Learning calculus: Derivative as a difference quotient, a limit and as a function, some historical origins and pedagogical considerations

Leo Rogers and Sue Pope

British Society for the History of Mathematics (BSHM); Manchester Metropolitan University

Calculus is at the heart of advanced school mathematics. This branch of mathematics has a rich and fascinating history that can enrich learning and humanise what is too often perceived as a set of routines for completing exam questions with little appreciation of how these techniques have developed over time and the sophisticated mathematical concepts underpinning them. We draw on notions of threshold concepts and troublesome knowledge to consider the pedagogical implications.

Keywords: history of mathematics, calculus, derivative, threshold concepts

Introduction

Calculus is what distinguishes advanced school mathematics from the National Curriculum. All the work on algebra, functions and graphs pre-16 are the building blocks for the development of differentiation and integration. Application of calculus is an essential tool to solve problems in further study of mathematics and subjects that use mathematics: engineering, physical sciences, economics etc. Compared to much school mathematics, calculus is relatively modern. Yet its roots are in the work of Archimedes (287-212 BCE).

In our paper on the history of quadratics (Rogers & Pope, 2015) we set out the rationale for using the history of mathematics in education. In this short article, we discuss some historical origins and pedagogical considerations for learning the derivative.

Threshold concepts

The teaching of mathematics often focusses on how to perform algorithms and the procedures associated with them. The rules governing the procedures have been established over time (in the history of the subject) and are made to become ‘habits of the mind’ rather than situations to be questioned and explored.

A threshold concept can be considered as a doorway or a portal that opens a new and previously inaccessible way of thinking about something. It represents a transformed way of understanding, interpreting, or viewing something without which the learner cannot progress. A consequence of comprehending a threshold concept may be a transformed internal view of subject matter, its landscape, or even world view. This transformation may be sudden or it may be protracted over time, with the transition to understanding proving troublesome (Meyer, Land & Baillie, 2010). ‘Troublesome Knowledge’ – or counter-intuitive knowledge – frequently appears in mathematics (non-commutativity, ‘imaginary’ numbers, space-time, etc.) yet we are asked to accept them and ‘carry on’ with the process of applying them.

Introducing the derivative as quotient

A common introduction to differentiation includes the use of a simple graphical image of the relation between x and *f(x)* (e.g. $y = x^2 + c$ often appears in introductory texts), where a chord is drawn from one point to another on the curve. We then name the gradient of the chord (sometimes called the secant line) in terms of finite differences:

$$\frac{\Delta y}{\Delta x} \rightarrow \frac{f(x_2) - f(x_1)}{x_2 - x_1} \rightarrow \frac{f(x+h) - f(x)}{h}$$

where the *increment* h is regarded as something *infinitely small.* The discourse that ensues entails a process of arguing that as the increment h gets smaller, the numerator of the fraction changes from a finite difference to something very small - almost l*ike a point*. This is the stage when it usually becomes necessary to introduce the concept of a *limit* or a *limiting value* where *h tends to zero* using the notation: $h \rightarrow 0$ and

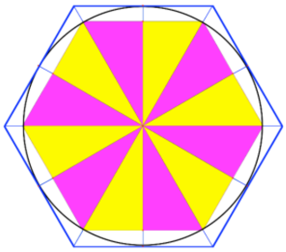
$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

If $h \rightarrow 0$ students often cannot see why the expression above does not simply disappear completely.

There are many different conceptions packed into this apparently simple statement and the problem of understanding what is meant by the *infinitely small* is a long story. Until the 19th century most of the conceptions were based on geometrical images, as we find in many contemporary calculus textbooks.

Early methods of approximating areas

The notion of an 'ever-decreasing quantity' dates back before Archimedes (287-212 BCE) who used Eudoxus' *method of exhaustion* to find bounds for the ratio of the circumference to the diameter by placing the circle between two measurable polygons:



inside polygon < circle < outside polygon

$$3\frac{1}{7} \langle \frac{diameter}{circumference} \langle 3\frac{10}{71}$$

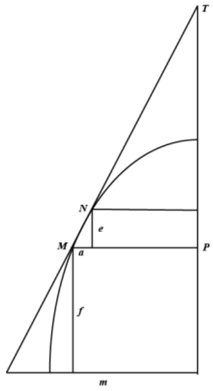
(approximately 3.1408) (approximately 3.1429)

Figure 1 Estimating the ratio of the circumference to the diameter.

