

Teaching Calculus with Infinitesimals

Rebecca Vinsonhaler

University of Wisconsin - Madison

Follow this and additional works at: <https://scholarship.claremont.edu/jhm>

Recommended Citation

Vinsonhaler, R. "Teaching Calculus with Infinitesimals," *Journal of Humanistic Mathematics*, Volume 6 Issue 1 (January 2016), pages 249-276. DOI: 10.5642/jhummath.201601.17 . Available at: <https://scholarship.claremont.edu/jhm/vol6/iss1/17>

©2016 by the authors. This work is licensed under a Creative Commons License.

JHM is an open access bi-annual journal sponsored by the Claremont Center for the Mathematical Sciences and published by the Claremont Colleges Library | ISSN 2159-8118 | <http://scholarship.claremont.edu/jhm/>

The editorial staff of JHM works hard to make sure the scholarship disseminated in JHM is accurate and upholds professional ethical guidelines. However the views and opinions expressed in each published manuscript belong exclusively to the individual contributor(s). The publisher and the editors do not endorse or accept responsibility for them. See <https://scholarship.claremont.edu/jhm/policies.html> for more information.

Teaching Calculus with Infinitesimals

Cover Page Footnote

Special thanks to Terry Millar for all his support and guidance.

Teaching Calculus with Infinitesimals

Rebecca Vinsonhaler

Department of Curriculum & Instruction, University of Wisconsin Madison, USA
vinsonhaler@wisc.edu

Synopsis

This article argues that first semester calculus courses for non-mathematics majors should be taught using infinitesimals. This applies to both high school and undergraduate calculus courses. The use of infinitesimals in calculus, though more intuitive than the approach developed in the 19th century, has been controversial for over two millennia. However, in the 20th century their use was shown to be equa-consistent with the approach developed in the 19th century. Here I first provide a brief history of infinitesimals, why they were controversial, and how they were finally put on a firm footing. Next I illustrate the intuitive nature of the use of infinitesimals. Thus I conclude that at least students not continuing on to more advanced analysis courses would be better served by learning calculus via infinitesimals.

Keywords: *calculus reform; history of mathematics; non-standard analysis*

1. Introduction

Infinitesimals not only could serve as an important intuitive aid to learning some key concepts in calculus, they also have a fascinating history. Infinitesimals were used as early as the Ancient Greeks by mathematicians such as Archimedes, were still being used by Leibniz in 1600 C.E. [14] and informally by physicists and mathematicians until at least the end of the 19th century. However, because it would not be until the 20th century that infinitesimals were shown to be realizable without contradiction, mathematicians usually sought alternative methods of proof when presenting their results to the “public” [10]. Thus, for example, Archimedes replaced his arguments

using “indivisibles” (found, for example, in his *The Method of Mechanical Theorems*¹) with extensions of Eudoxus’s method of exhaustion. Although Leibniz and Newton invented calculus by using infinitesimals, the latter were not rigorously established until the 1960s, forcing the development of a formal calculus along a different route. This formalized calculus, referred to as modern calculus, or *standard analysis*, was made possible by the work of such mathematicians as Cauchy and Weierstrass with the introduction of the “epsilon-delta” definition of limit [5].

There are deep ironies in this history. For example, standard analysis was certainly facilitated by the acceptance of a rigorous definition of real numbers. And this in turn was facilitated by the invention of set theory by Cantor. One twist here is that Cantor’s thesis advisor, Kronecker, rejected the concept of an infinite set, something critical to both set theory and the definition of reals that evolved. Another unexpected fact is that if historical weight counted for anything, then the moniker “standard analysis” would have to go to the infinitesimal approach, since it was used by some of the most prominent mathematicians and physicists working in the field, from Archimedes to those in the 18th century. Instead, the use of infinitesimals in calculus is dubbed “non-standard analysis”. Once standard analysis was developed, infinitesimals were exiled.

2. Standard Analysis vs. Non-Standard Analysis

Standard analysis uses the “epsilon-delta” definition of limit which in turn is used to define such key concepts as convergence, continuity, derivative and integral. Standard analysis is currently used as the rigorous foundation for a majority of traditional calculus courses. “Non-standard” analysis, on the other hand, uses infinitesimals and what are called “hyperreals”, and most importantly, does not depend on a notion of “limit” to define the fundamental concepts of elementary calculus. To speak loosely, the limit concept formalizes a concept of converging to small differences, whereas with the concept of the infinitesimal, one is “already there”.

¹This work was only rediscovered early in the 20th century in what has come to be called *Archimedes Palimpsest* — a 10th century Byzantine copy had been overwritten with Christian religious text by 13th century monks.

Let us look at the notion of the continuity of a function. Intuitively a function with domain and range that are subsets of the real numbers is continuous at a real number r if: 1) the number r is in the domain of the function; and 2) for real numbers s in the domain of the function that are close to r , the value of the function at s is close to the value of the function at r . Thus logically one must check a “for all” statement: “*for all real numbers s in the domain of the function close to r [statement]*”. And if one uses infinitesimals then this is exactly the logical complexity of the statement, since “ *s is close to r* ” is translated formally into “*the arithmetic difference of s and r is an infinitesimal*”; denoted by $r \approx s$. However, if infinitesimals are not available, one must search for a different translation of “ *s is close to r* ”. The solution introduced by Weierstrass required a definition whose logical complexity is “*for all $\epsilon > 0$ there exists a $\delta > 0$ such that for all s [statement]*” (where the missing statement is “*if the absolute value of the difference of s and r is less than δ , then the absolute value of the difference of the function values at s and r is less than ϵ* ”). This is the famous or infamous (depending on your experience in calculus) “epsilon-delta argument”.

In some sense the techniques introduced by Weierstrass for standard analysis reflected the method of exhaustion introduced by Eudoxus more than 2400 years earlier. Once it was understood by the ancient Greeks that there were geometric lengths whose ratios were not rational (such as the ratio of the hypotenuse to either of the other sides in an isosceles right triangle), a decision had to be made: tackle the question of the existence of numbers that were not rational, or find a way around the issue. Eudoxus did the latter. For he said that two lengths were equal if, using modern terminology, the sets of rational lengths less than both were equal, and the sets of rational lengths greater than both were equal. This comes very close to the definition of the real numbers introduced by Dedekind in the 19th century. Eudoxus’s method allowed the development of two and three dimensional versions used to prove such things as the relationship of the area and circumference of a circle or the area and volume of a sphere (all by Archimedes). Recall that Archimedes convinced himself of the truth of these relationships by using infinitesimals (and in the case of the sphere also using the Law of the Lever – see below), but then resorted to a version of the method of exhaustion probably because he recognized that an argument using infinitesimals would not be considered a proof.

We will see below in more detail how standard analysis avoids the question of the possible existence of a useful extension of the reals to what are now called the hyperreals by introducing an alternation of quantifiers that dynamically bounds the behavior of a function in a way similar to how Euclid, Archimedes, and others of that era statically captured certain lengths, areas, and volumes by variations on Eudoxus's method of exhaustion. The invention of the reals, though strongly resisted by constructivists such as Kronecker, facilitated the rigorous development of analysis and many other fields. The invention of the reals also simplified the number of cases necessary in some of the arguments by such luminaries as Euclid and Archimedes. It is the thesis of this paper that the invention of a sound system that incorporates infinitesimals, though passively resisted by most mathematicians of this and the previous century, could facilitate a more nuanced understanding of calculus by a larger part of humanity.

