2017

An Introduction to a Rigorous Definition of Derivative

Dave Ruch
ruch@msudenver.edu

An Introduction to a Rigorous Definition of Derivative

David Ruch*

November 22, 2021

1 Introduction

The concept of a derivative evolved over a great deal of time, originally driven by problems in physics and geometry. Mathematicians found methods that worked, but justifications were not always very convincing by modern standards. A co-inventor of calculus, Isaac Newton (1642–1727), is generally regarded as one of the most influential scientists in human history. Newton used the term *fluxion* for a special case of what we now call a derivative. He developed most of his methods around 1665, but did not publish them for many years. Newton's later writings about his fluxion methods influenced many mathematicians.

Newton and fluxions. The modern definition of a function had not yet been created when Newton developed his fluxion theory. The context for Newton's methods of fluxions was a particle "flowing" or tracing out a curve in the x-y plane. The x and y coordinates of the moving particle are *fluents* or flowing quantities. The horizontal and vertical velocities were the *fluxions* of x and y, respectively, associated with the flux of time. In the excerpt below from his first full exposition of the theory of fluxions in his *Tractatus de quadratura curvarum* (*Treatise on the Quadrature of Curves*), [Newton, 1704], Newton considered the curve $y = x^n$ and wanted to find the fluxion of y.

Let the quantity x flow uniformly, and let it be proposed to find the fluxion of x^n.

In the same time that the quantity x, by flowing, becomes $x + o$, the quantity x^n will become $(x+o)^n$, that is, $x^n + nox^{n-1} + \dfrac{n^2-n}{2}oox^{n-2} + \&c.$[1] And the augments o and $nox^{n-1} + \dfrac{n^2-n}{2}oox^{n-2} + \&c.$ are to one another as 1 and $nx^{n-1} + \dfrac{n^2-n}{2}ox^{n-2} + \&c.$

Now let these augments vanish, and their ultimate ratio will be 1 to nx^{n-1}.

*Department of Mathematical and Computer Sciences, Metropolitan State University of Denver, Denver, CO; ruch@msudenver.edu.

[1] In Newton's day, mathematicians frequently wrote out a variable squared as "oo" instead of "o^2", either for emphasis or for typeseting reasons.

$\boxed{\textbf{Task 1}}$ Write out the algebraic details of Newton's fluxion method for $n = 3$ using modern algebraic notation to find the fluxion of x^3. Sketch the curve $y = x^3$ and label the key quantities for Newton's fluxion method.

$\boxed{\textbf{Task 2}}$ Convert Newton's argument that $(x^n)' = nx^{n-1}$ to one with modern limit notation for the case where n is a positive integer. You may use modern limit laws.

It is important to remember that when Newton developed his fluxion method, there was no theory of mathematical limits. Critics of Newton's fluxion method were not happy with having augment o be a seemingly nonzero value at the beginning of the method, and then having the augment o "vanish" at the end of the argument. Indeed, the philosopher and theologian George Berkeley (1685–1753) wrote a 1734 book entitled *The Analyst; Or, A Discourse Addressed to an Infidel Mathematician*, [Berkeley, 1734], that attacked Newton's methods. Berkeley rejected:

> ... the fallacious way of proceeding to a certain Point on the Supposition of an Increment, and then at once shifting your Supposition to that of no Increment .... Since if this second Supposition had been made before the common Division by o, all had vanished at once, and you must have got nothing by your Supposition. Whereas by this Artifice of first dividing, and then changing your Supposition, you retain 1 and nx^{n-1}. But, notwithstanding all this address to cover it, the fallacy is still the same. And what are these Fluxions? The Velocities of evanescent Increments? And what are these same evanescent Increments? They are neither finite Quantities nor Quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities?

$\boxed{\textbf{Task 3}}$ What do you think of Berkeley's criticisms? Is the value of o in Newton's text both an increment and "not an increment"? How might Newton have responded?

Some historians argue that Berkeley was attacking calculus itself, while others argue that he was instead claiming that calculus was no more logically rigorous than theology. Certainly, many a Calculus I student has had the same feeling that Berkeley had about the limiting process! As we shall see in Section 2, by the early 1800s the mathematical community was closer to putting limits and derivatives on a firmer mathematical foundation.

