The Mathematics Enthusiast

Volume 7
Number 2 *Numbers 2 & 3*

Article 11

7-2010

Some reflections on Hernández and López’s reflections on the chain rule

Michael N. Fried

-COMMENTARY-

Some reflections on Hernández and López's reflections on the chain rule

Michael N. Fried[1]
Ben Gurion University of the Negev, Israel

To shed light on how history of mathematics can play a role in mathematics education, Hernández and López chose well in focusing on the chain rule as a case study. Generally speaking, it is, as they point out, a topic students find difficult to grasp: it is a rule involving a different order of complexity than, say, the rule for differentiating functions of the form $f(x)=x^n$, and, although it can eventually become intuitive, it is far from intuitive at the start. Harel *et al.* (2009) have pointed out that students' difficulties with the chain rule involve: "...coordinating three or more changing quantities [when viewed in a physical context]; executing the chain rule algebraically requires sophisticated abilities in recognizing algebraic form; making the translation between the physical and algebraic manifestations..." Others, reasonably enough, have seen the difficulties of the chain rule in conjunction with more general problems concerning the composition of functions (e.g. Cottrill, 1999). For their part, Hernández and López have looked for insight on the chain rule and how to confront the difficulties associated with it in one of the earliest instances where the rule seems to have been used, namely, in l'Hôpital's *Analyse des infiniment petits* (1696/1716).

L'Hôpital's point of departure was the notion that a curve could be thought of as an infinite number of infinitely small straight line segments, a polygon with an infinite number of sides. From that, it was an easy step for l'Hôpital to take the fundamental operation in the analysis of curves to be what he called the "difference," "the infinitely small portion by which a variable quantity increases or diminishes continually" (l'Hôpital, 1696/1716, p.2). For Hernández and López, accordingly, l'Hôpital's approach gives a historical justification for "...the cognitive advantages of defining the derivative as a difference arising from an infinitesimal change..." (p.6), that is, the cognitive advantages of a non-standard analysis framework, such as the one they used with success at the University of Puerto Rico. The connection is not absurd. On the one hand, Abraham Robinson, the founder of modern non-standard analysis himself recognized Leibniz as spiritual father of the subject, indeed, that his ideas vindicated Leibniz's (see Robinson, 1966, p.2). Robinson also mentions l'Hôpital's 1696 work, and l'Hôpital, himself, admits his own debt to Leibniz. So there is a certain undeniable confluence of

[1] mfried@bgu.ac.il

thought here. At the same time, one needs to take great care in searching the past for ways of thought offering us "cognitive advantages."

The danger is the temptation to make historical mathematical texts and historical mathematical thought work for us in ways that obscure their own idiosyncratic historical character. As I have pointed out elsewhere (Fried, 2008) and Hernández and López describe in their paper, semiotics can bring into relief the way the latter can occur, namely, that the synchronic structure of our sign system, in this case, how we employ and combine signs to speak and think about mathematics, presses and forces past mathematical thought, past synchronies into a modern mold. This also shows why the temptation—and, therefore, the danger—is so great and, sometimes, almost irresistible. It is so difficult to view the past on its own terms and not to see it as serving the present, a past that Michael Oakeshott has referred to as the "practical past" (see Oakeshott, 1999). Yet it is struggling against this temptation that historians of mathematics must forever engage in. Semiotics helps in this regard by making one aware of how meaning is constituted in the presence or absence of a sign or in how a sign is used with respect to other signs—and one should keep in mind that "function," "rule," "infinitely small," are as much signs as *f(x)*, $\frac{dy}{dx}$, *p*, and ε ; in fact, even continuous texts can function as signs (Tobin, 1989). Hernández and López speak in their paper, in a fine turn of phrase, about finding a "historical heuristic for mathematics education" (pp.5,8). There is much to think about in formulating such a historical heuristic, but what we learn from semiotics is that this heuristic must at very least make it a guiding principle that how signs are used now is different than how they were used in the past and that the sign systems of the past were as integral and coherent then as ours are now. This is perfectly consistent with how historians see their own pursuit. As the historian Herbert Butterfield puts it: "...it is better to assume unlikeness at first and let any likenesses that subsequently appear take their proper proportions in their proper context; just as in understanding an American it is wrong to assume first that he is like an Englishman and then quarrel with him for his unlikenesses, but much better to start with him as a foreigner and so see his very similarities with ourselves in a different light" (Butterfield, 1951, p.38).

