The Mathematics Enthusiast

Volume 7
Number 2 *Numbers 2 & 3*

Article 10

7-2010

A semiotic reflection on the didactics of the Chain rule

Omar Hernandez Rodriguez

Jorge M. Lopez Fernandez

A Semiotic Reflection on the Didactics of the Chain Rule[1]

Omar Hernández Rodríguez[2] and Jorge M. López Fernández[3]
Graduate School of Education, University of Puerto Rico at Río Piedras

Abstract. *According to (Fried, 2008), there is an intrinsic tension in trying to apply the history of mathematics to its didactics. Besides the widespread feeling that the introduction of didactic elements taken from the history of mathematics can detract the pedagogy of mathematics from the attainment of important goals, (Fried, 2008, p. 193) describes a pair of specific pitfalls that can arise in implementing such historical applications in mathematics education. The description in (Fried, 2008), is presented in the parlance of Sausserian Semiotics and identifies two semiotic "deformations" that arise when one fails to observe that the pairing between signs and meanings in a given synchronic "cross-section" associated with the development of mathematics need not hold for another synchronic cross section at a different time. In this exposition, an example related to an application of the history of the chain rule to the didactics of calculus is presented. Our example illustrates the semiotic deformations alluded by (Fried, 2008), and points out a possible explanation of how this may lead to unrealistic pedagogical expectations for student performance. Finally, an argument is presented for the creation of a framework for a historical heuristics for mathematics education, possibly beyond the bounds of semiotics.*

Keywords: chain rule; composition of functions; differentiation; historical heuristics; history of analysis; history of mathematics; Sausserian semiotics

1. Application of Sausserian Semiotics to Mathematics

The problem of applying the history of mathematics to mathematics education and its relation to semiotics is broadly discussed by (Fried, 2008). Specific examples are given of distortions that arise in the application of semiotics to the history of mathematics when failure to distinguish differences between synchronic and diachronic descriptions of the body of mathematics

[1] The authors are especially grateful to Robert L. Knighten, Adjunct Professor of Mathematics at Portland State University, for his help in clearing up some of the historical points treated in this article and related to the development of the chain rule as a statement of differentiability of the composition of two functions.
[2] *E-mail address: omar.hernandez4@upr.edu*
[3] *E-mail address: lopez.jorgem@gmail.com*

occurs. We shall not dwell on the elements of semiotics discussed there; in fact, the presentation in (Fried, 2008) depicts adequately the elements of semiotics relevant to the reading of this article, and also presents some of the details of the development of Sausserian Semiotics and its adequacy for framing the problem of applying the history of mathematics to mathematics education. Also, (Fried, 2008) dwells in a general fashion on some of the main ideas of semiotics, the contributions of Peirce to this field of knowledge, and also presents some examples related to the distortions that can arise in failing to differentiate between synchronic descriptions of the relations between signs and meanings (both in linguistics and mathematics) that occur at different time frames. In this article, we employ the framework put forth by Fried (2008) to discuss the application of the history of mathematics to the teaching of calculus, specifically to the didactics of the chain rule for the differentiation of the composition of two differentiable functions. It will be argued that failure to make the alluded distinction between diachronically distinct synchronic descriptions of the body of mathematics can result in unrealistic expectations regarding student understanding of the chain rule.

For the purposes of facilitating the exposition that follows, we review some of the descriptions presented by Fried (2008, p. 193) regarding two distortions that can arise when one fails to recognize the fact that the relations between signs and meanings in the history of mathematics, can be vastly different when diachronic differences between time periods are taken into account. The distortions, according to (Fried, 2008, pp. 193) are twofold. The first distortion consists of supposing that the synchronic relations between signs and meanings in a given historic period coincide with those thought to be the corresponding relations between the homologous signs and meanings of the present time (when mathematics education occurs). This distortion constitutes, in fact, the worst error a historian can make, that is, the error of anachronism. The error involves contriving non existent or false synchronic relationships between signs and meanings in the given historic period. The second distortion described by Fried, and related to the failure to recognize diachronic differences, is the fabrication of false inferences regarding the evolution of signs, meanings and their pairings throughout the history of mathematics. In (Fried, 2008) two examples of these distortions are given, one in linguistics, due to Saussure (1974) himself, and another one related to the notion of function in Euler's times (Fried, 2008, p. 194).