3. Historical Perspective - the development of both approaches and the reason infinitesimals are not popular

Today most mathematicians learn standard analysis first, few encounter non-standard analysis, and an even smaller number actually learn it. This is not surprising, since it took until the 1930s and 1960s respectively, for the work of logicians Kurt Gödel (Compactness Theorem) and Abraham Robinson (*Non-standard analysis*, North-Holland Publishing Co., Amsterdam 1966) to put the use of infinitesimals on a firm logical foundation. With Robinson's formalization non-standard analysis was proven sound, but the mathematical community did not readily warm to the concept. Although this formalization was not quickly accepted, it did eventually lead to attempts to teach calculus using non-standard analysis. Logicians such as Robinson and Jerome Keisler wrote calculus textbooks using infinitesimals and hyperreals, which are the basis of non-standard analysis. These texts more closely modeled the way Leibniz and others had considered the subject, and reduced the formal logical complexity of central notions in calculus, such as that of continuity, derivative, and integral (see below).

Although there is evidence (see Section 12 below) that students found the approach more intuitively accessible, the pedagogical approach did not take hold. A probable cause was that practicing mathematicians were unable to embrace the approach (Keisler, personal communication, April 17, 2014).

Standard analysis is currently used to teach almost all calculus courses at both the high school and college level. In this paper I argue that infinitesimal calculus should be taught in place of a standard first-semester calculus course for non-mathematics majors. Not only is there evidence that infinitesimals accord with the intuitions of students, but they also have a long and interesting history. Highlighting aspects of the intellectual history of mathematics can furthermore strengthen the idea that mathematics is not static and that it is not apolitical.

In what follows I will briefly explain the history behind infinitesimals, explain what they are and why they are intuitive, and reflect on some attempts of their integration into classrooms. The history of infinitesimals is an interesting one, and can be used by teachers to motivate lessons as well as to illustrate the usefulness of mathematics. Story telling can be a valuable tool when teaching mathematics [7]. Regardless of whether the stories are true, they capture our attention, and add to any lesson. The story that Newton's study of gravity was prompted by an apple falling on his head is fictitious. The story of Archimedes running naked from the baths in Sicily after realizing how to decide if the king's crown was made of pure gold may also be fictitious [14]. However, these stories help illustrate how and why mathematics was invented. They are entertaining and grab our attention.

Teaching calculus using infinitesimals offers us a great opportunity to introduce relevant historical and cultural contexts in mathematics education. Not only would students be more engaged in lessons, they could also develop an appreciation for the invention of mathematics. Furthermore, a historical approach serves as an opportunity to understand the important role mathematics plays in our society. By learning the history of calculus, students may recognize that mathematics is (at least partially) socially constructed, that alternate forms of mathematics exist, and that mathematics can and should be challenged. Students may also be excited by the notion that even some professional mathematicians do not fully understand or are not comfortable with the foundations of infinitesimal calculus, thus making it a bit of a thrill to learn.

In the overview of the history of infinitesimals below, there are many stories to be told, some about the mathematicians themselves, some about the history of mathematics, and others about the role mathematics has played in society.

4. The Mathematicians Who Believed in Infinitesimals

As mentioned above, Archimedes used infinitesimals as early as 200 B.C. [14]. First he set down axioms (for an area that is now called *statics*) to derive his Law of the Lever. He then found a clever way to position a fulcrum with “indivisible” slices of a right circular cylinder a certain distance from the fulcrum in equilibrium with two corresponding slices of a right circular cone and sphere at a fixed distance on the other side of the fulcrum, to prove the relation between the volume of a sphere and the (known) volumes of the cylinder and right circular cone [10]. He considered this his greatest achievement and most valuable contribution to mankind. Plutarch later wrote, “although he made many excellent discoveries, he is said to have asked his kinsmen and friends to place over the grave where he should be buried a cylinder enclosing a sphere, with an inscription giving the proportion by which the containing solid exceeds the contained” [14]. Thus, although his gravesite is still unknown, sources such as Plutarch imply that on the grave of Archimedes was an engraved picture of a sphere inscribed in a cylinder, with the relationships of the volumes and surface areas next to the illustration. Archimedes used what is now known as the method of indivisibles to derive this relationship. This process is considered one of the earliest steps toward integral calculus. The term “method of indivisibles” was not formally used until 1635 in Bonaventura Cavalieri’s text *Geometria indivisibilibus*, but the process is the same as that which was used by Archimedes (page 387). And as mentioned above, it is of interest that when Archimedes published his work on deriving this relationship, he used the method of exhaustion to solve the same problem, thus eliminating infinitesimals from the publication [14].

In an interview on his book on infinitesimals Amir Alexander asks, “How many parts can you divide a line into?” [11]. The idea is that it can be cut in half, and then in half again and so on until the pieces can’t be cut in half anymore. Of course, if you can’t cut a piece in half this would mean that the length of that piece was zero, because if it had any length at all then it could be cut in half again. This idea is similar to the notion of an infinitesimal. It is easy to see that this idea would be confusing and therefore problematic. In the 1600s it was the use of this concept that cast suspicion on the logical soundness of mathematics in general. In the 1600s mathematics was seen by many as logical and decisive and was meant to be a way to see

the world. Since the styles of reasoning of mathematicians were important to the church in the 1600s, the Jesuits were upset because they hoped to use this same sound logic to prove the Protestants wrong about their religion [11]. When infinitesimals came into play, the Jesuits began to question their belief in mathematics. The Jesuits figured that if mathematicians couldn't be trusted, then their methods of reasoning couldn't be trusted to help the Jesuits overthrow the Protestants [11]. Thus the Jesuits verbally battled mathematicians such as Galileo on the concept of infinitesimals. Luckily, infinitesimals were not completely drowned out, however, this complicated their use and laid the foundation for the negative reputation they still have within the mathematics community [5].

Although infinitesimals were still controversial during the 1670s and 1680s, Gottfried Wilhelm Leibniz used them in his development of infinitesimal calculus (Ely, 2010). After Leibniz laid down some rules and heuristics for working with infinitesimals his method of calculus was used into the 1700s, by many well-known mathematicians such as Euler. The most vehement attack on the infinitesimal came from Bishop Berkeley (of the Catholic church). In 1734 he wrote a famous paper attacking Newton and the infinitesimal. However, historians such as H. Bos ([2]; also see [1]) believed that the loose definition of infinitesimals is what allowed mathematics in the 18th century to advance so quickly. Thus the tremendous number of discoveries in mechanics, probability theory, astronomy, and of course calculus during that time period are likely in part a direct result of this flexible definition of infinitesimals [5]. However, in the early 1800s contradictory results in mathematics emerged and thus prompted mathematicians like Cauchy and Bolzano to work on proving things without the use of infinitesimals. In the 1860s Weierstrass ended the use of infinitesimals with his rigorously proven epsilon-delta definition of limit (Keisler, personal communication, April 17, 2014). It was at this time that mathematicians stopped using infinitesimals entirely. However, physicists and engineers continued to use them, and continued to find them useful as well as mathematically sound.

