that in the next century Bishop BERKELEY's overharsh and slightly misleading criticism of fluxions as "ghosts of departed quantities"[36] resulted in no immediate strengthening in rigour of presentation, but rather—more by counter-reaction—inspired COLIN MACLAURIN's beautifully systematised account of existing techniques.[37]

[35] PM_1: Book 2: Lemma 2: 224ff.

[36] Compare C.B. BOYER, *op. cit.* (note 1, chapter 8): ch. 6, especially 224ff.

[37] *Treatise of fluxions*, Edinburgh, 1742.

[38] Compare *Add.* 4004: 30V—33V (drafts from December 1664 to May 1665).

[39] *Add.* 3958, Section 3: 48R—63V, especially problem 2: 54V: *To find y^e quantity of crookednesse of lines.*

In the immediately following example 1 a procedure is sketched for calculating this radius of curvature at a point. His argument, obscurely phrased in terms of speeds of moving points—in particular $d^2y/dx^2 = dz/dx$, where $z = dy/dx$, is expressed in the concept of "velocity of y^e increase of y^e motion of ..."—is, in fact equivalent to a differential triangle method, and I will give it in that form. Take the curve ac defined by some relation between ordinate $bc = y$, and abscissa $ob = x$, and then draw one tangent at cn with $ec = cn$, cm the radius of curvature at the point c on the curve and eg, parallel to cm, meeting cg, parallel to ab, in g. Then consider the differential triangle $ce'f'$, where ce', cf' are the limit-increments of ce, cg: clearly $ng = p = y\dfrac{dx}{dy}$, $bd = v = y\dfrac{dy}{dx}$ with $cf' = dx$, $e'f' = dy$, or $cg' = cf' + f'g'\left(= \dfrac{e'f'^2}{cf'}\right) = \dfrac{dx^2+dy^2}{dx}$; again, taking cf perpendicular to cfg, we can show (by congruency) that $cf = nb$, $cg = bd$, or $cg = y\,\dfrac{dx^2+dy^2}{dx\cdot dy}$, and we can see the triangle ceg as "expanded" from $ce'g'$ by the proportion factor y/dy; and, finally, since $dk : dk' = cg : cg'$, $dk = \dfrac{y}{dy}\times dk'$. Now dk' is the limit-increment of od, or dk' = limit-increment of $ob(=dx)$ + "velocity of increase of d from $b''(=dv)$; so that $dk\left(= \dfrac{y}{dy}(dx+dv)\right) = y\left(\dfrac{dx}{dy}+\dfrac{dv}{dy}\right) = p + r$, where $r = y\dfrac{dv}{dy}$. The rest is immediate: since $\dfrac{cg-dk}{cd} = \dfrac{cg}{cm}$ and $\dfrac{cg-dk}{cb} = \dfrac{cg}{c\lambda}$,

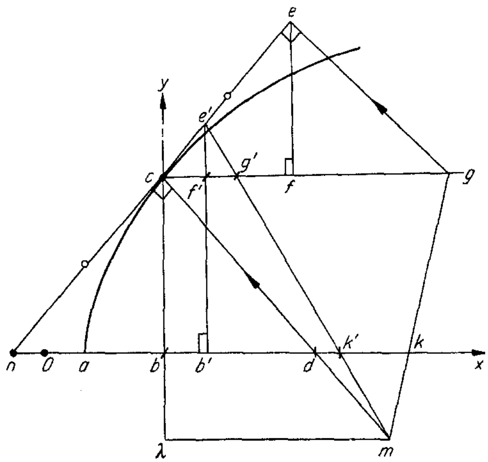


Fig. 118

$$cm = \frac{cd\times cg}{cg-dk} = \frac{(y^2+v^2)^{\frac{1}{2}}\times(p+v)}{v-r},$$

and

$$c\lambda = \frac{cb\times cg}{cg-dk} = \frac{y\times(p+v)}{v-r}.\ \star$$

This argument, especially as it relates to second order derivatives, is improved in the 1671 tract on analysis[40] where d^2y/dx^2 is now introduced by defining

---

⋆ And, expanding in terms of $p = y\dfrac{dx}{dy}$, $v = y\dfrac{dy}{dx}$, $r\left(= y\dfrac{dv}{dy}\right) = v + y^2\dfrac{dx}{dy}\dfrac{d^2y}{dx^2}$ or $p + v = y\dfrac{dx}{dy}\left(1+\left(\dfrac{dy}{dx}\right)^2\right)$, $v - r = -y^2\dfrac{dx}{dy}\dfrac{d^2y}{dx^2}$ and $y^2+v^2 = y^2\left(1+\left(\dfrac{dy}{dx}\right)^2\right)$, we have the more usual forms $cm = -\dfrac{\left(1+\left(\dfrac{dy}{dx}\right)^2\right)^{\frac{3}{2}}}{d^2y/dx^2}$, $c\lambda = -\dfrac{1+\left(\dfrac{dy}{dx}\right)^2}{d^2y/dx^2}$.

[40] *CUL Add.* 3960. Section 14: especially problem 5, 57–59 (·≡· Horsley 1: 445–6). In the original no dottage notation for fluxions is used (the fluxions of v, x, y, z being represented by l, m, n, r), and was first introduced by Colson in his (English) publication of the manuscript as *Method of fluxions and infinite series*, London, 1736, and kept by Horsley in his edition of the original Latin version in volume 1 of his *Newtoni opera*, London 1791.