L'Hôpital on the differential of a product. While Newton was working on his fluxion methods, Gottfried Leibniz (1646–1716) was independently developing calculus during the same time period. Many of Leibniz's ideas appeared in the book *Analyse des Infiniment Petits pour l'intelligence des Lignes Courbes* (*Analysis of the Infinitely Small to understand Curves*), [L'Hôpital, 1696] , written by Guillaume L'Hôpital (1661–1701). A key idea for Leibniz was the *differential.* Here is an excerpt from L'Hôpital's book.

Definition II. The infinitely small part whereby a variable quantity is continually increased or decreased, is called the differential of that quantity.

...

Proposition II. To find the differentials of the product of several quantities multiplied.

The differential of xy is $y\,dx + x\,dy$: for y becomes $y + dy$, when x becomes $x + dx$; and therefore xy becomes $xy + y\,dx + x\,dy + dx\,dy$. Which is the product of $x + dx$ into $y + dy$, and the differential thereof will be $ydx + x\,dy + dx\,dy$, that is, $y\,dx + x\,dy$: because $dx\,dy$ is a quantity infinitely small, in respect of the other terms $y\,dx$ and $x\,dy$: For if, for example, you divide $y\,dx$ and $dx\,dy$ by dx, we shall have the quotients y and dy, the latter of which is infinitely less than the former.

Task 4 Use L'Hôpital's definition of a differential, some algebra and your own words to explain why the differential of xy is $ydx + x\,dy + dx\,dy$.

Task 5 What do you think of L'Hôpital's argument that he could eventually ignore $dx\,dy$ because it is "a quantity infinitely small, in respect of the other terms $y\,dx$ and $x\,dy$"?

A number of mathematicians were stung by Berkeley's criticisms and attempted to define the derivative in a more logically satisfying manner. One of the leaders of this movement was Joseph-Louis Lagrange (1736–1813). While Lagrange was not entirely successful in his efforts, his ideas were influential, particularly with Augustin-Louis Cauchy (1789–1857), whom he knew in Paris. Cauchy developed a theory of limits that included the derivative and the definite integral in his textbook *Résumé des leçons données à L'École Royale Polytechnique sur le calcul infinitésimal* (*Summary of lessons given on the infinitesimal calculus at the Royal Polytechnic School*), [Cauchy, 1823], from which we will read in Section 2 of this project. After reading Cauchy in Section 2, you will have an opportunity to write and prove a modern version of L'Hôpital's Proposition II, using L'Hôpital's ideas.

2 Cauchy's Definition of Derivative

As a prelude to reading about Cauchy's definition of derivative, it is worthwhile to read an excerpt on limits from his 1823 *Calcul Infinitésimal*.

When the values successively attributed to the same variable approach a fixed value indefinitely, in such a manner as to end up differing from it by as little as we wish, this last value is called the *limit* of all the others.

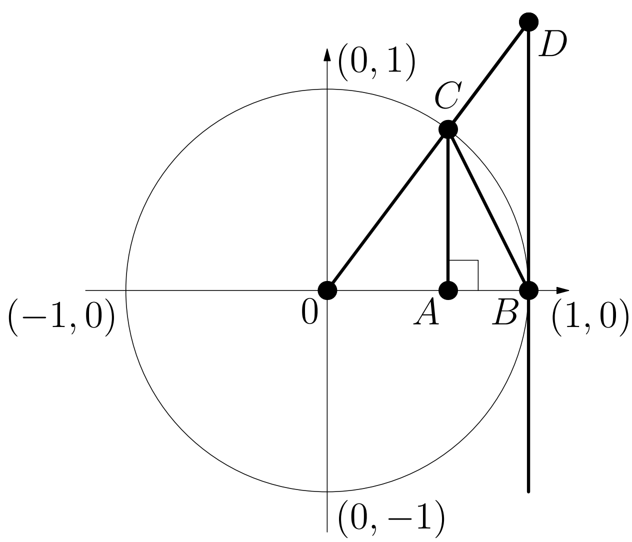
...