Archimedes also estimated the area of a parabolic segment by adding increasingly small areas in a geometric sequence; in neither case did he declare an end to these processes.

In the seventeenth century, many mathematicians were using infinitesimals while looking for an algorithm that would link the problem of finding tangents to curves with that of finding areas of various geometrical objects. In 1665, Newton, and in 1673 Leibniz independently devised the basic algorithms of the calculus and went on to give many demonstrations of their applications to different curves representing algebraic and trigonometric functions.

Typical of the attempts that were almost successful is the work of Isaac Barrow, Newton's tutor at Cambridge. In his triangle, Barrow introduced two infinitesimals *a* for the '*x*-coordinate' and *e* for the abscissa.



The coordinates of two points M (f, m) and N ($f + e$, $m - a$) on the curve MN are connected by the [arbitrary] relation

$$f^3 + m^3 = (f + e)^3 + (m - a)^3$$

Discarding terms containing powers of a and e higher than the first we have $0 = 3ef^2 - 3am^2$ so $\frac{e}{a} = \frac{m^2}{f^2}$

Thus if t denotes the sub-tangent (TP in the figure) then $\frac{TP}{TM}$ or $\frac{t}{m} = \frac{e}{a} = \frac{m^2}{f^2}$ (Baron, 1969 p.250) (Barrow, 1669 Lectiones Geometricae)

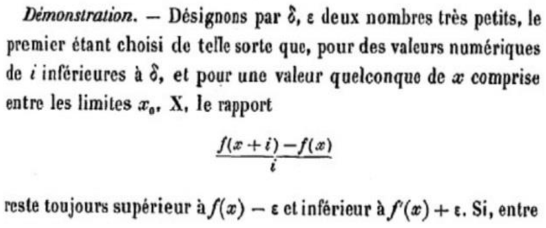
Figure 2 Barrow's infinitesimal triangle

Introducing the derivative as a limit

In the 19th Century, Cauchy defined certain terms in his Cours D'Analyse 'Preliminaries' (1821):

> When the values successively attributed to a particular variable indefinitely approach a fixed value in such a way as to end up by differing from it by as little as we wish, this fixed value is called the ***limit*** of all the other values. (p.44)
>
> When the successive numerical values of such a variable decrease indefinitely, in such a way as to fall below any given number, this variable becomes what we call ***infinitesimal***, or an ***infinitely small quantity***.
>
> When variable quantities are related to each other such that the values of some of them being given one can find all of the others, we consider these various quantities to be expressed by means of several among them, which therefore take the name independent variables. The other quantities expressed by means of the independent variables are called **functions** of those same variables.
>
> The function *f (x)* is a **continuous function** of *x* between the assigned limits if, for each value of *x* between these limits, the numerical value of the difference *f (x+a)− f (x)* decreases indefinitely with the numerical value of a . In other words, the function *f (x)* is continuous with respect to *x* between the given limits if, between these limits, an infinitely small increment in the variable always produces an infinitely small increment in the function itself.

This latter statement means there are infinitely many values of a variable i.e. infinitely many points on the curve, between the value *f(x + a)* and *f(x)* - a fact that many students do not appreciate. Cauchy's first Epsilon -Delta proof is in the Collection (Resumé) of lessons given in the *Ecole Polytechnique* (1826).


Démonstration. — Désignons par δ, ε deux nombres très petits, le premier étant choisi de telle sorte que, pour des valeurs numériques de i inférieures à δ, et pour une valeur quelconque de x comprise entre les limites x_0, X, le rapport

$$\frac{f(x+i)-f(x)}{i}$$

reste toujours supérieur à $f(x) - \varepsilon$ et inférieur à $f'(x) + \varepsilon$. Si, entre

Let δ and ε be two very small numbers: the first is chosen so that for all numerical [absolute] values of i less than δ , and for any value of x included between the limits x and X, the ratio $\frac{f(x+i)-f(x)}{i}$ will always be greater than $f(x) - \varepsilon$ and less than $f(x) + \varepsilon$

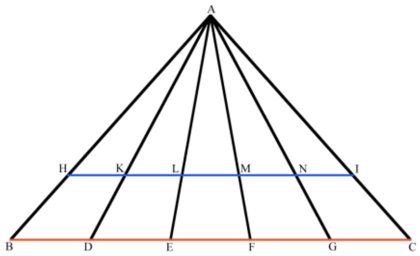
Fourier (1822) showed clearly that most of the misunderstandings that arose in the debate about the vibrating string (a practical experiment) - and other analytic representations of physical contexts - were the result of confusing two seemingly identical but actually vastly different concepts, namely that of *function* and its *analytic representation*.