a line-length $z = \dot{y}$ and considering its fluxion $\dot{z}$. Thus, in the following diagram (relettered to accord with the tract), where DC is the radius of curvature at point D of the curve (defined by some relation between AB and BD), take a second point D' on the curve (tangent DT) indefinitely near to D[41] (or $D'C$ is normal to the curve at D'): drawing the rectangle $DHCG$ and any parallel dfg to DFG (meeting the various elements of the figure as shown), we have $Cg:gd =$ subtangent $TB:DB$, $=$ "fluxio basis": "fluxio applicatae", and $DF = DE + EF\left(= \frac{D'E^2}{DE}\right)$ so that, denoting $AB = x$, $BD = y$, $gd = z$ and $Cg = 1$, $\dot{y}:\dot{x} = \dot{z}:1$ (or $z = \dot{y}/\dot{x} = dy/dx$). Now consider the limit-increases in an indefinitely small particle of time o: we have $DE = \dot{x}o$, $D'E = \dot{y}o$, $df = \dot{z}\,o$, or $DF = \left(\dot{x} + \frac{\dot{y}^2}{\dot{x}}\right)o$, and $(Cg =)\ 1 : CG = df : DF = \dot{z}o : \left(\dot{x} + \frac{\dot{y}^2}{\dot{x}}\right)o$, or $CG = \frac{\dot{x}^2 + \dot{y}^2}{\dot{x}\dot{z}}$; further, taking $\dot{x} = 1$ for simplicity $\left(\text{or } z = \dot{y} \text{ and } CG = \frac{1+z^2}{\dot{z}}\right)$, it follows that, as $D' \to D$, $DG\left(= CG \times \frac{dg}{Cg}\right) = CG \times z = \frac{(1+z^2)\,z}{\dot{z}}$, or $DC = \frac{(1+z^2)^{\frac{3}{2}}}{\dot{z}}$, which is an equivalent formula for the radius of curvature at D.

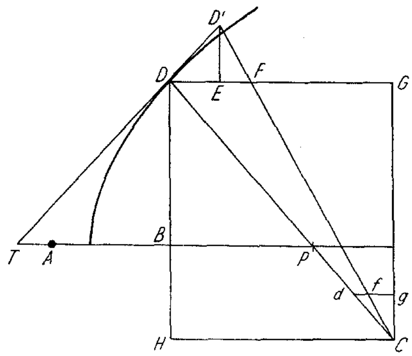


Fig. 119

It is curious that the more general corresponding formulas for a curve whose representing equation is given implicitly as $g(x, y) = 0$ were, in fact, given by NEWTON slightly earlier[42]. In fact NEWTON, taking $g(x, y) = 0$ by $)($, gives

$$DG\ (= CH) = \frac{\cdot)(\,(\cdot)(^2\, y^2 +)(\cdot^2\, x^2)}{\lambda x},$$

$$DH\ (= CG) = \frac{)(\cdot\,(\cdot)(^2\, y^2 +)(\cdot^2\, x^2)}{\lambda y}$$

and finally

$$DC = \frac{(\cdot)(^2\, y^2 +)(\cdot^2\, x^2)^{\frac{3}{2}}}{\lambda x y}, \quad \text{where} \quad -\lambda = \cdot)(^2)(: - 2\cdot)()(\cdot \cdot)(\cdot +)(\cdot^2 :)(.\,^\star$$

⋆ Here, as we have seen in the previous chapter,

$$\cdot)(= x\frac{\partial g}{\partial y}, \quad :)(= x^2\frac{\partial^2 g}{\partial x^2}, \quad)(\cdot = y\frac{\partial g}{\partial y}, \quad)(: = y^2\frac{\partial^2 g}{\partial y^2},$$

and

$$\cdot)(\cdot = x y \frac{\partial^2 g}{\partial x\, \partial y};$$

so that NEWTON gives the correct form

$$-DC = \frac{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}{\left(\frac{\partial g}{\partial x}\right)^2 \frac{\partial^2 g}{\partial y^2} - 2\frac{\partial g}{\partial x}\,\frac{\partial g}{\partial y}\,\frac{\partial^2 g}{\partial x\,\partial y} + \left(\frac{\partial g}{\partial y}\right)^2 \frac{\partial^2 g}{\partial x^2}}.$$

[41] NEWTON says: "sit DD' infinite parvum intervallum".

[42] See *Add.* 4004: 48R–49R, dated 21st May 1665—that is, over a year before the algorithm of October 1666 was drafted.

It is tempting in analogy with the typical modern proof* to assume that NEWTON in some way developed these theorems from his generalization of HUDDE's tangent-method, and so we find. Though it is clear[43] that NEWTON derived his general theorem by numerical induction over a large number of particular cases, yet if we look closely at the structure of his argument we see each time an application of the HUDDE rule.

From this partial-difference formula he deduces immediately a general test for inflexion at a point—clearly, assuming smooth continuity, the radius of curvature must vary continuously and so at an inflexion point be infinite, or ·)(^{2})(: — 2·)()(··)(·+)(·^{2}:)(=0[44], and, similarly, less important theorems follow as further corollaries. However, the limitation remains that, strictly, the various partial derivatives of)(are defined by a method which is valid only for)(algebraic**, and in considering non-algebraic curves a little ingenuity was needed. Typical is NEWTON's treatment of the cycloid,[45] where he finds the radius and centre of curvature from the subtangent property which is, equivalently, $\frac{dy}{dx}=\left(\frac{a-y}{y}\right)^{\frac{1}{2}}$. Specifically, taking coordinates $OX=x$, $XP=y$ at any point P on the cycloid

---

* Specifically, using modern notation, HUDDE's rule for $g(x, y)=0$ gives

$$[0=]\ \frac{dg}{dx}=\frac{\partial g}{\partial x}+\frac{\partial g}{\partial y}\frac{dy}{dx};$$

or, applying it a second time

$$[0=]\ \frac{d^2g}{dx^2}=\frac{d}{dx}\left(\frac{\partial g}{\partial x}\right)+\frac{d}{dx}\left(\frac{\partial g}{\partial y}\right)\frac{dy}{dx}+\frac{\partial g}{\partial y}\frac{d^2y}{dx^2}$$

$$=\frac{\partial^2 g}{\partial x^2}+2\frac{dy}{dx}\frac{\partial^2 g}{\partial x\,\partial y}+\left(\frac{dy}{dx}\right)^2\frac{\partial^2 g}{\partial y^2}+\frac{\partial g}{\partial y}\frac{d^2y}{dx^2}.$$

From this we can express $\frac{dy}{dx}$, $\frac{d^2y}{dx^2}$ in appropriate partial differential form, and substitution in $\frac{-\left(1+\left(\frac{dy}{dx}\right)^2\right)^{\frac{3}{2}}}{\frac{d^2y}{dx^2}}$ yields NEWTON's formula for the radius of curvature DC, and similarly for DG and DH.