It was not until the 20th century that mathematicians could explain infinitesimals rigorously and reliably. In short, mathematicians could not formalize infinitesimals not because infinitesimals were contradictory, but because the language of logic had not developed enough to do so. The first step was what could be called the formalization of semantics – the world of ref-

erents for mathematical entities – that started in the 1870s with the work of two mathematicians, Cantor and Dedekind, when they initiated the study of what is now set theory [6]. Most mathematicians now agree that (formally) the objects of their various studies are sets and therefore governed in principle by the axioms of set theory. For example, though mathematicians who study number theory or geometry typically have very non-set theoretic intuitions (in other words they visualize patterns of numbers or shapes of figures), in principle they pay homage to set theory. Set theory allowed mathematicians to uniformly define and group objects and describe relationships between two sets of mathematical objects.

The “flip side” of semantics, or the world of referents, is symbol or the world of language, syntax, and proof. The emergence of tools sufficient to prove the relative consistency of infinitesimals required the development of what is now called mathematical logic. Although Aristotle gave us the formal study of syllogisms and Leibniz is credited with beginning the creation of symbolic logic, it wasn’t until the mid-1800s that symbolic mathematical logic really came into existence [6]. There are many different ways to formalize logic and there are choices to be made in terms of such things as the quantifiers that are used (for example first order quantifiers versus higher order quantifiers). For mathematics to date, the most useful tool in this toolbox is known as first-order predicate logic, because of what it enables one to prove.

The formal developments connecting these two sides of the coin of mathematics finally set the stage for the formalization of infinitesimals. The missing key to the stage was Kurt Gödel’s Completeness theorem and its corollary the (logical) Compactness theorem that made the connection between the two sides of the coin [10]. Gödel is well known for his Incompleteness Theorem, but arguably his Completeness theorem is at least as profound, and in a constructive rather than a limiting way (the Incompleteness theorem is about the constraints to proof whereas the Completeness theorem is about the affordances of proof). The fundamental technique in the proof of the Completeness theorem (and this is more easily seen in subsequent simplifications using, for example, what is known as a *Henkin construction*), is to convert the referents for syntactic objects (the language of a theory) into the referents for semantic objects, and in a logically sound manner (T. Millar, personal communication, March 20, 2014).

Thus, to speak metaphorically, the consistency of infinitesimals is demonstrated (through a relatively easy application of the Compactness theorem) by starting with a logically sound way to syntactically describe infinitesimals and, almost literally, converting that description (in first order logic) into a semantic object (in set theory) in which those same infinitesimals “reside”. Both Archimedes and the Jesuits of the 17th century might have appreciated the humor in this development! But this development lay relatively dormant until the 1960s when Robinson realized what could be done with the stage that had been unlocked by Gödel’s key. It is important to note that this was nearly 2300 years after Archimedes had initially used infinitesimals. Finally, Robinson was able to prove infinitesimals were logically sound, and once his work was accepted as error-free, he wrote a textbook on infinitesimal calculus.

In 1966, *Non-standard Analysis* by Abraham Robinson was published. Although non-standard analysis had now been formally established, it was not easily accepted by the mathematical community. However, from this original text a few mathematicians, and in particular logicians, began work on teaching calculus using non-standard analysis. The mathematical logician Jerome Keisler for instance both helped prove the power of the approach theoretically and also wrote a very readable calculus textbook based on the use of infinitesimals.

Keisler tells that his book, *The Infinitesimal Approach to Calculus*, stirred up something of a hornet’s nest among mathematicians, with strong supporters and opponents (personal communication, April 17, 2014). This could be expected at first, because as mentioned before, in making calculus rigorous starting in the 1800s, the work of Weierstrass and others effectively eliminated the need for the use of infinitesimals. This elimination was seen as a very positive development by many prominent mathematicians, philosophers, and logicians as far as into the 20th century (for example Bertrand Russell famously said, “Infinitesimals . . . must be regarded as unnecessary, erroneous, and self-contradictory”). However, it is surprising that even though the infinitesimal approach is now universally accepted as mathematically sound, it is rarely taught.

5. Examples of Infinitesimal Use and Teaching Differences

In the following section I elaborate on infinitesimal calculus, and present examples of the principles of calculus and the different methods to proving them using a standard versus non-standard approach. First we note that a *positive infinitesimal* is a quantity greater than zero and less than any positive real number (with a similar definition for a *negative infinitesimal*). Together with the standard real numbers, the infinitesimals make up what is called the *hyperreals*. Just as the real numbers were a useful abstraction of the rationals, especially in areas such as continuous functions, the same is true of the hyperreal numbers as an extension of the reals. I will state here, without proof, some of the useful properties of the hyperreal numbers:

- (i). Every first order statement in the usual language of the real numbers is true about the reals numbers if and only if it is true about the hyperreal numbers.
- (ii). There are hyperreal numbers, called *positive infinitesimals*, which are larger than 0, but smaller than any positive real number.
- (iii). The additive inverse of a positive infinitesimal is called a *negative infinitesimal*.
- (iv). For any infinitesimal, its multiplicative inverse exists (by (i)), and in absolute value, is larger than any real number ((i) and (ii)).
- (v). For every finite hyperreal number there is exactly one real number such that the arithmetic difference of the two is an infinitesimal.
- (vi). Every function f on the real numbers “naturally” extends to a function f^* on the hyperreal numbers.

We will also use the following generalizations of very intuitively accessible notions of finite, very small, and very large:

- a) the sum of two infinitesimal numbers is infinitesimal;
- b) the inverse of an infinitesimal is an infinite hyperreal;
- c) the product of an infinitesimal with a finite hyperreal is infinitesimal.

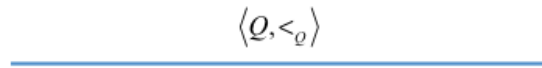
The direct construction of a useful model of the hyperreals using the language of arithmetic (in other words with a binary predicate symbol for “less than” and two binary function symbols for “addition” and “multiplication”) is quite complicated, though it can be done in a way that generalizes Cantor’s definition of real numbers (as equivalence classes of converging sequences of

rationals). However, it is relatively easy to visualize such a model if we restrict the language to just “less than”. Therefore we will define such a model and confirm that properties (i), (ii), (v) and (vi) hold in the model. Since we are not allowing binary function symbols for addition and multiplication in this language, the model will be “silent” on (iii) and (iv).