In this note we discuss some of the signs and meanings associated with the notions of derivative and composition of functions, as related to the chain rule[4]. The issue here is the history of the chain rule since the publication of L'Hospital *Analyse des Infiniment Petits pour la Intelligence des Lignes Courbes* in 1696. Succinctly stated, the modern statement of the chain rule is taken to be one that relates the derivative of the composition of two functions with the individual derivatives of the functions composed (provided, of course, certain conditions are satisfied). Since the idea of composition of functions seems to have appeared in the literature at least a century after the publication of *Analyse des infiniment petits*[5], it is impossible that the signs and meanings relevant to the statement of the chain rule in the seventeenth century are the same as those associated with the present version of the chain rule.

2. History of the Chain Rule

We now present a brief relation of the evolution of the mathematical ideas and relations that have come to be known as the "chain rule". The present day statement of the Chain Rule is a rather sophisticated one and presupposes the confluence, and consolidation of many mathematical ideas. In fact the modern statement of the Chain Rule is the following:

---

[4] that is, to the rule for differentiating the composition of two differentiable functions.

[5] As far as we can tell, the first "modern" version of the chain rule appears in Lagrange's 1797 *Théorie des fonctions analytiques*, (Lagrange, J. L., 1797, §31, pp. 29); it also appears in Cauchy's 1823 *Résumé des Leçons données a L'École Royale Polytechnique sur Le Calcul Infinitesimal,* (Cauchy, A. L., 1899, Troisième Leçon, pp. 25).

Theorem 1.

If g is differentiable at c, and f is differentiable at g(c), then $f \circ g$ *is differentiable at c and*

$$(f \circ g)'(c) = (f' \circ g)(c) \cdot g'(c).$$

The possibility of the succinct and beautiful statement contained in Theorem 1 presupposes a great deal of evolution of the underlying mathematical ideas and a commensurable amount of "negotiations" related to the corresponding signs and meanings. Here, the functions of the statement are assumed to be defined on neighborhoods of their points of differentiability, and the corresponding limits for the difference quotients are supposed to exist as real numbers. Furthermore, once the "correct" definition of the derivative for Euclidean spaces was discovered, the chain rule was extended to state a relation about the differentiation of composite functions on Euclidean spaces, thus changing very little the formal statement of Theorem 1; see (Dieudonné, 1960).

In (L'Hospital, 1696, p. 2), the difference of a variable *y* depending on an independent variable *x* is defined as the infinitesimal increment in *y* when *x* changes by an infinitesimal amount $dx \neq 0$. In modern notation that needs little explanation: $dy = [y(x+dx) - y(x)]dx$. In fact, in Analysis des infiniment petits (L'Hospital, 1696), curves are considered as polygons of an infinite number of sides of infinitesimally small lengths, so that if we were to extend the "side" of the curve $y(x)$ that joins the points $(x, y(x))$ and $(x+dx, y(x+dx))$ in the graph of *y* as a function of *x*, we would, in fact, obtain the tangent line to the curve at (x, y), whose slope is, no more and no less, than the quotient *dy/dx*. It should be noted that L'Hospital (1696) used a geometric argument that employs the similarity of infinitesimal triangles to show that the value of the desired slope is infinitely close to the indicated value *dy/dx*. If one follows mathematical convention and writes $y'(x)$ for the quotient *dy/dx* (and this is, indeed, a quotient!) then, the relation $dy = y'(x)dx$ holds for all infinitesimals *dx*. In fact, in *Analyse des infiniment petits* the calculus of derivatives is really the calculus of "differences" of variables.

It may come as a surprise to the reader that nowhere in *Analyse des infiniment petits*, (L'Hospital, 1696), is the chain rule stated explicitly. This mystery is rather significant in more than one way. First, if we have differentiable variables *y* depending on *u* and *u*, in turn, depending on *x*, then $dy = y'(u)du$ and $du = u'(x)dx$ are the basic relations for the differences at the appropriate points, so that $dy = y'(u)du = y'(u)u'(x)dx$. From this, again, it is clear that $dy = y'(u)u'(x)dx$, and this is the chain rule.