** Indeed, NEWTON assumes this explicitly, defining his)(as the polynomial

$$g(x)=\sum_{0\leq i\leq m}\ \sum_{0\leq j\leq n}(a_{ij}\,x^i\,y^j).$$

[43] See especially 48R—48V, where NEWTON works out the radius of curvature for a large number of cases $p+qy^\lambda$, $\lambda=1, 2, 3, 4, \ldots$, where p and q are functions of x (and outlines at 48V the extension which treats $p+qy^\lambda+ry^\mu+\cdots$, $\lambda, \mu=0, 1, 2, 3, 4, \ldots$, with $p, q, r, \ldots$ functions of x). Though NEWTON never formally generalised his argument to include all functions)($\equiv g(x, y)=0$, it is not difficult to fabricate a general proof structurally identical with that used by NEWTON in each particular case.

[44] See *Add.* 3958. 3: 56R: "Prob. 3^d: to find y^e points distinguishing twist y^e concave and convex portions of crooked lines."

[45] *Add.* 3960.14: "exempl. 4^t: 61—63 (≡HORSLEY: *Newtoni opera* **1**: 448—450). The manuscript scrap, *Add.* 3958. 1: 37V (dated 30 October 1665), shows that the basic properties of curvature of a general point on the cycloid were known to NEWTON in 1665, and they are used in a manuscript on motion in a cycloid (1670?) (*Add.* 3958: 89R—91V).

arc $\widehat{OPA}$, we can show[46] that the tangent at P is parallel to the corresponding chord $A'Q$ of the generating semicircle $A'QA$ (where PQR is a general parallel to the base OA)—that is, where $AA'=a$, the slope of the cycloid arc at P is measured by $\frac{A'R}{QR}=\frac{a-y}{(y(a-y))^{\frac{1}{2}}}\left[=\frac{dy}{dx}\right]$ and so, defining $z=\frac{dy}{dx}$, we have

$$\frac{dz}{dy}\left(=\frac{1}{z}\,\frac{dz}{dx}\right)=\frac{d}{dy}\left(\left(\frac{a-y}{y}\right)^{\frac{1}{2}}\right),$$

or the radius of curvature at P is

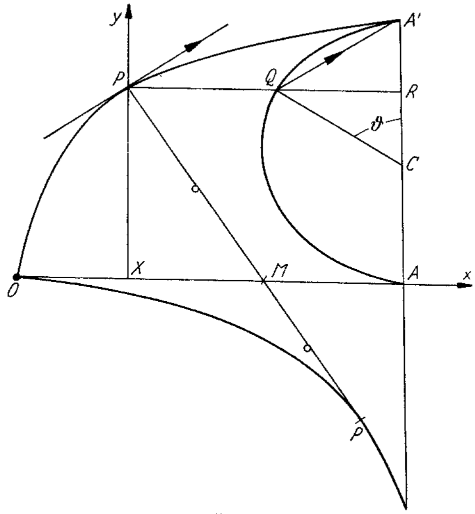


Fig. 120

$$Pp=\frac{(1+z^2)^{\frac{3}{2}}}{\frac{dz}{dx}}=\frac{\left(1+\frac{a-y}{y}\right)^{\frac{3}{2}}}{\left(\frac{a-y}{y}\right)^{\frac{1}{2}}\frac{d}{dy}\left(\left(\frac{a-y}{y}\right)^{\frac{1}{2}}\right)}$$

$$=2(ay)^{\frac{1}{2}},\ =2y\frac{A'Q}{QR}\left(=\frac{MP}{PX}\right)$$

$$=2MP,$$

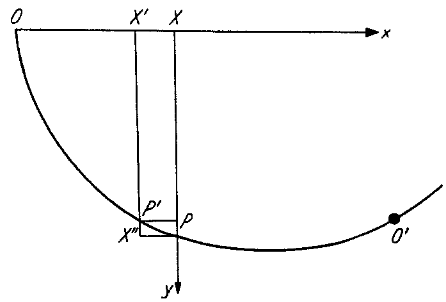


Fig. 121

where MX is the subnormal at P (and, finally, we easily show that p is on a second cycloid Opa', congruent but contraposed to the first, of which it is the evolute; and that Pp is tangent at p).

This example is important when we consider a curious but illuminating dispute which arose at the end of the century over various solutions given to the brachistochrone problem: given two points O, O' in the same vertical plane, to find the path of point P which falls from rest at O to O' in minimum time under gravity (that is, such that the square of its speed at point P is proportional to its vertical distance PX below O). JOHN BERNOULLI, who had proposed the problem[47], gave a neat resolution[48] which pictures the motion of the point under gravity as a point of light moving through a medium in which the speed of light varies as the square root of its distance below the horizontal, PX. Taking the sufficient (but non-necessary) condition for minimum path that each arc, however small, also be a minimum path between its end-points, we can apply SNELLIUS' law for

[46] By any procedure equivalent to differentiating the representing equation

$$\tfrac{1}{2}\pi a-x=(y(a-y))^{\frac{1}{2}}+\tfrac{1}{2}a\cos^{-1}\left(\frac{y-\frac{1}{2}a}{\frac{1}{2}a}\right),\ =\tfrac{1}{2}a(\sin\vartheta+\vartheta),$$

where $\vartheta=\widehat{QCA'}$ (and so $y=\frac{1}{2}a(1+\cos\vartheta)$). Compare previous chapter.

[47] In AE (1696): 269.