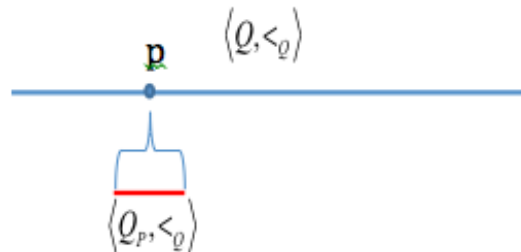
It is common practice in mathematics to define a new number system by taking ordered pairs of a “simpler” system and defining new relations and functions for the new system in terms of the relations and functions of the simpler system. For example, rational numbers are usually defined as equivalence classes of order pairs of integers (excluding 0 from the denominator), and the complex numbers are usually defined as ordered pairs of reals. Formally, the universe of our model will be all ordered pairs of rationals. We will call elements of our model “hybrid-rationals”. Before we give the formal definition of the less than relation on our hybrid-rationals, we will describe a direct way to construct them:

6. Construction of the hybrid-rationals

Start with the rationals:



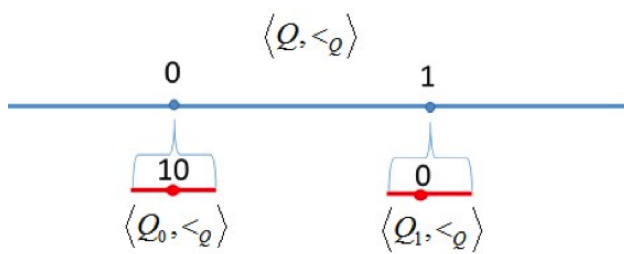
Now replace every rational p by a copy of the rationals “tagged” by the rational p :



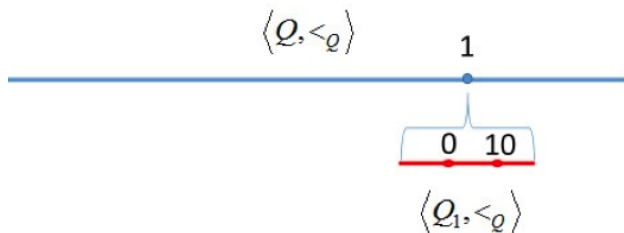
Formally the universe of the new model is the set $\{\langle p, q \rangle \mid p, q \in \mathcal{Q}\}$ and we define the order on the universe as the lexicographic order:

$\langle p, q \rangle <_{HQ} \langle s, t \rangle$ just if either 1) $p <_Q s$ or 2) $p = s$ and $q <_Q t$.

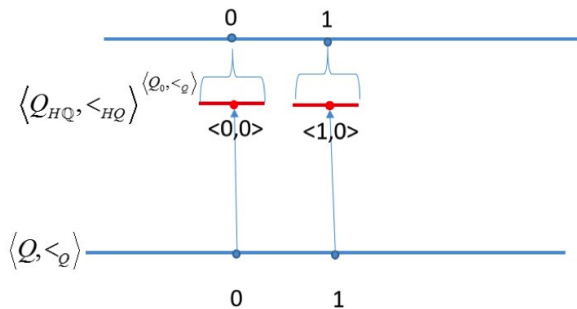
Here is an example of case 1): $\langle 0, 10 \rangle <_{HQ} \langle 1, 0 \rangle$:



And here is an example of case 2): $\langle 1, 0 \rangle <_{HQ} \langle 1, 10 \rangle$:



Now in a way similar to how we interpret the integers in the rationals or the rationals in the reals, we can interpret the rationals in the hybrid-rationals by interpreting the rational p as say $\langle p, 0 \rangle$. Here are two examples with $p = 0$ and $p = 1$:



It is straightforward to prove that this model is a dense linear order without endpoints. It also is countable. Therefore it is isomorphic to $\langle Q, <_Q \rangle$. And thus for every statement in the language, the statement is true in $\langle Q, <_Q \rangle$ if and only if it is true in $\langle HQ, <_{HQ} \rangle$. This is equivalent to (i) in the list above of properties of hyperreals.

Let us say arbitrarily that any element in $\langle HQ, <_{HQ} \rangle$ that is greater than $\langle 0, 0 \rangle$ is positive, and any element less than $\langle 0, 0 \rangle$ is negative. It follows that there are positive hybrid-rationals that are less than any positive rational — call those positive infinitesimals, and similarly for negative infinitesimals. This corresponds to (ii) in the above list.

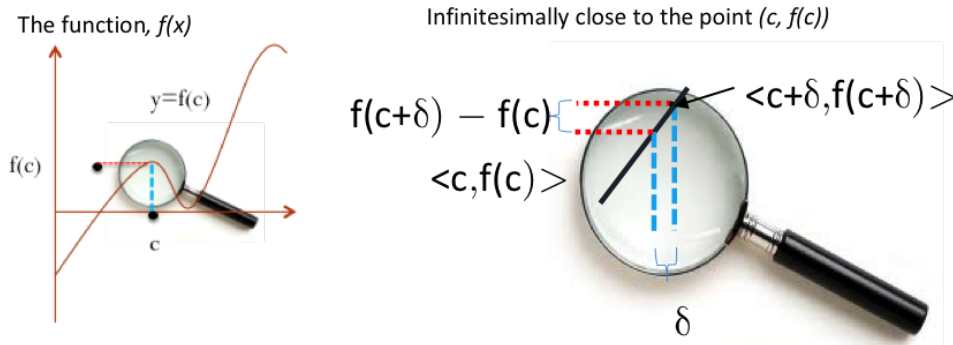
In general, say that two hybrid-rationals are infinitesimally close if they have the same first coordinate and different second coordinates. Then it also is easy to see that for any hybrid-rational, there is exactly one rational such that the two are either the same or infinitesimally close. This corresponds to (v) in the above list.

Finally, in this example without the arithmetic operations and simply treating set theory as the meta-theory, it is easy to show that (vi) holds. For example, for any function f (defined in the meta-theory) with domain and range the rationals, define

$$f^* (\langle p, q \rangle) = \langle f(p), f(q) \rangle.$$

7. Why might we want to do all this?

A primary advantage to the use of infinitesimals is that “nice” functions (for example differentiable) are “linear” at the infinitesimal scale in the neighborhood of any real number:



The first example I offer is the Chain Rule. The Chain Rule states that given two functions f and g where f is continuous on an interval $[a, b]$, $f'(x)$ exists at some point $x \in [a, b]$, g is defined on an interval I which contains the range of f , and g is differentiable at the point $f(x)$, if $h(t) = g(f(t))$ where $a \leq t \leq b$, then h is differentiable at x and $h'(x) = g'(f(x)) \cdot f'(x)$. [15, page 105].

Consider the following oversimplified example. Assume that the value of a particular home increases at a linear rate of 10% of its value per year. Next assume that the property tax (PT) on the homes in the district of this home is a base amount (BT) plus an amount of 5% of the value of the home. Given these assumptions, what is the annual rate of change of the property tax paid on this home? Let t be the time variable for number of years. Then let $V(t)$ be the values of the home, and $PT(V(t))$ the property tax as functions of time.

We then have:

$$\begin{aligned} V(t+1) &= V(t) + 0.1 \cdot V(t) \\ PT(V(t)) &= BT + 0.05 \cdot V(t) \end{aligned}$$

Therefore, with substitution and simplifying the expression we have:

$$\begin{aligned} PT(t+1) &= BT + 0.05 \cdot V(t+1) = BT + 0.05 \cdot (V(t) + 0.1 \cdot V(t)) \\ &= (BT + 0.05 \cdot V(t)) + 0.05(0.1 \cdot V(t)) \\ &= PT(t) + (0.05 \cdot 0.1) \cdot V(t) \end{aligned}$$

Thus, as is intuitively clear to many people without having to go through the formalism, the rate of change of the property tax on that home is the product of the rate of change of the property tax (as a function of home value) with the rate of change of the value of the home.