Furthermore, this is true whether *dx* is zero or not. It may be even more surprising to realize that the statement of the chain rule is also absent in all of Euler's analysis books, *Introductio in analysin infinitorum*, (Euler, 1748, Vol. 1), (Euler, 1748, Vol. 2), and *Institutiones calculi differentialis*, (Euler, 1755). Furthermore, Euler did define the notion of a function in (Euler, L., 1748, Vol. 1), but he never treated the topic of the composition of functions in any of his writings, (Euler, 1748, Vol. 1), (Euler, 1748, Vol. 2) and (Euler, 1755).

As far as we can tell, the first mention of the Chain Rule[6] in the literature of calculus seems to be due to Leibniz (Child, 2007, p. 126), and it appears in a 1676 memoir (with various mistakes) in which he calculated $d\sqrt{a+bz+cz^2}$ by means of the substitution $\mathrm{x} = \mathrm{a} + \mathrm{bz} + \mathrm{cz}^2$. In *Analyse des infiniment petits* (L'Hospital, 1696, pp. 3-4), the rules for calculating the differences of the basic algebraic combination of (differentiable) variables are given. L'Hospital posed the problem of calculating the difference of x^r for any "perfect or imperfect" power r (that is, for any rational power *r*) and he answers his question by proving that $dx^r = rx^{r-1}dx$. In keeping with the style of *Analyse des infiniment petits*, after proving general rules, L'Hospital gave instances of

[6] as a rule for finding differences of expressions by means of substitutions.

the application of the rule to specific examples. In this case, the first example of the general rule $dx^r = rx^{r-1}dx$ given by L'Hospital is the calculation of the difference $d(ay - x^3)^3$. The calculation is, as expected, a direct application of the chain rule and requires no expansion of the cube. No comment is made by L'Hospital to the effect that the application of the general rule (differentiation of the cube) to more complicated expressions necessitates an application of a special rule (the chain rule) whose statement or demonstration is nowhere to be found in his work (Campistrous, Lopez, and Rizo, 2009).

In our view, the example provided illustrates dramatically that the anachronisms that ensue from failing to understand the diachronic differences between the mathematics of different times can betray the existence of pitfalls and ill practices in the didactics of mathematics. Informal experiments performed in an introductory non standard calculus course at the University of Puerto Rico have shown that students have significant difficulties in identifying the composed functions before they are able to correctly apply the chain rule. On the other hand, the level of understanding of the chain rule improves when the algorithm is presented as differentiation after a substitution of variables. To the distant observer this may seem to be a trivial difference, but the history of mathematics shows that the notion of composition somehow requires a higher level of abstraction for its understanding. Similar remarks apply to the understanding of the chain rule by students when it is presented in nonstandard analysis parlance as contrasted with the usual standard analysis presentation, which requires arguments often seemed as much to do about nothing. Perhaps, it should be remarked as a sobering thought, that even if all the diachronic and synchronic semiotic deformations in the history of mathematics can be avoided, there will still remain what in our view is the most interesting part of the history of mathematics (and, also, the part most related to mathematics education), and that is the inferences that can be made from it regarding optimal strategies for classroom teaching.

3. Towards a Historical Heuristics for Mathematics Education

After (Toeplitz, 1963), it has been amply regarded that the so called "genetic approach" to mathematics education has special advantages. Toeplitz (1963) carefully points out the difference between history in general, as a compilation of facts, and the history of mathematics in particular, as a source of ideas for teaching mathematics. He remarks: "It is not history for its own sake in which I am interested, but the genesis, at its cardinal points, of problems, facts and proofs" (Toeplitz, 1963, p. xi). In view of the semiotic considerations of this work, we venture to suggest the need of a sort of historic heuristics for mathematics education, in the vein, perhaps, of the heuristics of (Polya G., 1945) for problem solving, but which attend to the pairings the human mind makes between signs and meanings for the purpose of advancing mathematics knowledge. In our opinion, in the case of the chain rule, a strong argument can be made for the cognitive advantages of defining the derivative as a difference arising from an infinitesimal change, just like in (L'Hospital, 1696). To this, in our view, we owe the absence of explanations and the familiar and informal handling of the chain rule in (L'Hospital, 1696).