[48] In AE (1697): 208—209.

each indefinitely small arc $\widehat{PP'}$, deriving the condition that $\sin PP'X$: point-speed at P be constant for all points P: thus, denoting $OX = x$, $PX = y$, $\sin \widehat{PP'X} = \frac{dx}{ds}$ and speed at $P = (Ky)^{\frac{1}{2}}$ where K is some constant; so that $\frac{dx}{ds} : (Ky)^{\frac{1}{2}} = \lambda$, constant, or, where $\frac{1}{a} = \lambda^2 K$, $\frac{ds}{dx} = \left(\frac{a}{y}\right)^{\frac{1}{2}}$, and so $\frac{dy}{dx} = \left(\frac{a-y}{y}\right)^{\frac{1}{2}}$: which defines the brachistochrone to be a cycloid with origin at O and base along OX. Two years later FATIO DE DUILLIER[49] gave a solution (long and tedious if equally ingenious) which uses virtually second order differentials: for ϱ the radius of curvature at point P of the path, he derives from an equivalent minimal path condition the defining equation $\frac{ds}{dx} = \frac{\varrho}{2y}$ (which, comparing it with the NEWTON cycloid example, again proves the path cycloidal). FATIO's book, *lineae brevissimi descensus* ..., for other reasons aroused a petulant controversy which filled AE during the period 1699 to 1701, and has, indeed, been urged with little justice as the origin of the fluxion priority dispute by those who would whitewash KEILL; and in the angry remarks which were passed LEIBNIZ[50] made the criticism that FATIO's solution is inferior to JOHN BERNOULLI's in that it involves a second order derivative (in introducing the concept of curvature-radius) as against BERNOULLI's first order differential equation. Though his criticism has been supported in recent times[51], the two are exactly equivalent and FATIO's solution is immediately reducible to BERNOULLI's differential equation* —rather, this refusal to admit their equivalence and claiming the one approach superior to the other on such ill-argued grounds reflects the uncertainty and lack of sure insight which accompany immaturity and lack of familiarity with abstract calculus operations.

---

* Thus, substituting

$$-\varrho = \frac{\left(1+\left(\frac{dy}{dx}\right)^2\right)^{\frac{3}{2}}}{\frac{d^2y}{dx^2}} = \frac{(1+z^2)^{\frac{3}{2}}}{\frac{dz}{dx}}$$

where

$$z = \frac{dy}{dx} \quad \left(\text{or } \frac{dz}{dy} = \frac{1}{z} \times \frac{dz}{dx}\right),$$

FATIO's solution is

$$\varrho : 2y = (1+z^2)^{\frac{1}{2}} : 1 \quad \text{or} \quad -\frac{(1+z^2)}{z} = 2y \times \frac{dz}{dy};$$

so that, since $z = \frac{dy}{dx}$ is 0 when $y = a$,

$$-\int_{y=a}^{y=y} \frac{2z}{1+z^2} \cdot dz = \int_a^y \frac{dy}{y}$$

or

$$-\log(1+z^2) = \log(y) - \log(a),$$

and finally $\frac{1}{1+z^2} = \frac{y}{a}$, for some constant a, which is a form of BERNOULLI's solution $\left(z^2 = \frac{a-y}{y}\right)$.

[49] See his *lineae brevissimi descensus investigatio geometrica duplex* ..., London, 1699: 6–8, 11–12.

[50] In AE (1700): 201, where he tries to pursue the analogy of using a conic to solve a problem where a line-construction suffices.

[51] D. J. STRUIK in *Outline of a history of differential geometry*, **1** in Isis **19** (1933): 92–120, especially 98.

In contrast with all this comparatively advanced use of differential techniques, the development of corresponding inverse (integral) procedures—for which, unlike the standard differential algorithms which had been formulated, few general (and no universal) methods are applicable—lagged far behind except in treating the simplest cases. Though for example, NEWTON constructed fairly elaborate tables of integrals in the middle 1660's[52] such tables of standard forms did not appear in printed form till the early 1700's (when, suddenly, they sprout in profusion in every textbook). In particular, the concept of finding a general solution to an analytically given differential equation had hardly crystallized out by the end of the 17th century—so, while NEWTON[53] makes some attempt to deal with simple types of first-order linear differential equations, in true NEWTONIAN style he merely assumes a solution is possible which can be expressed in the form of a convergent sum-sequence, $\frac{y}{x^\lambda} = \sum_{0 \leqq i \leqq n} (a_i x^i)$ for some λ, where each a_i is determined by substituting for y in the given $F\left(x, y, \frac{dy}{dx}\right) = 0$ and then comparing coefficients.

Indeed, before general methods of reducing differential equations to integrable form could be evolved, the subtleties of the concept of transforms allowable under the operation of differentiation had of necessity to be thoroughly understood (and it mattered less whether such transforms were defined analytically or on a suitable geometrical model). However, even by 1700 the concept of transforming a variable (or line-coordinate) was still accepted as a difficult problem, particular forms of which, when proved, could be looked upon with the admiration of achievement. Only in that light can we appreciate LEIBNIZ' overwhelming enthusiasm and pride in his transform $z = y - x\frac{dy}{dx}$, by which he derived the sum-sequence $\frac{\pi}{4} = \lim_{n\to\infty} \sum_{0 \leqq i \leqq n} \left((-1)^i \frac{1}{2i+1}\right)$, and upon which he wrote a whole manuscript treatise *de quadratura arithmetica circuli et hyperbolae, cujus corollarium est trigonometria sine tabulis* (1676)[54]; or NEWTON's equally minute geometrical analysis[55] of the transform $\int \frac{dy}{dx} \cdot dx = y$ used in the construction of his tables of standard integrals in the middle 1660's.

Of course. the systematic generalizations of existing geometrical methods written up from the middle of the century, especially—in England—JAMES GREGORY's *GPU* and ISAAC BARROW's *LG*, contained implicitly many valuable

[52] See *Add.* 4000 and *Add.* 3958. 3: *passim.* These formed the basis of the general tables published (partially) in 1704 in his tract on quadrature of curves, and more fully in 1736 when COLSON first published (an English version of) his 1671 tract on analysis.