Hernández and López have certainly heeded this principle in highlighting l'Hôpital's reliance on infinitely small differences. They also pay attention to the crucial fact that the chain rule is never stated explicitly by l'Hôpital. The other crucial fact about l'Hôpital's *Analyse des infiniment petits*, which I think Hernández and López are aware of but do not mention, is that while l'Hôpital uses the word "variable" everywhere, he uses the word "function" nowhere. The two facts are, I believe, deeply related and give a clue as to the great differences between l'Hôpital's conception of the calculus and ours. These differences, in my view, ought to be stressed somewhat more than the apparent resemblance between l'Hôpital's infinitely small differences and Robinson's non-standard analysis, which, despite Robinson's own claims alluded to above remains moot. Let me, then, say a few words about these differences.

The complete title of l'Hôpital's work is *Analyse des infiniment petits pour l'intélligence des lignes courbes* (*Analysis of the infinitely small for the understanding of curved lines)*. Although in the published work itself, the title becomes abbreviated to *Analyse des infiniment petit* and is subtitled *Du calcul des differences* (*On the calculus of differences*), it is the second part of the main title that reveals l'Hôpital's ultimate intention, namely, the understanding of curved lines. And I use the word "intention" here almost in the Husserlian sense: curved lines provide the direction of l'Hôpital's thought, and the understanding of curved lines and lines related to them informs the meanings of all his other terms. Thus the definition of a "variable" begins the book since a curve is given by an equation that relates, algebraically, the lengths of certain lines. Such lengths, or other cogent geometrical objects related to curves, are the variables l'Hôpital really has in mind even though in principle an algebraic equation can be a purely symbolic expression. If there is any doubt about this, just consider that immediately after defining "variables" and "constants" he writes, "Thus in a parabola the ordinates (*les appliquées* [*par ordre*]) and the lines cut off [from the diameter by the ordinate] are the variable quantities and instead of the parameter [the *latus rectum* of the parabola] is the constant quantity" (p.1). Immediately afterwards, moreover, when l'Hôpital defines the "difference," the examples given are again related to a curve and are not only lengths, but also triangles as differences of segments of the curve and rectangles as differences of areas contained by the curve, diameter, and ordinates.

Curves and their equations, then, are always present in l'Hôpital's view of his subject, sometimes implicitly, usually explicitly; his notion of infinitely small differences is suggested by his understanding of curves and, in turn, is meant to be a tool in furthering his understanding of curves. Functions, as I have said, are nowhere in l'Hôpital's work. In Leibniz, whom l'Hôpital read and admired, the word "function" does appear, but, there too, it is only in relation to curves: functions are relations connected to tangents, normals, subtangents, subnormals, and so on. Adding the sign "function" in mathematics, which occurred in the following century, meant that the kind of calculus described by l'Hôpital could be thought of as being about the relations connected with a single abstract object, the function, rather than the almost endless variety of individual curves. In fact, if one sees the idea of an infinitesimal difference coming out of a conception of curves, one can understand why an approach to the calculus by way of functions is set against one inspired by infinitesimals. This can be seen strikingly in the title of Lagrange's treatise on the calculus written exactly one century after l'Hôpital's and partly a response to it (see the introduction to Lagrange, 1797): *Théorie des fonctions analytiques, contenant les principles du calcul différential, dégagés de toute consideration d'infiniment petits ou d'évanouissans, de limits ou de fluxions, et réduits a l'analyse algébrique des quantités finies* (*Theory of Analytic Functions: Containing the Principles of the Differential Calculus Free of All Consideration of the Infinitely Small or Evanescents, Limits, or Fluxions, and Reduced to the Algebraic Analysis of Finite Quantities*) (1797).