> A general function *f(x)* is a sequence of values or ordinates, each of which is arbitrary...It is by no means assumed that these ordinates are subject to any general law; they may follow one another in a completely arbitrary manner, and each of them is defined as if it were a unique quantity. (Fourier, 1822: 552)

Prior to Fourier, no distinction was drawn between the concepts of *function* (which at this time was a geometrical object) and *analytic representation*, (that is, representing an expression entirely in algebraic terms – based on arithmetical conceptions).

Over the 19th century mathematicians began to introduce a number of important modifications which refined and clarified these ideas through two major avenues: the *construction of real numbers* and the refinement of *the concept of continuity* in analysis. However, it became clear that the two areas were inter-related where clearer concepts of real numbers were fundamental for advances in analysis.

Among the artists of the early Renaissance, a new problem had emerged. Euclid (300 BCE in I, IV) established that similarity is a transitive relation. Piero della Francesca (1415?-1492) showed that if a pair of unequal parallel segments are divided into equal parts, lines joining corresponding points converge to a vanishing point.



Della Francesca's argument was based on the fact that each of the pairs of triangles ABD and AHK, ADE and AKL etc., are similar, because HK is parallel to BC and the ratio AB to BC is the same as AH to HI. So, all the converging lines meet at A, the vanishing point.

Figure 3 Piero della Francesca's Theorem

This diagram also demonstrates the equivalence of linear segments for a rational division of the real line, but this idea does not appear as a problem concerning real numbers before Bolzano (1781-1848) who anticipated Cantor's theory of infinite sets.

> The most common kind of proof depends on a truth borrowed from geometry, …… There is certainly nothing to be said against the correctness nor against the obviousness of this geometrical proposition. But it is also equally clear that it is an unacceptable breach of good method to try to derive truth of pure (or general) mathematics (i.e. arithmetic, algebra, analysis) from considerations which belong to a merely applied (or special) part of it, namely, geometry. (Bolzano in Russ, 2004: 254 -257)

Cauchy did not distinguish clearly between continuity and uniform continuity on an interval. In his 1821 *Cours d'analyse*, Cauchy argued that the (pointwise) limit of (pointwise) continuous functions was itself (pointwise) continuous, a statement interpreted as being incorrect by many scholars. The correct statement is that the *uniform limit of continuous functions is continuous* (also, the uniform limit of uniformly continuous functions is uniformly continuous). This required the concept of *uniform convergence*. Weirstrass saw the importance of *uniform convergence*, formalised it and applied it widely throughout the foundations of calculus

> $f(x)$ is continuous at $x = x_0$ if $\forall \varepsilon > 0\ \exists\ \delta > 0$ such that for every x in the domain of f, $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$. (Grattan-Guiness, 1980)

While teaching calculus, Dedekind (1831-1916) developed the idea of a cut - now a standard definition of the real numbers. Any irrational number divides the rational numbers into two classes (sets), with all the numbers of one class (greater) being strictly greater than all the numbers of the other (lesser) class. For example, $\sqrt{2}$ defines all the negative numbers and the numbers whose squares are less than 2 into the lesser class, and the positive numbers the squares of which are greater than 2 into the greater class. Every location on the number line continuum contains either a rational or an irrational number. There are no empty spaces, or discontinuities. Dedekind's theorem states that if there existed a one-one correspondence between two sets, then the two sets are similar. He invoked similarity to give the first precise definition of an infinite set: a set is infinite when it is similar to a proper part of itself, (e.g. integers $\aleph$ and their squares $\aleph^2$). Not only does he accept this system as an 'actual infinity', i.e. a complete infinite set as a mathematical object in itself; he also considers it structurally, as an example of a linearly ordered set closed under addition and multiplication (i.e. an ordered field). Consequently, Dedekind's view of a function was a one-one mapping of one set onto another.

Once accepted, these modifications produced a situation where there was 'no turning back' and once the new ideas were accepted, it was impossible to 'un-think' them so these threshold concepts transformed our way of understanding, interpreting, or viewing analysis, without which the contemporary learner cannot progress.