[53] Especially in problem 2 of his 1671 tract on analysis: *exposita aequatione fluxiones quantitatum involvente, relationem quantitatum inter se invenire* (· ≡ · HORSLEY: *Newtoni opera* 1: 412—428), which is referred to briefly in NEWTON's letter of 1692 (now lost) which WALLIS added in 1693 to the Latin translation of his *Algebra* (see *opera mathematica* 2: 392—396).

[54] Now in Hanover Royal Library—see J.E. HOFMANN: *Entwicklungsgeschichte der Leibnizschen Mathematik* ...: 32ff., and compare G. LORIA: *Pseudo-versiera e quadratice geometrica* Bibliotheca Mathematica$_3$ 3 (1902—1903): 127—130.

[55] In *CUL Add.* 4000: 134V—135R. His proof is apparently modelled on VAN HEURAET's rectification method, printed in (ed. FRANZ SCHOOTEN) DESCARTES: *geometria* 1 (Amsterdam, 1659): 517—520.

results, but to a large extent the corresponding analytical variable transforms had to be rediscovered in the 1700's and only too infrequently can we show an influence of geometrical upon analytical approaches. So proposition 11 of JAMES GREGORY'S *GPU*[56], which defines the curve OQQ' from the given convex curve OPP' such that PQ, parallel to OX, $=$ subtangent TX, and then shows area $(OX'P'PO) =$ area $(OPP'Q'QO)$, is equivalent to the definite integral transform

$$\int_{x=\beta}^{x=\alpha} \left(y\,\frac{dx}{dy}\right) \cdot dy = \int_{\beta}^{\alpha} y \cdot dx,$$

where $OX = x$, $PX = y$ $\left(\text{and so } TX = PQ = y \times \frac{dx}{dy}\right)$; but the theorem seems never to have been accepted in the 17[th] century as more than a conveniently rigorous proof of a geometrical transform which generalized particular methods of ROBERVAL and TORRICELLI.

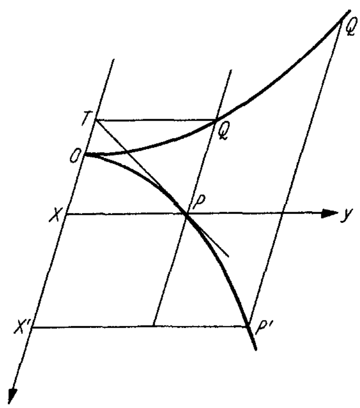


Fig. 122

There do exist, however, some infrequent examples which prove the rule, and especially that of JOHN CRAIG, who based[57] a general method of integration on transforms derived from BARROW'S *LG*. Especially theorem 1 of his *methodus figurarum* ... is based on an equivalent result in *LG*: lectio 11,[58] which shows

$$\int \left(y\frac{dy}{dx}\right) \cdot dx \left[= \int y \cdot dy\right] = \frac{1}{2}\,y^2:$$

geometrically, if curve OP is defined by some relation between $OX = x$ and $PX = y$ and from it curve OQ is defined by ordinate $XQ =$ subnormal XN, then area $(OXQ) = \frac{1}{2} \cdot PX^2$. But even there the application is a little artificial and the basic geometrical model is easily—and preferably—eliminated.*

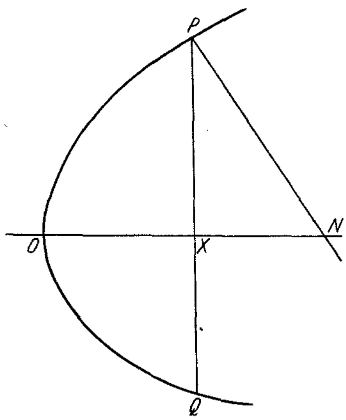


Fig. 123

* Significantly it is completely eliminated in his fluxional calculus text-book of 1718, *de calculo fluentium libri duo*.

[56] See chapter 9.

[57] In his *methodus figurarum* ..., London 1685: theorem 1: 2—3, and especially in his *tractatus mathematicus* ..., London 1693: pars prima, *passim*. As he says in the former (p. 3): "I owe this theorem to Dr. Barrow who has innumerable and sublime theorems on the properties of curve-lines, nor has it been my fortune ... to have seen anyone ... who with so much judgment and success has treated and promoted this rather abstruse and less cultivated part of geometry." It is interesting to compare JOHN BERNOULLI'S criticism of the method in a letter to L'HOSPITAL of May and June 1696—see *Briefwechsel von Johann Bernoulli* (ed. O. SPIESS) **1** (Basel, 1955): 286ff., 293ff.

[58] *LG*: lectio **11**: § 1: 85. Further, *tractatus mathematicus* ..., lemma 2: 20ff. ≡ LG: lectio **11**: § 19: 90, while the more original theorem below (pp. 36—37)—, which shows that $dx/dz = u/y$ implies $y \cdot \int dx = \int u \cdot dz$—is but a slight generalization of ideas worked out in BARROW *LG*: lectio **11**: *passim*.

In a very strong sense the crystallization out of standard algorithmic calculus techniques was inevitable, and the blunt answer to that favourite 19th century query of how such important advances could be made on such inadequate bases is that it begs the question: the bases were not inadequate, and problems of rigour, consistency and existence were all answered, if suitable analytical justification was not forthcoming, by direct appeal to the visual plausibility of a geometrical model.