We obviously have, for very small numerical values[2] of α,

$$\frac{\sin\alpha}{\sin\alpha} > \frac{\sin\alpha}{\alpha} > \frac{\sin\alpha}{\tan\alpha}.$$

By consequence, the ratio $\frac{\sin\alpha}{\alpha}$ always contained between the two quantities $\frac{\sin\alpha}{\sin\alpha} = 1$, and $\frac{\sin\alpha}{\tan\alpha} = \cos\alpha$, the first of which serves to limit the second, will itself have unity for a limit.

Task 6 Use the following diagram with some basic trigonometry to justify the inequality above. Recall that the area of a sector with radius 1 and angle α is $\alpha/2$.



Task 7 Comment on Cauchy's definition of limit and his proof that $\lim_{\alpha\to 0}\sin\alpha/\alpha = 1$. What adjustments, if any, are needed to conform to the modern definition of limit?

Here is Cauchy on the derivative. In this passage from [Cauchy, 1823], when Cauchy used the term "between" two values, he meant to include the two values.

[2] Cauchy means the magnitude of α here. Cauchy was working before the creation of *absolute value* by mathematicians.

When the function $y = f(x)$ remains continuous between two given limits of the variable x, and that we assign to this variable a value contained between the two limits in question, an infinitely small increment attributed to the variable produces an infinitely small increment of the function itself. By consequence, if we then set $\Delta x = i$, the two terms of the *ratio of differences*

$$\frac{\Delta y}{\Delta x} = \frac{f(x+i) - f(x)}{i} \tag{1}$$

will be infinitely small quantities. But, while these two terms indefinitely and simultaneously will approach the limit of zero, the ratio itself may be able to converge toward another limit, either positive or negative. This limit, when it exists, has a determined value for each particular value of x; but, it varies along with x. Thus, for example, if we take $f(x) = x^m$, m designating an integer number, the ratio between the infinitely small differences will be

$$\frac{(x+i)^m - x^m}{i} = mx^{m-1} + \frac{m(m-1)}{1 \cdot 2} x^{m-2} i + \cdots + i^{m-1},$$

and it will have for a limit the quantity mx^{m-1}, that is to say, a new function of the variable x. It will be the same in general, only the form of the new function, which will serve as the limit of the ratio $\frac{f(x+i)-f(x)}{i}$, will depend on the form of the proposed function $y = f(x)$. To indicate this dependence, we give to the new function the name of *derivative function*, and we represent it, with the help of an accent mark, by the notation

$$y' \quad \text{or} \quad f'(x).$$

Let's introduce some modern notation and terminology for Cauchy's definition. First assume, like Cauchy, that $f : [a,b] \to \mathbb{R}$ is continuous on $I = [a,b]$. If f' exists at some x, we say f is **differentiable** at x. If f is differentiable at each x in an interval I, we say f is differentiable on I.

Task 8 We have seen two explanations of the derivative rule $(x^n)' = nx^{n-1}$.

(a) Compare and contrast the arguments of Cauchy and Newton.

(b) Rewrite Cauchy's argument in your own words with modern terminology, using modern properties of limits.

After the excerpt above, Cauchy gave some derivative examples, including:

$$\text{For } y = \sin x, \quad \frac{\Delta y}{\Delta x} = \frac{\sin \frac{1}{2} i}{\frac{1}{2} i} \cos\left(x + \frac{1}{2} i\right), \quad y' = \cos x \tag{2}$$

Task 9 (a) Use a trigonometric identity to justify Cauchy's expression for $\frac{\Delta y}{\Delta x}$ in (2).

(b) Use Cauchy's expression for $\frac{\Delta y}{\Delta x}$, limit properties and the continuity of the cosine function to show that $y' = \cos x$.

Cauchy defined the derivative $f'(x)$ when the limit $\lim_{i\to 0} \frac{f(x+i)-f(x)}{i}$ exists. Implicitly this means $f'(c)$ may not exist for some c values.

Task 10 Define $f(x) = |x|$.

(a) Show that $f'(x)$ is 1 for $x > 0$, and is -1 for $x < 0$.

(b) Show that $f'(0)$ does not exist.