The reason this kind of example is relatively straightforward (no pun intended) is that the relations are linear. But what if the functions are not linear? In some sense, both a traditional proof and an infinitesimal proof of the Chain Rule in the general case exploit the fact that differentiable functions have reliable linear approximations. But in the traditional proof the limit arguments necessary to make the argument rigorous often obscure for the student the underlying linear intuition, whereas an infinitesimal proof is more likely to maintain the linear intuition for the student.

8. What might Newton have said?

His *Method of Fluxions* was written in 1671, but was not published until 1736. Newton considered a curve as generated by the continuous motion of a point. A changing quantity is called a **fluent** (a flowing quantity), and its rate of change is called the fluxion of the fluent. If a fluent, such as the ordinate of the point generating a curve, be represented by y , then the **fluxion** of this fluent is represented by \hat{y} . Newton also introduces another concept, which he calls the **moment** of a fluent; it is the infinitely small amount by which a fluent such as x increases in an infinitely small interval of time o . Thus the moment of the fluent x is given by the product $o\hat{x}$. Newton remarks that we may, in any problem, neglect all terms that are multiplied by the second or higher power of o and thus obtain an equation between the coordinates x and y of the generating point of a curve and their fluxions \hat{x} and \hat{y} . As an example, he considers the cubic curve $x^3 - ax^2 + axy - y^3 = 0$. [6, page 400]

Recall that Newton's approach to orbital analytics was geometric through and through, with the addition of fluents, fluxions, and infinitesimals. Therefore I will take the liberty of rephrasing part of the quote from Eves:

Newton also introduces another concept, which he calls the **moment** of a fluent; it is the infinitesimal amount by which a fluent such as x increases in an infinitesimal time o .

Or in symbols, changing x to f and t to x : for δ any infinitesimal and f as in the statement of the Chain Rule:

$$f(x + \delta) = f(x) + f'(x) \cdot \delta + \text{something to ignore.}$$

Now of course this begs the question of how to formalize the 'something to ignore', and limits are one way to do that. But if infinitesimals are made rigorous, then one can have what Keisler [8] calls the Increment Theorem:

The Increment Theorem: If x is real, $y = f(x)$, $f'(x)$ exists, and Δx is a nonzero infinitesimal, then $\Delta y = f'(x)\Delta x + \epsilon\Delta x$ for some infinitesimal ϵ , (i.e. $\Delta y \approx dy$ compared to Δx).

Proof: Since f is differentiable at x , it follows by definition that there is a real number r (and $f'(x)$ is then defined to be that real number, i.e. $f'(x) = r$) such that for any infinitesimal δ , r and $\frac{f(x+\delta)-f(x)}{\delta}$ differ by an infinitesimal. Calling that infinitesimal ϵ gives us

$$f(x + \delta) = f(x) + f'(x) \cdot \delta + \delta \cdot \epsilon$$

as desired.

This is what we will use to prove the Chain Rule in the case of using infinitesimals. However, we first consider the standard approach. Note that it comes at the problem in a somewhat similar manner.

9. Proving the Chain Rule

I reproduce Rudin's proof [15] of the Chain Rule so that the reader can compare and contrast an excellent standard proof with an infinitesimal proof. I use Rudin's proof because his books are known for their clarity and conciseness, and because it uses the notion of 'tends to zero'. First let us introduce the definition in that book of the derivative and an infinitesimal definition:

Definition from Rudin: Let f be defined (and real-valued) on $[a, b]$. For any $x \in [a, b]$ form the quotient $\phi(t) = \frac{f(t)-f(x)}{t-x}$ (for $a < t < b, t \neq x$) and define $f'(x) = \lim_{t \rightarrow x} \phi(t)$ provided the limit exists.

Infinitesimal definition: Let f be defined (and real-valued) on $[a, b]$. For any $x \in [a, b]$ if there is a real number r such that for all infinitesimals δ the difference of r and $\frac{f(x+\delta)-f(x)}{\delta}$ is 0 or an infinitesimal, then define $f'(x) = r$.

Note that both definitions involve ratios of 'rise over run' for points close to x . However the definition in Rudin defines an auxiliary function $\phi(t)$ (that is then used, in essence, to later introduce error terms that 'tend to zero') and the limit notion to define the derivative of f at x , whereas the infinitesimal definition 'goes directly there' and simply requires that there is a unique real infinitesimally close to the value of the ratio for any approximation (hyper-real) point infinitesimally close to x .

Standard Proof of the Chain Rule [15, page 105] Let $y = f(x)$. By the definition of the derivative, we have

$$\begin{aligned} I. \quad & f(t) - f(x) = (t - x) \cdot [f'(t) + u(t)] \\ II. \quad & g(s) - g(y) = (s - y) \cdot [g'(s) + v(s)] \end{aligned}$$

where $t \in [a, b]$, $s \in I$, $u(t) \rightarrow 0$ as $t \rightarrow x$ and $v(s) \rightarrow 0$ as $s \rightarrow y$. Let $s = f(t)$. Using first II. and then I., we obtain

$$\begin{aligned} h(t) - h(s) &= g(f(t)) - g(f(x)) = [f(t) - f(x)] \cdot [g'(y) + v(s)] \\ &= (t - x) \cdot [f'(x) + u(t)] \cdot [g'(y) + v(s)] \end{aligned}$$

or if $t \neq x$

$$III. \quad \frac{h(t) - h(x)}{(t - x)} = [f'(x) + u(t)] \cdot [g'(y) + v(s)]$$

Letting $t \rightarrow x$, we see that $y \rightarrow s$, by the continuity of f , so that the right side of III. tends to $g'(f(x)) \cdot f'(x)$, which gives the desired result.

To prove the Chain Rule this way one must use the limit definition of the derivative. The limit definition is somewhat counterintuitive because it requires students to visualize the denominator, h , going towards zero, and it also uses the 'error estimate' functions $u(t)$ and $v(s)$.

10. Proof of the Chain Rule with Infinitesimals

Now let us prove the Chain Rule using infinitesimals. We must show that there is a real number r such that for any infinitesimal δ , r and the hyperreal $\frac{g(f(x + \delta)) - g(f(x))}{\delta}$ differ by 0 or an infinitesimal, and that real number r will be equal to (by definition) $\frac{f(x + \delta) - f(x)}{\delta}$.

Start by fixing an arbitrary infinitesimal δ . Then by the Increment Theorem there is an infinitesimal ϵ such that $f(x + \delta) = f(x) + f'(x) \cdot \delta + d \cdot \epsilon$. Therefore $g(f(x + \delta)) = g(f(x) + f'(x) \cdot \delta + d \cdot \epsilon)$. Since $f'(x)$ is a real number, the expression in the brackets is zero or an infinitesimal. (Since an infinitesimal times a finite number results in 0 [if the finite number is 0] or an infinitesimal; the sum of an infinitesimal plus either 0 or another infinitesimal is an infinitesimal). Therefore by the Increment Theorem applied to g (since $g'(f(x))$ exists), there exists an infinitesimal λ such that $g(f(x + \delta)) = g(f(x)) + g'(f(x)) \cdot [f'(x) \cdot \delta + d\epsilon] + \lambda[f'(x) \cdot \delta + d\epsilon]$. Using substitution we have

$$\begin{aligned} &\frac{(g(f(x + \delta)) - g(f(x)))}{\delta} \\ &= \frac{(g(f(x)) + g'(f(x)) \cdot [f'(x) \cdot \delta + d\epsilon] + \lambda[f'(x) \cdot \delta + d\epsilon] - g(f(x)))}{\delta} \end{aligned}$$

$$\begin{aligned}
&= g'(f(x)) \cdot [f'(x) + \epsilon] + \lambda[f'(x) + \epsilon] \\
&= g'(f(x)) \cdot f'(x) + [g'(f(x)) \cdot \epsilon + \lambda \cdot f'(x) + \lambda\epsilon].
\end{aligned}$$

And once again the expression in the brackets is an infinitesimal, and therefore $\frac{g(f(x+\delta))-g(f(x))}{\delta}$ is infinitesimally close to $g'(f(x)) \cdot f'(x)$, and since δ was an arbitrary infinitesimal, the proof is complete.