In fact—and in summary—what was done in 17th century mathematics (and, even more so, what was sketched in or hinted at) was sufficient to provide rich pickings for 18th century mathematicians seeking a lead into the unknown. In the case of EULER, particularly, it is enlightening to see how much of his work improves and generalises the obscurer but richer parts of the published work of DESCARTES, FERMAT, WALLIS and NEWTON—and I do not mean thereby to decrease EULER'S status as a creative mathematician of the first order. Perhaps we tend to underestimate the 17th century mathematical achievement, over-impressed by the greater self-confidence and technical mastery of the 18th and 19th centuries or disillusioned by the heavy numerical bias of the immediately preceding 16th century. In fact, the foundations for two centuries of mathematical advance were laid in the 17th century, and only recently have we, in our newfound preference for the exhaustive axiomatic treatment, passed to a higher plane of mathematical thought. But though the profoundest achievements of the 17th century be now no more than schoolroom mathematics, the headspinning excitement of first discovery which fills the pages of its great works will never quite be lost, and the genius and brilliance of its individual mathematicians will always stand out.

Select Bibliography of primary sources

Note: For conciseness of reference many of the primary texts quoted in the body of this essay have been cited—though not always—by a code-reference system which adapts that used for many years by J.E. HOFMANN in his various books and articles. Its use should be clear. Thus, where note 15 of Chapter 5 cites (JAMES GREGORY) *EG*: part 2: 9—13 the reference is to JAMES GREGORY: *exercitationes geometricae*. London, 1668: part 2: pages 9—13; and note 28 of Chapter 3 (PT **3** (1668); 645—649) refers to Philosophical Transactions, Volume **3** (year 1668): pages 645—649.

In tabling these code-references it is convenient also to collect the main primary sources, both printed and manuscript, which have been consulted. Secondary texts, commentaries and standard histories, insofar as they enlighten or reinforce the argument, are cited in the notes to individual chapters, and there seems little point in repeating them here—indeed, there exist several excellent and up-to-date bibliographies which it is unnecessary to duplicate. These include in particular

RUSSO, F.: *Histoire des sciences et des techniques. Bibliographie.* Actualités sc. et ind. **1204**. Paris, 1954 (with supplement 1955),

but above all the critical bibliographies to be found in Isis and (since 1940) Mathematical Reviews (History section) together with the copious references and citations of

HOFMANN, J.E.: *Geschichte der Mathematik.* Berlin 1953 —.

1. (1953). *Von den Anfängen bis zum Auftreten von Fermat und Descartes.*

2. (1957). *Von Fermat und Descartes bis zur Erfindung des Calculus und bis zum Ausbau der neuen Methoden.*

3. (1957). *Von der Auseinandersetzung um den Calculus bis zur Französischen Revolution.*

A. Periodicals and collections

PT Philosophical Transactions: giving some Account of the present Undertakings, Studies and Labours of the Ingenious in many Considerable Parts of the World. London, 1666–.

AE acta eruditorum, Leipzig, 1682–1779.

GERHARDT (*B*) C. J. GERHARDT: *Der Briefwechsel von Gottfried Wilhelm Leibniz mit Mathematikern*, Band **1**, Berlin, 1899.

RIGAUD (*C*) S. J. RIGAUD: *Correspondence of scientific men of the XVII*th *century*, Oxford, 1841 (2 vols.).

Commercium epistolicum D. Johannis Collins et aliorum de analysi promota. London, 1712/3 (augmented in $_2$1722).

B. Individual works

1. English

BARROW, ISAAC: *Euclidis Elementorum libri xv breviter demonstrata.* Cantabrigiae, 1655.

LM *lectiones mathematicae xxiii in quibus principia matheseos generalia exponuntur ... habitae Cantabrigiae 1664—1666*, London 1685.

LG *lectiones xviii Cantabrigiae in scholis publicis habitae; in quibus opticorum phaenomenwn genuinae rationes investigantur ac exponuntur. annexae sunt lectiones aliquot geometricae*, London, 1670.

Archimedis opera: Apollonii Pergaei conicorum libri iiii: Theodosii sphaerica: methodo nova illustrata et succincte demonstrata. London, 1675.

BRIGGS, HENRY:

AL *arithmetica logarithmica, sive logarithmorum chiliades triginta. ... quorum ope multa perficiuntur arithmetica problemata et geometrica*, London, 1624.

TB *trigonometria britannica, sive de doctrina triangulorum libri duo*, Gouda, 1633.

Viscount BROUNCKER, WILLIAM: None of BROUNCKER's writings exist in separate form. I have published a full list of the scattered fragments of his mathematical work in *Notes and Records of the Royal Society*, Tercentenery issue, July 1960 (especially p. 157).

CRAIG, JOHN:

methodus figurarum lineis rectis et curvis comprehensarum quadraturas determinandi, London, 1685.

tractatus mathematicus de figurarum curvilineorum quadraturis et locis geometricis, London, 1693.

theologiae christianae principia mathematica, London, 1699.

de calculo fluentium libri duo, London, 1718.

DUILLIER, FATIO DE:

lineae brevissimi descensus investigatio geometrica duplex. cui addita est investigatio geometrica solidi rotundi in quod minima fiat resistentia, London, 1699.

GREGORY, DAVID:

exercitatio geometrica de dimensione figurarum, sive specimen methodi generalis dimetiendi quasvis figuras, Edinburgh, 1684.

GREGORY, JAMES:

VCHQ *vera circuli et hyperbolae quadratura, in propria sua proportionis specie inventa et demonstrata*, Padua, $_1$1667, $_2$1668.

GPU *geometriae pars universalis, inserviens quantitatum curvarum transmutationi et mensurae*, Padua, 1668.

EG *exercitationes geometricae. appendicula ad veram circuli et hyperbolae quadraturam. N. Mercatoris quadratura hyperbolae geometrice demonstrata. analogia inter lineam meridianam planisphaerii nautici et tangentes artificiales geometrica demonstrata, seu quod secantium naturalium additio efficiat tangentes artificiales. item, quod tangentium naturalium additio efficiat secantes artificiales. quadratura conchoidis. quadratura cissoidis. methodus facilis et accurata componendi secantes et tangentes artificiales*, London, 1668.

GREGORY *TV* *James Gregory Tercentenary Memorial Volume, containing his correspondence with John Collins and his hitherto unpublished mathematical manuscripts ...* (ed. H.W. TURNBULL), London, 1939.