The subtleties of differentiability were not universally understood in the mathematics community for some time after Cauchy. A fine example of this comes from the correspondence between the mathematicians Gaston Darboux (1842–1917)and Guillaume Jules Hoüel (1823–1886) during the 1870s.[3] Hoüel was writing a calculus textbook and asked for Darboux's feedback. Darboux was unhappy with Hoüel's proof rigor, and he introduced the function $M(x) = x^2 \sin(1/x)$ near $x = 0$ to make his point. This strange function, termed a "monster function" by Henri Poincaré (1894–1912), appeared again in an 1884 correspondence between mathematicians Giuseppe Peano (1858–1932) and Camille Jordan (1838–1932) regarding rigor in a proof of the Mean Value Theorem. Clearly an analysis of this function is worthwhile!

Task 11 Make a sketch of the graph of Darboux's function M near zero, using technology as needed, and use it to help answer the following.

(a) Identify the x-intercepts for the graph of $y = M(x)$.

(b) Identify the x values for the points where the graph of $y = M(x)$ coincides with the graph of $y = x^2$ or the graph of $y = -x^2$.

(c) Explain the relationship between the graphs of $y = M(x)$, $y = x^2$ and $y = -x^2$.

Task 12 Consider Darboux's function M.

(a) Find a value for $M(0)$ so that M is continuous at 0. Prove your assertion.

(b) Find a value for $M'(0)$ so that M is differentiable at 0. Prove your assertion with the definition of derivative.

Task 13 Consider Darboux's function M.

(a) Use standard calculus rules to find $M'(x)$ for $x \neq 0$.

(b) Is M' continuous at $x = 0$? Prove your assertion.

Hint: Consider M' at the intercepts of M.

[3] A more detailed discussion of the correspondence is given in the primary source project "Rigorous Debates over Debatable Rigor: Monster Functions in Introductory Analysis" [Barnett, 2016].

After trading letters with Hoüel, Darboux wrote him the following in 1875:

You have not addressed the nature of my objection. ... I have told you that, according to (your) rule of composite functions, we obtain

$$\frac{dy}{dx} = 2x \sin \frac{1}{x} - \cos \frac{1}{x},$$

an expression that is indeterminate for $x = 0$, even though, according to first principals, the derivative is perfectly determined, it is zero. For your methods to be sound, you will need to explain very clearly what part of your reasoning is deficient in this particular case. Without that your proofs are not proof.

Task 14 How would you advise Hoüel to respond, based on your work above?

Cauchy's definition of derivative allows mathematicians to carefully analyze strange function behavior, as we have seen with Darboux's function M. It can also be used effectively to rigorously prove important derivative rules such as the product rule.

Task 15 Suppose that $f, g : [a, b] \to \mathbb{R}$ and both f and g are differentiable on $[a, b]$. Using L'Hôpital's proof ideas from his Proposition II (page 3), we can construct a modern proof that $f \cdot g$ is differentiable on $[a, b]$. Viewing f and g as functions of t, we can associate L'Hôpital's x with $f(t)$ and his y with $g(t)$. Using a small, finite increment Δt in place of L'Hôpital's "infinitely small" increase, we can associate L'Hôpital's dx with Δf defined by $\Delta f = f(t + \Delta t) - f(t)$. Similarly associate L'Hôpital's dy with Δg defined by $\Delta g = g(t + \Delta t) - g(t)$.

(a) Note that $f(t + \Delta t) = f(t) + \Delta f$ and write a similar expression for g.

(b) Use part (a) to show that

$$f(t + \Delta t) \cdot g(t + \Delta t) - f(t) \cdot g(t) = g(t) \Delta f + f(t) \Delta g + \Delta f \cdot \Delta g.$$

(c) Use part (b) and modern limit properties to show that

$$\lim_{\Delta t \to 0} \frac{f(t + \Delta t) \cdot g(t + \Delta t) - f(t) \cdot g(t)}{\Delta t} = g(t) f'(t) + f(t) g'(t).$$

Hint: You will need to use the fact that differentiable functions are necessarily continuous, an important property that is proven in most first-year calculus textbooks.

(d) Summarize your results from parts (a)–(c).