Kitcher (1983, p. 229) presents some compelling arguments for what we consider to be the cognitive advantages of what can be called "Newton's kinematic metaphors" (thinking of fluents and fluxions as positions and velocities, respectively, of moving objects; see (Kitcher, 1983, p. 232)), and the appropriateness of infinitesimals as a cognitive vehicle for the "initial calculus[7]", striving to describe mathematically the idea of "change". In fact, in spite of all objections to the unclarities of the initial calculus, as voiced (mainly) by Bishop George Berkeley (Wilkins, 2002), the calculus was rapidly accepted and it developed in an unprecedented way. In the words of

[7] the calculus developed by Newton Leibniz, the Bernoullis, L'Hospital and Euler

(Kitcher, 1983, p. 230): “To understand how the power of the methods introduced by Newton and Leibniz outweighed the unclarities which attended them, we must begin with the problems which interested the mathematicians of the early seventeenth century.”; and further ahead on the same page: “Both Newton and Leibniz introduced new language, new reasonings, new statements and new questions into mathematics. Some of the new expressions were not well understood and the workings of some of the new reasonings were highly obscure. Despite of these defects, the changes they proposed were accepted quite quickly by the mathematical community, and the acceptance was eminently reasonable.” Thus, in spite of all logical difficulties, the methods of the initial calculus were intensely exploited and they yielded a dramatic development of mathematics in general and the calculus in particular. From reading (Kitcher, 1983) one cannot escape the feeling that the meanings associated to the idea of an infinitesimal were quite adequate to capture the underlying idea of “change” and to transform it into the body of knowledge we know today as calculus. In fact, there are good reasons to believe that the Leibnizian “signs” for the calculus and the formulation of change in terms of infinitesimals were responsible for the faster development of European continental mathematics when compared to its counterpart in England; see (Grabiner, 1997). In our view these arguments strongly suggest a clear cognitive advantage in “thinking change” in terms of infinitesimals and, also, explain why the chain rule in the language of infinitesimals is obvious to the point of not requiring explicit justification.

Kitcher (1983, p. 154) argues the contention that mathematical knowledge is cumulative when compared with scientific knowledge, which is regarded as being transformed in a more disruptive fashion. For instance, in resolving that the Lorenz transformations (which render invariant Maxwell’s laws of electromagnetism) are the basic transformations of physics, strictly speaking, it is necessary to admit the incorrectness of Newton's laws, which remain invariant under the Galilean transformations. This is of course consistent with (Kuhn, 1970). In mathematics, on the other hand, in dealing with the famous error of (Cauchy, 1882) regarding the continuity of the limit of a series of continuous functions, analysis suffered a very profound transformation which brought about the ε - δ definition of limits and the notion we know today as uniform converge. But, as opposed to physics, in mathematics, the previous body of knowledge of the calculus can be reformulated in terms of limits, and all of the "theorems" of the initial calculus continue to be valid in the new version of mathematical analysis that ensued from Cauchy, Weierstrass and others. Hence, in this sense, mathematical knowledge is cumulative. However, any teacher of calculus can attest to the fact of the great amount of difficulty that the Cauchy-Weierstrass theory of limits presents to students. This is perhaps to be expected as it took roughly a century from the time the initial calculus was invented to the formulation of the theory of limits to deal with Cauchy's "error". It thus seems reasonable to suggest that when paradigms change in mathematics (as the change towards the theory of limits after the infinitesimal approach) they must have a cognitive advantage for dealing with pressing unsolved problems, but this advantage does not necessarily extend for mathematics education. In fact, teaching the old body of mathematical knowledge with the new paradigms, in our view, adds a heavy overhead to the pedagogy of the subject matter.

Reflections related to observed advantages in student understanding when the calculus is presented in the language of infinitesimals appear in education journals; studies on this very topic, using Keisler's book *Foundations of Infinetisimal Calculus* (Keisler, 1976) as a textbook, have been made, and the observed results appear discussed in the literature (see, for example, Sullivan (1976)).