HALLEY, EDMUND: *Apollonii Pergaei de sectione rationis libri duo ... accedunt ejusdem de sectione spatii libri duo restituti, opus analyseos geometricae studiosis apprime utile* ..., Oxford, 1706.
Apollonii Pergaei conicorum libri octo, Oxford, 1710.
HARRIS, JOHN: *A New short Treatise of Algebra, with the Geometrical Construction of Equations ..., together with a Specimen of the Nature and Algorithm of Fluxions*, London, 1703.
HAYES, CHARLES: *A Treatise of Fluxions, or an Introduction to Mathematical Philosophy containing a full explication of that Method by which the most celebrated Geometers of the present age have made such vast advances in Mechanical Philosophy. A Work very useful for those that would know how to apply Mathematicks to Nature*, London, 1704.
KERSEY, JOHN: *The Elements of that Mathematical Art commonly called Algebra expounded in four Books*, London, 1673.
MERCATOR, NICOLAUS:
Log *logarithmotechnia, sive methodus construendi logarithmos nova, accurata et facilis*, London, 1668.
Euclidis elementa geometrica, novo ordine ac methodo fere demonstrata, una cum Nicolae Mercatoris in Geometriam introductione brevi, qua magnitudinum ortus ex genuinis principiis et ortarum affectiones ex ipsa genesi derivantur, London, 1678.
NAPIER, JOHN: *mirifici logarithmorum canonis descriptio ejusque usus* ..., Edinburgh 1614.
mirifici logarithmorum canonis constructio et eorum ad naturales ipsorum numeros habitudines, Edinburgh, 1619.
NAPIER *TV Napier Tercentenary Memorial Volume* (ed. C. G. KNOTT), London, 1915.
NEWTON, ISAAC: (A mass of unpublished mathematical manuscript exists in the Cambridge University Library, especially *CUL Add.* 3958—3964, 4000, 4004.)
To resolve Problems by Motion. (October 1666.) (*CUL Add.* 3958.3: 48v—63v.)
analysis per aequationes numero terminorum infinitas, London, 1711 (but written 1668/9).
methodus fluxionum et serierum infinitarum (1671). (*Add.* 3960.14: printed by S. HORSLEY as *geometria analytica, sive artis analyticae specimina* in his *Newtoni opera quae exstant omnia*, 1 London, 1779; and in English by J. COLSON, London, 1736.)
AU arithmetica universalis, sive de compositione et resolution arithmetica liber, London, 1707 (printed from his Lucasian lectures of 1673—1683 ·=· *CUL Dd.* 9.68).
MD methodus differentialis, London, 1704 (written c. 1675).
tractatus de compositione locorum solidorum (c. 1675). (*CUL Add.* 3963.8/12/13).
PM philosophiae naturalis principia mathematica, London, $_1$1687, $_2$1713, $_3$1726 (based largely on his Lucasian lectures of 1684—1687 · = · *CUL Dd. Dd.* 9.46/.4.18).
tractatus de quadratura curvarum, London, 1704 (but written in 1691/2 as part of a projected treatise on geometry).
enumeratio linearum tertii ordinis, London 1704 (but written c. 1695—fuller manuscript versions are at *CUL Add.* 3961 and, of the projective classification of cubics, at *Add.* 4004: 153—159.)
regula differentiarum (c. 1692). (*CUL Add.* 3964.5, printed by D.C. FRASER, London, 1927.)
RAPHSON, JOHN:
analysis aequationum univeralis, seu ad aequationes algebraicas resolvendas methodus generalis et expedita, ex nova infinitarum serierum methodo deducta et demonstrata, London, $_1$1690, $_2$1697, $_3$1704.
SR de spatio reali, seu Ente infinito conamen mathematico-metaphysicum (added to 2nd and 3rd editions of his *analysis aequationum universalis*).
WALLIS, JOHN:
operum mathematicorum pars prima/altera, Oxford, 1657/6. Part 1 (1657) has *MU mathesis universalis, seu opus arithmeticum*. Part 2 (1656) includes: *philogice et mathematice traditum, arithmeticam numerosam et speciosam aliague continens.*

AI *arithmetica infinitorum, sive nova methodus inquirendi in curvilineorum quadraturam aliaque difficiliora matheseos problemata,* and

SC *de sectionibus conicis nova methodus expositis tractatus.*

CE *commercium epistolicum de quaestionibus quibusdam mathematicis habitum,* Oxford, 1658.

tractatus duo, prior de cycloide et corporibus inde genitis; posterior epistolaris in qua agitur de cissoide et corporibus inde genitis: et de curvarum tum linearum εὐθύνσει, tum superficierum πλατυσμῷ. Oxford, 1659.

mechanica, sive de motu tractatus geometricus. London, 1670/1.

A Treatise of Algebra both historical and practical, showing the original, progress and advancement thereof from time to time, and by what steps it hath attained to the heighth at which it now is ..., London, 1685.

Op *opera omnia mathematica,* Oxford, **1** (1695); **2** (1693); **3** (1699).

WREN, CHRISTOPHER: (As with BROUNCKER none of WREN's writings exist in separate printed form, and the issue of *Notes and Records* of the Royal Society there mentioned (on p. 111) contains a list of the mathematical fragments as I know them.)

2. French

ARNAUD, ANTOINE:

Nouveaux élémens de géométrie; contenant, outre un ordre tout nouveau et de nouvelles démonstrations des propositions les plus communes, de nouveaux moyens de faire voir quelles lignes sont incommensurables, de nouvelles mesures des angles dont on ne s'estoit point encore avisé, et de nouvelles manières de trouver et de demontrer la proportion des lignes. Paris, 1667.

DESARGES, GIRARD:

Brouillon Proiect d'une Atteinte aux Evenemens des Rencontres du Cône avec un Plan, Paris, 1639. (The only copy now known is in the Bibliothèque Nationale at Rés. V. 1209—*cf.* R. TATON: *L'Oeuvre mathématique de G. Desargues,* Paris, 1951.)