3 Conclusion

Mathematicians figured out how to define and apply the derivative well before they put the concept on a firm mathematical foundation. After giving his definition of the derivative, Cauchy went on to prove many theorems with it, including the all important Mean Value Theorem. This result is fundamental in proving many results from an introductory calculus course. Unfortunately, Cauchy and other great mathematicians occasionally fell into subtle traps while trying to use the definition of derivative. From a modern point of view, Cauchy's proof of the Mean Value Theorem appears to be flawed due to these subtleties, and a completely correct proof was not published for several more decades.

References

J. H. Barnett. Rigorous debates over debatable rigor: Monster functions in introductory analysis. 2016. Primary Source Project available at `https://digitalcommons.ursinus.edu/triumphs_analysis/10/`.

G. Berkeley. *The Analyst; Or, A Discourse Addressed to an Infidel Mathematician.* J. and R. Tonson and S. Draper, London, 1734.

D. Cates. *Cauchy's Calcul Infinitésimal – A Complete English Translation.* Fairview Academic Press, Walnut Creek CA, 2012.

A. L. Cauchy. *Résumé des leçons données à l'École royale polytechnique sur le calcul infinitésimal.* De Bure, Paris, 1823. Translation by Dennis Cates in [Cates, 2012].

G. L'Hôpital. *Analyse des Infiniment Petits pour l'intelligence des Lignes Courbes (Analysis of the Infinitely Small to understand Curves).* Montalant, Paris, 1696. English translation by E. Stone appears (with some modification) in [Struik, 1969, pp. 303–312].

I. Newton. *Tractatus de quadratura curvarum (Treatise on the Quadrature of Curves).* [S.l.] : [s.n.], London, 1704. English translation by J. Stewart appears in [Struik, 1969, pp. 313–316].

Notes to Instructors

PSP Content: Topics and Goals

This Primary Source Project (PSP) is designed to introduce the derivative with some historical background for a course in Analysis. More specifically, its content goals are to help students to:

1. Learn the basics of Newton's fluxion method and Berkeley's criticisms.
2. Learn the basics of the differential method with the product rule.
3. Develop a modern derivative definition based on Cauchy's definition.
4. Develop facility with the modern definition of derivative by using it to rigorously find derivatives of some elementary functions, and to prove the necessity of continuity for differentiability.
5. Analyze subtleties of the derivative definition using Darboux's monster function $M(x) = x^2 \sin(1/x)$.

Student Prerequisites

This project assumes students have done a rigorous study of limits and continuity for real-valued functions.

PSP Design and Task Commentary

This project consists of two main sections, followed by a brief conclusion.

Section 1 Introduction

This material is included mostly to motivate the need for a rigorous definition of the derivative, with some historical perspective. Newton's example of $(x^m)'$ is also discussed by Cauchy later in the project. L'Hôpital's argument for the Product Rule with differentials is used to motivate a modern proof of the Product Rule for derivatives in Task 2 near the end of the project.

Section 2 Cauchy

Cauchy is generally credited with being among the first to define and use the derivative in a near modern fashion. His definitions of limit and $\lim \sin \alpha/\alpha$ are given so that students have some background on his writing style and mathematical language. This latter limit reappears later in the project, in finding $(\sin x)' = \cos x$.

Instructors may sample the task set after Cauchy's definition for classroom examples or homework problems. Most investigate nice properties, and then subtle behavior is investigated using Darboux's monster function $M(x) = x^2 \sin(1/x)$. More detailed exploration is given in Janet Heine Barnett's project "Rigorous Debates over Debatable Rigor: Monster Functions in Introductory Analysis" [Barnett, 2016]. This monster function M appears in nearly every Introductory Analysis text, but its interesting historical significance is rarely mentioned.

Suggestions for Classroom Implementation

This is roughly a one week project under the following methodology (basically David Pengelley's "A, B, C" method described on his website (`https://www.math.nmsu.edu/~davidp/`):

1. Students do some advanced reading and light preparatory exercises before each class. This should be counted as part of the project grade to ensure students take it seriously. Be careful not to get carried away with the exercises or your grading load will get out of hand! Some instructor have students write questions or summaries based on the reading.