Notice that there is no additional layer of abstraction here having to do with error functions or ‘tends to zero’ or ‘ $t \rightarrow x$ ’. With infinitesimals we already are there, and when there is a fixed real number for any such infinitesimal, then we have a proof.

Of course if one is not familiar with infinitesimals, then the layers of abstraction in the two proofs may appear about the same. However, conceptually and pedagogically the rules of understanding the relevant interactions (addition, multiplication, etc) of infinitesimals with reals is really quite straightforward. The key idea in ultimately disregarding the infinitesimal is that we are only interested in the behavior of functions on real numbers, but we use infinitesimals to discover properties of functions when restricted to real numbers.

Clearly the two methods give the same result. The second method removes the need to define new (error) functions, and is defined using a formula and concept students are very comfortable with. The idea that a number is so close to another number that they are virtually the same is similar to the idea of looking under a microscope and zooming in extremely close to one point on a curve. This concept is intuitive to students and helps simplify definitions and concepts such as slope [5].

One final point of difference between the two proofs. Note that the statement of the theorem in [15] includes an assumption of continuity of f on $[a, b]$, and that continuity is cited in the proof when Rudin writes: “we see that $y \rightarrow s$, by the continuity of f ”. This additional assumption is odd since it is not needed, and in fact on the page prior to the Chain Rule we find the following theorem.

Theorem. Let f be defined on $[a, b]$. If f is differentiable at a point $x \in [a, b]$, then f is continuous at x . (page 104)

Before writing the proof, note some of its implications. First, it implies that the assumption of continuity in the Chain Rule is superfluous. Recall that the assumption of continuity is not made in the Increment Theorem nor

is it used in the infinitesimal proof above. So why did Rudin include that assumption? It could have simply been an oversight, but is it perhaps because of the function $s(t) = f(t)$ that is introduced in the proof and then used in the statement: “we see that $y \rightarrow s$, by the continuity of f ”? (Observe that the statement of the theorem and the proof treat “ x ” as a fixed real, which is the reason for introducing $s(t) = f(t)$, since to make the limit argument the proof requires a function whose limit is 0 as $t \rightarrow x$). Finally, note that no such functions are needed in the infinitesimal case – it is simply a matter of verifying that $g'(f(x)) \cdot \epsilon + \lambda \cdot f'(x) + \lambda \epsilon$ is an infinitesimal under the assumptions of the theorem (excluding continuity), and assuming ϵ and λ are infinitesimals.

Now, let us look at Rudin’s proof of the above theorem, taken verbatim from [15, page 105].

Proof. As $t \rightarrow x$, we have, by Theorem 4.4²,

$$f(t) - f(x) = \frac{f(t) - f(x)}{t - x} g(t - x) \rightarrow f'(x)g(0) = 0$$

Next, consider the infinitesimal proof:

Proof. Fix any infinitesimal δ . It is sufficient to show that $f(x + \delta) - f(x)$ is an infinitesimal. By the assumption of differentiability at x , we know that $\frac{f(x + \delta) - f(x)}{\delta} - f'(x)$ is an infinitesimal, denote it ϵ . Therefore $f(x + \delta) - f(x) = \delta(x - t) + \delta\epsilon$ and since the term on the right hand side is an infinitesimal so is $f(x + \delta) - f(x)$.

²Rudin’s Theorem 4.4 is the following assertion:

Suppose $E \subset X$, a metric space, p is a limit point of E , f and g are complex functions on E , and

$$\lim_{x \rightarrow p} f(x) = A \quad \text{and} \quad \lim_{x \rightarrow p} g(x) = B.$$

Then

1. $\lim_{x \rightarrow p} (f + g)(x) = A + B$;
2. $\lim_{x \rightarrow p} (fg)(x) = AB$;
3. $\lim_{x \rightarrow p} (f/g)(x) = A/B$ if $B \neq 0$.

Again, for many readers the limit notation (which informally masks the formally necessary epsilon-delta details) obscures this argument in a way that the infinitesimal argument does not.

11. Defining Continuity

Next I look at another major component of calculus, which is the definition of continuity. Again, infinitesimal calculus uses a much nicer definition than standard calculus.

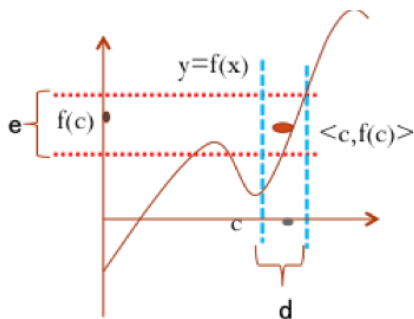
In a standard calculus class students learn the definition of continuity of a function $f(x)$ at the point $x = c$ through the first application of the definition of the limit definition. That is, “ $f(x)$ is continuous at the point $x = c$ ” is defined to mean that “ $f(c)$ exists and for all $\epsilon > 0$ there exists a $d > 0$ such that for all x , if $|x - c| < d$, then $|f(x) - f(c)| < \epsilon$ ”. This definition involves multiple quantifiers and conditions.

A logically simpler definition would be one using infinitesimals. Define $a \sim_{inf} b$ to mean that $|a - b|$ is zero or a positive infinitesimal. With the properties of the hyperreals, this is provably an equivalence relationship. Then the definition of $f(x)$ being continuous at $x = c$ can be stated as follows:

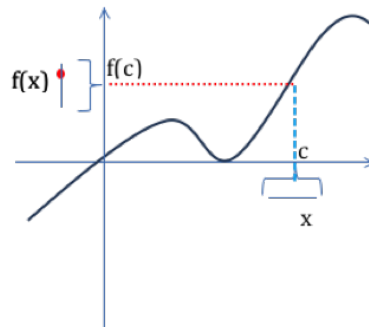
$f(c)$ exists and for all x (if $x \sim_{inf} c$ then $f(x) \sim_{inf} f(c)$).

And that certainly better corresponds to the standard verbal way of describing continuity: “If x is close to c then $f(x)$ is close to $f(c)$ ”.

To emphasize the multiple quantifiers in the standard definition and the more intuitive approach of the non-standard definition the pictures below illustrate the standard and non-standard definitions respectively.