DESCARTES, RENÉ:

La Géométrie (3^{rd}) appendix, pp. 297—413, of his *Discours de la Methode pour bien conduire sa Raison et chercher la Vérité dans les Sciences,* Leyden, 1637. *geometria ... anno 1637 Gallice edita, una cum notis Florimondi de Beaune ...,* Leyden, 1649. (A further augmented edition appeared at Amsterdam, 1659/61, $_2$1683.)

OE *Oeuvres de Descartes* (ed. CH. ADAM & P. TANNERY) (12 vols) Paris, 1897—1910.

FERMAT, PIERRE:

OE *Oeuvres de Fermat* (ed. P. TANNERY & CH. HENRY) (4 vols.) Paris, 1891—1912.

Op *varia opera mathematica,* Toulouse, 1679.

LAHIRE, PHILIPPE DE:

Observations ... sur les Points d'Attouchement de trois lignes droites qui touchent la Section d'un Cône sur quelques-uns des Diamètres, et sur le Centre de la mesme Section, mises en lumière par A. Bosse. Paris 1672. (See R. TATON: *La première œuvre géométrique de Philippe de La Hire.* Révue d'Histoire des Sciences *6* (1953): 73—111.

Nouvelle Méthode en Géométrie pour les Sections des Superficies Cylindriques qui ont pour bases des Cercles ou des Paraboles, des Ellipses et des Hyperboles. Plus les Planiconiques, Paris, 1673.

Nouveaux Elémens des Sections Coniques, les Lieux géométriques, la Construction ou Affectation des Equations, Paris, 1679.

sectiones conicae in novem libros distributae, in quibus quidquid hactenus observatione dignum cum a veteribus, tum a recentioribus geometrie traditum est, novis contractisque demonstrationibus explicatur ..., Paris, 1685.

LALOUVÈRE, ANTOINE DE (ANTONIUS LALOVERA):

Quadratura circuli et hyperbclae segmentorum, ex dato eorum centro gravitatis, una cum inventione proportionis et centri gravitatis in portionibus sphaerae plurimorumque periphericorum, nec non tetragonismo absoluto certa cujusdam cylindri partis et aliorum ..., Toulouse, 1651.

veterum geometria promotu in septem de cycloide libris ..., Toulouse, 1660.

PASCAL, BLAISE:

Essay pour les Coniques, Paris, 1640 (privately printed and circulated).

Lettres de A. Dettonville, contenant quelques unes de ses Inventions en Géométrie ... L'Egalité entre les lignes Spirale et Parabolique demonstrée a la manière des Anciens ..., Paris, 1659.

Traité du Triangle arithmétique, avec quelques autres petits traités sur la mesme matière ..., Paris, 1665.

OE Oeuvres des Blaise Pascal (14 Vols.) (ed. L. BRUNSCHVICG, P. BOUTROUX & F. GRAZIER), Paris, 1908—1925.

ROBERVAL, GILLES PERSONE DE: (All ROBERVAL's mathematical work as it then existed was collected after his death and printed in vol. **6** of the Mémoires de l'Académie royale des sciences (1666—1699), Paris, 1693, = (2nd ed.) 1730: pp. 1—478. That still remains our only source, apart from minor correspondence with contemporaries.)

VIÈTE, FRANCOIS (FRANCISCUS VIETA):

Op opera mathematica (ed. F.v. SCHOOTEN), Leyden, 1646.

3. Italian

CAVALIERI, BONAVENTURA:

GI geometria indivisibilibus continuorum nova quadam ratione promota, Bologna, $_1$1635, $_2$1653. *exercitationes geometricae sex*, Bologna, 1647.

MENGOLI, PIETRO:

novae quadraturae arithmeticae, seu de additione fractarum, Bologna, 1651.

via regia ad mathematicas per arithmeticam, algebram speciosam, planimetriam, Bologna, 1655.

GS geometria speciosa, Bologna, 1659.

circolo, Bologna, 1672.

TORRICELLI, EVANGELISTA:

opera geometrica, Florence, 1644.

opere (ed. G. LORIA & G. VASSURA), (4 vols.) Faenza, 1919/1944.

de infinitis spiralibus (manuscript printed by E. CARRUCCIO, Pisa, 1955.)

4. German/Dutch

ST. VINCENT, GREGORY (GREGORIUS a SANCTO VINCENTIO):

OG opus geometricum quadraturae circuli et sectionum coni, Antwerp, 1647.

HUYGENS, CHRISTIAAN:

theoremata de quadratura hyperboles, ellipsis et circuli ex dato portionum gravitatis centro, Leyden, 1651.

de circuli magnitudine inventa, Leyden, 1654.

OE Oeuvres complètes (22 vols.), The Hague, 1888—1950.

KEPLER, JOHANNES:

nova stereometria doliorum vinariorum, Linz, 1615.

LEIBNIZ, GOTTFRIED WILHELM:

Many of LEIBNIZ's mathematical manuscripts, now in the Royal Library at Hanover, remain to be published. A few were printed by C.J. GERHARDT at the end of the 19th century—of which J.M. CHILD's *The early mathematical manuscripts of Leibniz*, Chicago, 1916 is a convenient collection (in English translation)—but for the rest we have to rely on the description given by J.E. HOFMANN in his *Die Entwicklungsgeschichte der Leibnizschen Mathematik während des Aufenthaltes in Paris (1672/76)*, München, 1949. (All LEIBNIZ's important mathematical ideas were given to his contemporaries—if at all—in periodical articles which are far too numerous to list here.)

SARASA, ALPHONS ANTOINE DE:

solutio problematis a R.P. Marino mersenno minimo propositi: datis tribus quibuscunque magnitudinibus rationalibus vel irrationalibus, datisque duarum ex illis logarithmis, tertiae logarithmum geometrice invenire, Antwerp, 1649.

St. Catharine's College
Cambridge, England

(Received January 20, 1961)