Felix Klein (1849-1925) in Vol III of Elementary Mathematics from a Higher Standpoint (2016) draws the distinction between Precision Mathematics (Pure) and Approximation Mathematics (Applied) and writes about the Abstract and Empirical Definition of a Function. The most general definition of a function that we have reached in modern mathematics starts by fixing the values that the independent variable *x* can take on. We define that *x* must successively pass through the points of a certain 'point set'. The language used is therefore geometric, but thanks to the Cantor-Dedekind Axiom the arithmetical meaning is defined as well. Within this point set *y* is called a function of x, in symbols $y = f(x)$, if for every x of the set there is a specific y (x and y are intended as precisely defined numbers, that is, as decimal fractions with well-defined digits). Usually, *x* passes continuously through the points of a part of the axis, which correspond to the set of all points between two fixed ones *m* and *n*. Such a point set is also called an interval *mn*; we speak of a closed interval if the endpoints belong to the interval, of an open interval if the endpoints do not belong to the interval. We thus obtain the older definition of a function, like the one used for instance by Lejeune-Dirichlet: y is called a function of x in an interval, if for every numerical value x in the interval, there is a well-defined numerical value of y. (Klein 2016 Vol. 3: 15)

The derivative as a function

Bourbaki (1939) introduced differentiation as a functional relation between a set of elements (the x-coordinates) and the results of the limiting process $\frac{dy}{dx}$ and today, in axiomatic set theory, the primitive notions are set, class, and membership so we *define* 'function' as a relation, *on* a set of ordered pairs, and *define* an ordered pair as a set of two 'dissymetric' sets.

Conclusion

We are used to manipulating geometrical objects and concrete models of the number system, and we are quite remarkable in being able to visualise their transformations in various ways (ATM**,** 2016)**.** However, visualisation takes us only so far, and we have taken physical tools and turned them into mental tools. Dealing with *threshold* *concepts* can draw us away from the 'natural world' into a new and exciting world of the imagination where new things (objects, actions, relations) become possible, but we need to be careful to distinguish which world we inhabit.

Klein's distinction between Precision and Approximation mathematics is important to keep in mind. Moving from the 'real world' into the 'theoretical world' of analysis is not psychologically easy. Robinson (1966) in his Nonstandard Analysis showed that a system of infinitesimals could be logically consistent and Keisler (1976) demonstrated a viable approach using set theory. Tall (1980), and other educators, discussing this version of Nonstandard Analysis, maintain that the use of infinitesimals is more intuitive and more easily grasped by students than the so-called 'epsilon-delta' approach to analytic concepts. This approach can sometimes provide easier and more intuitive proofs of results than the corresponding epsilon-delta formulation, but few university teachers have taken up this challenge.

This reminds us of a remark by John von Neumann, "In mathematics you don't understand things. You just get used to them." (Zukav, 1979: 208 footnote) - How true has that been for you?

References

ATM (2016). Imagery *Mathematics Teaching* 253.

Baron, M. (1969). *The Origins of the Infinitesimal Calculus*. Oxford & London: Pergamon

Barrow, I. (1669). *Lectiones Geometricae*. Cambridge

Bourbaki, N. (1939). *Elements of Mathematics: The Theory of Sets.* Paris: Hermann.

Cauchy, A.L. (1821). *Cours D'Analyse de l'ecole Royale Polytechnique*. Paris: De Bure.

Fourier, J. (1822). *La Theorie Analytique de la Chaleur.* Paris.

Grattan-Guinness, I. (1980). *From the Calculus to Set Theory1630-1910*. London: Duckworth

Keisler, H. J. (1976). *Foundations of the Infinitesimal Calculus*. Boston, Mass.: Prindle Weber & Schmit

Klein, F. (2016). *Elementary Mathematics from a Higher Standpoint. Vol. III.* (Tr.Menghini &Baccaglini-Frank). Berlin: Springer

Meyer, J.H.F, Land, R. & Baille, C. (2010). *Threshold concepts and transformational learning*. Rotterdam: Sense Publishers

Robinson, A. (1966). *Non-standard Analysis*. Amsterdam : North Holland

Rogers, L. & Pope, S. (2015). A brief history of quadratic equations for mathematics educators. *Proceedings of the British Society for Research into Learning Mathematics,* 35(3), 90-95.

Russ, S. (2004) (Trans.) *The Mathematical Works of Bernhard Bolzano.* Oxford: O.U.P.

Tall, D. O. (1980). *Intuitive infinitesimals in the calculus.* Accessed 9 January 2017
https://homepages.warwick.ac.uk/staff/David.Tall/pdfs/dot1980c-intuitive-infls.pdf

Zukav, G. (1979). *The Dancing Wu Li Masters: An overview of the New Physics*. New York: Bantam Books