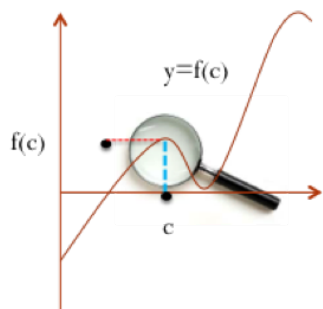
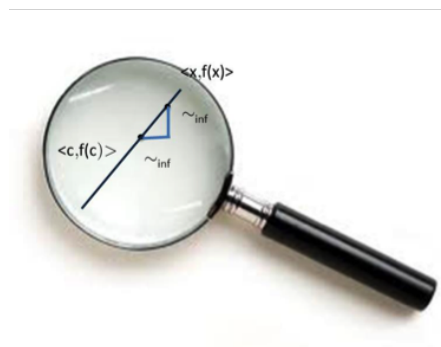


The standard definition of continuity.



The infinitesimal definition of continuity.

With the non-standard approach, one must zoom-in, as a microscope would do, in order to see that $x_0 + d$ is infinitesimally close to x_0 . This is intuitive, even for students who have already taken courses in standard analysis [18]. Even more surprising is that students who studied infinitesimal calculus actually better understood standard analysis. Tall [18] writes, “comparing students following Keisler’s approach with a control group, Sullivan demonstrated that those using infinitesimal techniques subsequently had a better appreciation of ϵ - δ techniques as well” (page 3). One reason is perhaps that the concept of two numbers being infinitesimally close is not completely removed from standard analysis. Once students develop the notion of being able to continuously zoom in on a specific point, thus getting infinitesimally close, they are more equipped to see the same process taking place within an epsilon-delta argument. The next two illustrations show how one would zoom in on a graph until seeing it as a continuous straight line.

The function $f(x)$.Infinitesimally close to the point $(c, f(c))$.

12. The Benefits of Infinitesimals

These are only two simple examples of how infinitesimals are used, but they give a clearer picture as to why infinitesimals may be more intuitive to students. In 2010 Robert Ely conducted a survey on students' intuitions about important concepts in calculus. His original study was a questionnaire he gave to 233 university calculus students to gauge their knowledge on limits, functions, continuity and the real number line. Ely did not conduct the survey expecting to find student conceptions about nonstandard analysis, thus when looking at the data he was surprised to find student conceptions of nonstandard analysis [5, page 126]. Not only did he realize students believed

in infinitesimals, but through the questionnaire as well as student interviews, he also found that students were consistent within their responses to questions about things being “infinitely small” or “infinitely close” (page 127). Eighty-three percent of students surveyed responded *true* to the statement “it is possible to choose 2 different points on the real number line that are infinitely close to one another” (page 127). In standard analysis two numbers are either the same number or a finite distance apart. Since all of the students surveyed have only taken standard analysis, it is interesting that they still believe two numbers can be infinitely close.

In one interview, one student Sarah, explains her thinking with the notion of zooming in and cutting something into pieces, similar to the concept mentioned earlier. Sarah gives the example of the number $3.999\dots$ repeating forever and 4 being infinitely close, but she also claims there is a number in between these two quantities. She claims that there are infinitely many digits followed by even more digits, which is the same concept used to think about the hyperreals and infinitesimals. What is even more telling is that Sarah states that $3.999\dots$ repeating forever is not actually a “Real” number [5, page 128].

Ely’s interviews also reveal that students are consistent with their beliefs. Of critical importance is the fact that Sarah, a sophomore in college, has taken two semesters of calculus; nonetheless, she explicitly states that she thinks her ideas are wrong [5]. Even with two semesters of calculus her natural intuitions have been unaltered, and her “misconceptions” have been unchanged by her formal mathematics training. It is safe to assume that she is learning mathematics through memorizing and following the instruction of her professors. What if instead she followed her own intuitions; might she have a better grasp on calculus?

When Ely compared student beliefs to the systems Leibniz and Robinson worked in, he found that they were very similar [5, page 139]. He concludes by asking, “Why do Sarah’s conceptions reflect those of mathematicians whose ideas she has not seen, and which are not part of the standard curriculum?” (page 142).

Thus there is reason to think that non-standard analysis is a better method for teaching first-year calculus. Although I argue here for such a course, few universities offer such a class (Millar, personal communication, March 20, 2014).

13. Institutional Difficulties with Teaching Infinitesimals

There are various reasons as to why colleges continue to teach standard calculus. First, universities need someone to initiate teaching a new calculus course. One university that had someone tackle this project was the University of Wisconsin Madison. Jerome Keisler started teaching a one-semester honors infinitesimal calculus course at University of Wisconsin Madison in 1969. He used his own mimeographed notes for the first year, and then taught using a two-semester preliminary version of what would eventually become his textbook (Keisler, personal communication, April 17, 2014). The first full 3-semester edition of his text was printed in 1976. The book was used at Wisconsin for approximately 20 years, and was even used for some large calculus sections (approximately 250 students) (Keisler, personal communication, April 17, 2014). During this time Keisler was able to recruit about 9 other members of the math department to teach the course. According to him, one reason he had trouble finding other teachers was that it is much easier for a teacher to use the approach they are familiar with than to do extra work to learn a new approach. In particular he noted that at the college level calculus is considered a ‘service course’ and thus a professor would rather spend her time working on her own research than preparing to teach calculus.

Although Keisler’s comment is about college professors, learning a new method of instruction would be an issue for high school teachers as well. It is safe to assume that instructors learn calculus from the standard approach, and once they have prepared materials and taught the course using one method, it is unlikely they will want to do the work to change.

Keisler (personal communication, April 17, 2014) believes that another reason the course did not catch on is that it requires learning something new. Most graduate math students do not take a logic course and therefore do not have a good understanding of infinitesimals. Keisler, Terry Millar, and Joel Robbin who all taught the course at Madison believe that many professors were uncomfortable with the material and therefore not confident in their ability to teach it (Keisler, Millar, and Robbin, personal communication). Again, high school and college calculus teachers would need time to learn the material and understand it in order to teach an effective course.

Finally, Keisler (personal communication, April 17, 2014) noted that math departments think it is dangerous to tamper with their calculus course. Calculus is a required course for many college graduates. Most mathematics departments currently have full control of these calculus classes, but if they began teaching calculus with the infinitesimal approach, how would the other departments react? What if other departments were then not satisfied and decided to teach calculus as they desired to their own students? Losing control of calculus could be detrimental to mathematics departments themselves. It could even lead to a large decrease in the number of positions within mathematics departments.

At Madison the infinitesimal calculus course lasted 20 years, but it was mainly taught as an experimental or honors course (Keisler, personal communication, April 17, 2014). The course was taught by volunteer teachers who, Keisler (personal communication, April 17, 2014) noted, “sometimes faced hostility from their colleagues”. Keisler saw his colleagues’ main incentive to volunteer to be learning the infinitesimal approach to calculus. Since the number of volunteers eventually ran out, the course was not sustainable.

Thus, although calculus can be taught using infinitesimals, it is perhaps understandable why so few college math departments implement this type of course. However, it is being taught elsewhere, especially in contexts where cost is a factor. Keisler’s text is free online and Dover (2012) sells a low-cost version of the text, so for homeschoolers, small and short courses, or where cost is a major factor the text is still being used (Keisler, personal communication, April 17, 2014).