Lagrange's work presents the chain rule as a rule (Lagrange, 1797, p. 30), as Hernández and López point out. Lagrange can do this precisely because he has shifted the object from curve to function. What this shift allows, in effect, is an entire range of second-order concepts such as a "function of a function" and a "derivative of a function" (and it is worth reminding that in this same work, Lagrange invents the term "derivative" to mean the function *derived* from the "primitive function" from the series $f(x+i)=f(x)+Pi+Qi^2+...$). One cannot express the chain rule as a rule without such second order ideas. It is a subtle point. l'Hôpital can speak of the kinds of first order relations among variables that produce equations of curves, relations such as x^2, x^3, xy, and for *these* he can provide rules; however, he cannot give rules for relations of relations, for it is curves and not relations themselves that that form the domain of his intentionality. So, it is only in examples of the rule for "perfect and imperfect" powers, that is, for integral and fractional powers, that we discern our chain rule in l'Hôpital's work. This rule, Rule IV, is the rule stating that the difference of x^m is $mx^{m-1}dx$, and the examples are as follows: the difference of $(ay - xx)^3$ (the one quoted by Hernández and López), the difference $\sqrt{xy + yy}$, the difference of $\sqrt{ax + xx + \sqrt{a^4 + axyy}}$, and finally, $\frac{\sqrt{ax+xx}}{\sqrt{xy+yy}}$. One might argue that the chain rule is truly implicit in Rule IV since one can consider the difference of u^m which is $mu^{m-1}du$ and then consider the difference du in terms of the difference dx. Of course we desperately want to say, "where u is a function of x," but this, precisely, is what we cannot say.

The difficulty of these examples in comparison to the simple powers referred to in Rule IV itself give them more the air of problems based on the rule rather than simple examples of it. And in his solutions (even though these might be in fact more to Bernoulli's credit than l'Hôpital's) one senses that l'Hôpital is showing off—a kind of one-upmanship not at all untypical in 17th century mathematics. But this too brings us further away from the chain rule as a rule. If we do something that helps us solve a problem, a clever manipulation or representation, we generally do not see ourselves as following a rule. In fact, to the extent we say we are following a rule, we lessen our claim to be solving a problem. For a problem is a problem precisely when there are no means known in advance of solving the problem. We prove our problem-solving mettle by pulling something out of a hat.

To return to mathematics education, what is it really that we learn from l'Hôpital and the chain rule? It is not, I think, a different approach to the chain rule, even with his concentration on the infinitesimals with its seeming identity with the modern alternative of non-standard analysis. It is rather a different intentionality altogether with regards to basic objects of mathematical analysis. By studying l'Hôpital with an eye to the ways he differs from us, students and teachers can begin to see how the introduction of "function" can give rise to a "rule" as a rule. They begin to see how much our own way of talking about analysis—and therefore our understanding of analysis—is shaped by these differences. This kind of insight into our own understanding of things derived from looking at an understanding different than ours is in general what one wants learn from history and, no doubt, its clearest educational value. Let me then end by quoting Butterfield again writing about what we gain from history, or rather, what we lose when our history fails to distinguish past from present:

> If we turn our present into an absolute to which all other generations are merely relative, we are in any case losing the truer vision of ourselves which history is able to give; we fail to realise those things in which we too are merely relative, and we lose a chance of discovering where, in the stream of the centuries, we ourselves, and our ideas and prejudices, stand. In other words we fail to see how we ourselves are, in our turn, not quite autonomous or unconditioned, but a part of the great historical process; not pioneers merely, but also passengers in the movement of things (Butterfield, 1951, p.63)

References

Butterfield, H. (1951). *The Whig Interpretation of History*. New York: Charles Scribner's Sons.

Cottrill, J. F. (1999). *Students' understanding of the concept of chain rule in first year calculus and the relation to their understanding of composition of functions.* Ph.D. Dissertation, Purdue University.

Fried, Michael N., (2008). History of Mathematics in Mathematics Education: A Saussurean Perspective. *The Montana Mathematics Enthusiast, 5*(2), 185-198.

Harel, G., Holdener, J., McCallum, W., McClure, J., Thompson, P. (2009). The Chain Rule. Presentation at the Institute for Mathematics and Education, *Workshop on Mapping the Calculus Curriculum*. Available at the website: http://math.arizona.edu/~ime/2008-09/Calculus/Group6_0419.pdf (accessed, 15 April, 2010).

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

L'Hospital, G. (1696/1716). *Analyse des infinitement petits pour l'intélligence des lignes courbes,* 2nd edition. Paris: François Montalant.

Oakeshott, M. (1999). *On History and Other Essays*. Indianapolis, Indiana: Liberty Fund, Inc.

Robinson, A. (1966). *Non-Standard Analysis*. Princeton: Princeton University Press.

Tobin Y. (Ed), (1989). *From Sign to Text: A Semiotic View of Communication*. Amsterdam: John Benjamins Publishing Company