2. Class time is largely dedicated to students working in groups on the project - reading the material and working exercises. As they work through the project, the instructor circulates through the groups asking questions and giving hints or explanations as needed. Occasional student presentations may be appropriate. Occasional full class guided discussions may be appropriate, particularly for the beginning and end of class, and for difficult sections of the project. I have found that a "participation" grade suffices for this component of the student work. Some instructors collect the work. If a student misses class, I have them write up solutions to the exercises they missed. This is usually a good incentive not to miss class!

3. Some exercises are assigned for students to do and write up outside of class. Careful grading of these exercises is very useful, both to students and faculty. The time spent grading can replace time an instructor might otherwise spend preparing for a lecture.

If time does not permit a full implementation with this methodology, instructors can use more class time for guided discussion and less group work for difficult parts of the project.

LaTeXcode of this entire PSP is available from the author by request to facilitate preparation of 'in-class task sheets' based on tasks included in the project. The PSP itself can also be modified by instructors as desired to better suit their goals for the course.

Sample Implementation Schedule (based on a 50-minute class period)

Full implementation of the project can be accomplished in 3 class days, as outlined below.

Students read through the introductory Newton material and do Task 1 before the first class. After discussing their results at the beginning of Class 1, students do Task 2, read the Berkeley excerpt and start Task 3 (philosophical instructors may want to take more time here). They next read the L'Hôpital excerpt and work on Tasks 4, 5. Polishing Tasks 3 and 5 can be done for homework.

As preparation for Class 2, students read the first Cauchy excerpt and do Tasks 6, 7. After discussing their results at the beginning of Class 2, students read the second Cauchy excerpt, then work on Tasks 8, 11, 12. Tasks 9, 10 can be assigned for homework.

As preparation for Class 3, students work on Task 13. After discussing their results at the beginning of Class 3, students read the Darboux excerpt and work on Tasks 14, 15. Depending on time, polishing any of these assignments can be done for homework.

Connections to other Primary Source Projects

The following additional projects based on primary sources are also freely available for use in an introductory real analysis course; the PSP author name for each is listed parenthetically, along with the project topic if this is not evident from the PSP title. Shorter PSPs that can be completed in at most 2 class periods are designated with an asterisk (*). Classroom-ready versions of the last two projects listed can be downloaded from `https://digitalcommons.ursinus.edu/triumphs\_topology`; all other listed projects are available at `https://digitalcommons.ursinus.edu/triumphs\_analysis`.

- *Why be so Critical? 19th Century Mathematics and the Origins of Analysis** (Janet Heine Barnett)
- *Investigations into Bolzano's Bounded Set Theorem* (David Ruch)
- *Stitching Dedekind Cuts to Construct the Real Numbers* (Michael Saclolo)
  Also suitable for use in an Introduction to Proofs course.
- *Investigations Into d'Alembert's Definition of Limit** (David Ruch)
  A second version of this project suitable for use in a Calculus 2 course is also available.
- *Bolzano on Continuity and the Intermediate Value Theorem* (David Ruch)
- *Rigorous Debates over Debatable Rigor: Monster Functions in Real* (Janet Heine Barnett; properties of derivatives, Intermediate Value Property)
- *The Mean Value Theorem*(David Ruch)
- *The Definite Integrals of Cauchy and Riemann* (David Ruch)
- *Henri Lebesgue and the Development of the Integral Concept** (Janet Heine Barnett)
- *Euler's Rediscovery of e** (David Ruch; sequence convergence, series & sequence expressions for e)
- *Abel and Cauchy on a Rigorous Approach to Infinite Series* (David Ruch)
- *The Cantor set before Cantor** (Nicholas A. Scoville)
  Also suitable for use in a course on topology.
- *Topology from Analysis** (Nicholas A. Scoville)
  Also suitable for use in a course on topology.

Acknowledgments

The development of this student project has been partially supported by the TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS) Program with funding from the National Science Foundation's Improving Undergraduate STEM Education Program under Grant Nos. 1523494, 1523561, 1523747, 1523753, 1523898, 1524065, and 1524098. Any opinions, findings, and conclusions or recommendations expressed in this project are those of the author and do not necessarily represent the views of the National Science Foundation.

For more information about TRIUMPHS, visit `https://blogs.ursinus.edu/triumphs/`.