14. More Thoughts on Teaching with Infinitesimals

What is interesting about the implementation of the course is that the reaction of the students was very favorable. Millar, Keisler, and Robbin, as mentioned above, are all former teachers of the course. They agree that one of the major advantages of the course is that one can often use “microscopes and telescopes” to illustrate definitions, concepts, and proofs. These visuals make it much easier for students to grasp the concepts. Scientists and engineers appreciate this as a method for discovering and understanding concepts (Millar, personal communication, March 20, 2014). As mentioned earlier, infinitesimals also eliminate the need for students to learn about limits before continuity.

One argument against teaching infinitesimal calculus is that more advanced analysis courses require knowledge of standard epsilon-delta proofs (J. Robbin, personal communication, March 27, 2014). Although this means students taking infinitesimal calculus and pursuing math degrees will have to learn the concepts later, it is important to note that infinitesimal calculus does not alter the major calculus concepts that students are learning.

Moreover, most students taking calculus will not go on to pursue a degree in mathematics. According to a recent study at Duke University, only 2% of their graduates each year major in mathematics [17]. However, 80% of Duke's first-year undergraduates take calculus. Also of importance is the fact that many mathematics majors come in with calculus credit, thus they are not included in the 80%. Therefore it is safe to assume over 80% of Duke's graduating students will have taken calculus, but the majority of those students will never take a more advanced analysis course. Based on this information, if infinitesimal calculus is more intuitive for students, then the majority of Duke's college graduates would have benefited from an alternative calculus course.

Stanley Ocken [12] at C.U.N.Y College reported that in the fall of 2000, 1,156,000 high school graduates began their freshman year at a four-year college in the United States. In the fall of 2000, 463,000 students in four-year U.S. institutions enrolled in first-semester calculus. This data shows 40% of college freshman took calculus. Ocken reports that many of those students were retaking calculus, as they had failed their first attempt at the course. This problem is not a new phenomenon. In 1988, Science Magazine reported that at some universities up to 50% of first-semester calculus students either drop out or fail [3]. At this point in time the NSF had allotted \$1 million to calculus reform, and many universities began experiments in order to receive grant money. However, even now in the 21st century this issue has not been resolved.

Jessica Ellis and her colleagues at San Diego State are also tackling the problem of first-year calculus courses. Although they are not advocating for an infinitesimal approach, they are interested in calculus reform. In 2010 they conducted a survey of over 5,300 students who were intended STEM majors, finding that 12.5% of those students switched majors and 31.4% of students who switched cited their negative experience in Calculus I as a factor in their decision [4]. If calculus is such a barrier to STEM fields it is important that the mathematical community do something to alleviate it.

15. Conclusion and Call to Arms

In conclusion, an infinitesimal approach to teaching calculus may not be initially easy, but it could have a large impact on how much students learn and engage in the course. A course that uses infinitesimals provides many opportunities to include the history of mathematics and to help change student attitudes that mathematics is static. There is also evidence that the course may be more intuitive for students, and thus easier. When considering that many high school students are pushed to take calculus in order to get into colleges, that many colleges require calculus even for non-STEM majors, and that calculus as it is currently taught on college campuses is often a serious roadblock, why would we not try to take a more intuitive approach in teaching it?

References

- [1] Bell, John L., *Continuity and Infinitesimals*, Stanford Encyclopedia of Philosophy. 2005 / 2013. <http://plato.stanford.edu/entries/continuity/>.
- [2] Bos, H. “Differentials, higher order differentials and the derivative in the Leibnizian Calculus,” *Archive for History of Exact Sciences*, Volume 14 (1974), pages 1–90.
- [3] Cipra, B. A., “Calculus Crisis Looms in Mathematics’ Future: Researchers and educators are debating how calculus should best be taught to increasingly recalcitrant students”, *Science*, Volume 239 Issue 4847 (1988), pages 1491–1492.
- [4] Ellis, J., Rasmussen, C., and Duncan, K., “Switcher and persister experiences in Calculus 1”, in *Sixteenth Annual Conference on Research in Undergraduate Mathematics Education*, Denver, February 2013. Available at http://pzacad.pitzer.edu/~dbachman/RUME_XVI_Linked_Schedule/rume16_submission_93.pdf.
- [5] Ely, R., “Nonstandard student conceptions about infinitesimals”, *Journal for Research in Mathematics Education*, Volume 41 Number 2 (March 2010), pages 117–146.

- [6] Eves, Howard, *An Introduction to the History of Mathematics*, 6th edition, Cengage Learning—Saunders College Publishing, Philadelphia, 1990.
- [7] Kasube, Herb, *Gauss' Cherry Tree: The Use of Anecdotes in the History of Mathematics*, lecture presented at the History and Pedagogy of Mathematics Spring Conference, Illinois State University, Normal, Illinois, April 12, 2014.
- [8] Keisler, H. Jerome, *Foundations of Infinitesimal Calculus*, 2011, online edition available at <https://www.math.wisc.edu/~keisler/foundations.html>.
- [9] McDonough, Jeffrey K., *Notes to Leibniz's Philosophy of Physics*, Stanford Encyclopedia of Philosophy. 2014. <http://plato.stanford.edu/entries/leibniz-physics/notes.html>.
- [10] Millar, Terry, *History of Mathematics: January 2014–April 2014*, Lecture notes, University of Wisconsin-Madison.
- [11] NPR staff, *Far From 'Infinitesimal': A Mathematical Paradox's Role In History*, 20 April 2014. <http://www.npr.org/2014/04/20/303716795/far-from-infinitesimal-a-mathematical-paradoxs-role-in-history>, accessed on January 29, 2016.
- [12] Ocken, Stanley, “Why students fail calculus,” *NYC Hold*, The City College of C.U.N.Y. 2001. <http://www.nychold.com/ocken-calculus01.pdf>, accessed on January 29, 2016.
- [13] Paulos, John Allen, “The 16th Century's Line of Fire: Infinitesimal, a Look at a 16th-Century Math Battle”, *New York Times*, April 7, 2014: D5. <http://www.nytimes.com/2014/04/08/science/infinitesimal-looks-at-an-historic-math-battle.html>, accessed on January 29, 2016.
- [14] Rorres, Chris, *Archimedes in the 21st Century*, New York University, 1995. <https://www.math.nyu.edu/~crrorres/Archimedes/Tomb/Cicero.html>, accessed on April 5, 2014.
- [15] Rudin, Walter, *Principles of Mathematical Analysis*, 3rd edition, International Series in Pure and Applied Mathematics, McGraw-Hill, 1964.

- [16] Russell, Bertrand, *Principles of Mathematics*, 2nd edition, W.W. Norton & Company, New York, 1938.
- [17] Smith, David A., “Renewal in Collegiate Mathematics Education”, *Documenta Mathematica*, Extra Volume ICM III (1998), pages 777–786. An expanded version available at <http://www.math.duke.edu/~das/essays/renewal/>.
- [18] Tall, David, “Intuitive infinitesimals in the calculus”, Poster presented at the Fourth International Congress on Mathematical Education, Berkeley, 1980. Available at <http://homepages.warwick.ac.uk/staff/David.Tall/pdfs/dot1980c-intuitive-infls.pdf>.
- [19] Tropp, Joel A., “Infinitesimals: History and Application,” Master’s Thesis, University of Texas at Austin, Austin, Texas. 1999.