In (Kitcher, 1983, p. 155), the following qualified remark is found: "Unfortunately, the history of mathematics is underdeveloped, even by comparison with the history of science". Clearly the topics for the explorations suggested by this brief exposition need a framework for the history of mathematics that lies beyond the bounds of semiotics, and these explorations are crucial for gaining a better understanding of the cognitive workings of the human mind as it strives to understand mathematics.

It would be, indeed, a framework that must include the discussion of pairings of signs and meanings validated by the history of mathematics as being effective, but it must also include the discussion of issues like the ones raised here. This framework, a sort of historical heuristics for mathematics education, should set the stage for exploring the cognitive workings of the human mind as it grapples with signs and meanings in its quest for advancing mathematical knowledge.

References

Campistrous, L. A., Lopez, J. M., and Rizo, C. (2009). *Reflexiones sobre la didáctica del cálculo a propósito de una lectura del primer texto publicado sobre esta materia por el Marqués Guillaume François Antoine L'Hospital: Analyse des infiniment petits pour l'intelligence des lignes courbes*, Submitted for publication.

Cauchy, A. L. (1882). *Cours d'Analyse*, Série II, Tome III, Œvres Complétes, 25 vols, Paris, 1882-1932.

Cauchy, A. L. (1899). *Résumé des Leçons Données a L' Ecole Royale Polytechnique sur Le Calcul Infinitesimal*, Série II, Tome IV, Œvres Complétes.

Child, J. M. (translator), Leibniz, G.W. (Author). (2007). *The Early Mathematical Manuscripts of Leibniz*. Merchant Books.

Dieudonné, J. (1960). *Foundations of Modern Analysis*. NewYork: Academic Press,

Euler, L. (1748). *Introductio in analysin infinitorum*, Vol 1, Edición Facsimilar, SAEM "Thales". Sevilla: Real Sociedad Matemática Española.

Euler, L. (1748). *Introductio in analysin infinitorum*, Vol 2, Introduction to the Analysis of the Infinite (translation), Book II. New York: Springer Verlag.

Euler, L. (1755). Institutiones *calculi differentialis*, Foundations of Differential Calculus (translation), New York: Springer Verlag.

Fried, M. N. (2008). History of mathematics in mathematics education: A Saussurean perspective. *The Montana Mathematics Enthusiast*, *5*(2&3), 185-198.

Grabiner, J. (1997). Was Newton´s calculus a dead end? The continental influence of Maclaurin´s Treatise of Fluxions. *American Mathematical Monthly*, *104*(5), 393-410.

Keisler, H. J. (1976). *Foundations of Infinitesimal Calculus*. Creative Commons Attribution-Noncommercial-Share Alike, 2.5 Licence, http://creatvecommons.org/licenses/by-nc-sa/2.5/.

Kitcher, P. S. (1983). *The nature of mathematical knowledge.* London: Oxford University Press.

Kuhn,T. S. (1970). *The structure of scientific revolutions*. Chicago: University Press.

Lagrange, J. L. (1797). *Théorie des fonctions analytiques*. Paris : L'Imprimerie de la république.

L'Hospital, G. F. A. (1696). *Analyse des infiniment petits, Pour l'intelligence des lignes courbes*. Paris: L'Imprimerie Royale.

Polya, G. (1945). *How to Solve It: A New Aspect of Mathematical Method*. Princeton: Princeton University Press.

Saussure, F. (1974). *Cours de Linguistique Générale*. C. Bally and Albert Sechehaye (eds.) with the collaboration of A. Riedlinger, Paris: Payot.

Sullivan, K. (1976). The teaching of elementary calculus using the nonstandard analysis approach, *The American Mathematical Monthly*, *83*(5), 370-375.

Toeplitz, O. (1963). *The calculus: A genetic approach*. Chicago: The Mathematical Association of America and The University of Chicago Press.

Wilkins, D. R. (Ed.). (2002). *The analyst; or a discourse addressed to an infidel mathematician*, Printed by and for S. Fuller at the Globe in Meath-street, and J. Leathly Bookseller in Dames-street. 1734, http://www.maths.tcd.ie/pub/HistMath/People/Berkeley/Analyst/Analyst.pdf.