In Defense of Differential Calculations

ALAIN SCHREMMER
Department of Mathematics, Community College of Philadelphia, Philadelphia, PA 19130
FRANCESCA SCHREMMER
Department of Mathematical Sciences, West Chester State University, West Chester, PA 18383

0. Introduction. That "fat calculus texts", and therefore calculus courses, are topics based is probably the single most important cause for the "cookbook" approach and it is difficult to see how the "*fragmentation problem* [...] *can be resolved by adding an additional semester to the calculus course*" [**2**]. For instance, why should the definition of continuity, a property of *f*, be followed with the definition of the derivative, a function, rather than with the definition of differentiability, a stronger property of *f*? But then, the currently prevailing emphasis in Differential Calculus on the derivative at the expense of differentiability does not result from mathematical or pedagogical considerations and, as a result, does little for an harmonious development: Why should Differential Calculus be based on the notion of derivative in dimension 1 when it is based on that of differentiability in all other dimensions? Typically, students are warned that "[while] *for functions of one variables, the terms "differentiable" and "has a derivative" are synonymous,* [...] *for functions of two variables differentiability is a more stringent requirement than the existence of partial derivatives*" [**1**] but little explanation is offered and the subsequent treatment of the total derivative is generally disappointing.

To quote Dieudonné [**9**] at some length, "*the fundamental idea of Calculus* [is] *the 'local' approximation of functions by* linear *functions. In the classical teaching of Calculus, this idea is immediately obscured by the accidental fact that, on a one-dimensional vector space, there is a one-to-one correspondence between linear forms and numbers, and therefore the derivative at a point is defined as a* number *instead of a* linear form. *This slavish subservience to the shiboleth of numerical interpretations at any cost becomes much worse when dealing with functions of several variables: one thus arrives, for instance, at the classical formula giving the partial derivatives of a composite function, which has lost any trace of intuitive meaning, whereas the natural statement of the theorem is of course that the (total) derivative of a composite function is the composite of their derivatives, a very sensible formulation when one thinks in terms of linear approximations*."

Indeed, "*the theme of approximation, it* [is] *agreed, is central to calculus—what it is and what it does*" [**23**] and I. Bivins has by now received *two* prizes for his article "What a Tangent Line is When it isn't a Limit" [**4**]. The committee's citation for the Polya prize reads in part: "*By defining the tangent line as the best linear approximation to the graph of a function near a point,* [Bivins] *has narrowed the gap, always treacherous to students, between an intuitive idea and a rigorous definition. The subject of this article is* fundamental to the first two years of college mathematics *and should simplify things for students....*" (Emphasis added).

But, if the idea of linear approximation is mentioned in just about every text, it is always as an afterthought and the "intuitiveness" of the tangent-as-limit-of-a-secant remains unquestioned. For instance, in an article advocating Carathéodory's definition of the derivative [**14**], the linear approximation definition is merely mentioned as a "*variation*" to be found for instance in [**5**], [**21**] even though"[t]*his approach has the intended additional benefit of making transparent the linear approximation of the tangent line*".

The reason may be that, to be really useful, the idea of approximation should be extended to the use, algebraically, of best polynomial approximations and, geometrically, of osculating curves of degree *n*.

1. Asymptotic expansions. For real valued functions to represent the way situations change, the *differential* calculus, the "mathematics of change", must derive *local information*—about (mostly) gradual[1] changes—from punctual information. The desired information can be *qualitative*—is f near x_0 positive/negative, increasing/decreas-ing, concave up/concave down?— or *quantitative*—what is the approximate value, rate of change, concavity of f near x_0? But, qualitatively, we might also want to know whether, at x_0, f is continuous or differentiable while, quantitatively, we might ask what the jump or the slope is.

To study functions by way of their local polynomial approximations is then considerably more natural than, to quote Lagrange, "*seeing derivatives in isolation*", if only because it is quite reminiscent of the use of decimal approximations in arithmetic. And the advantages are that: **i.** it organizes, unifies and simplifies the differential calculus [[**13**], **ii.** it extends, naturally, to the Frechet derivative in multivariable calculus, to Banach Spaces, and to jets in Differential Topology and, last but not least, **iii.** not only are the "naïve" proofs in this setting natural and plausible, but they are easily made rigorous.

---

[1] A theory of *abrupt* changes is provided by Catastrophe Theory.

A function f can be as simple as a polynomial or as complicated as a fractal. In arithmetic, we might approximate as complicated a number as π by a simple decimal expansion such as 3.14, writing $\pi = 3.14 + (...)$ where the ellipsis represents the complicated but small difference between π and 3.14. Thus, a natural idea when studying $f(x_0+h)$, the value of f near a point x_0, is to try to separate a *principal part*, that is a part smooth enough to be relevant to the information being sought, from a *remainder*, a part too small to be significant in that regard.

The simplest approach is to distinguish $P^{(n)}(x_0,h)$, a polynomial part of degree n in $h = x - x_0$, from a remainder $R^{(n)}(x_0,h)$ small enough that, compared to $P^{(n)}(x_0,h)$ and for the given purpose, it can be neglected. By this, we mean that is $R^{(n)}(x_0,h) = \mathrm{o}[h^n]$ which we read as saying that $R^{(n)}(h)$ approaches 0 faster than h^n, that is $\lim_{h\varnothing 0} \dfrac{R^{(n)}(h)}{h^n} = 0$; graphically, this means that the graph of $|R^{(n)}(h)|$ is under the graph of $|h|^n$ in a neighborhood of 0. Also, since h is small, it can be thought of as 0.1 so that h^2, h^3, … can be thought of as 0.01, 0.001, … .

This use of approximation is not only quantitative (numerical) but also qualitative as an approximation should be determining in the sense that the approximating function should be equivalent in some specified manner to the function being approximated, say, for instance, up to a diffeomorphism. When using polynomials as approximating functions, they are called the jets of the function [**6**].

In other words, we are using *asymptotic* expansions (as opposed to *series* expansions as in [**18**]) with the power functions as *gauge* functions [**10**]. For beginning students it is enough to point out that the principal part carries the relevant quantitative information and that the remainder carries only the qualitative information that $P^{(n)}(h)$ differs from $f(x_0+h)$ by a small amount and just write $f(x_0+h) = P^{(n)}(h) + (...)$.

2. An example. Since "[*g*]*raphing rational functions is one of the more difficult tasks in calculus*" [**19**] let us consider, for instance, how the function $f(x) = \dfrac{x}{x^3-1}$ can be studied by way of approximations.

We study f near its pole: $f(1+h) = \dfrac{1+h}{3h + \mathrm{o}[h]} = \dfrac{1/3}{h} + \mathrm{o}[h]$ by division in *ascending* powers and, at ∞, we get, by division in *descending* powers, $f(x) = \dfrac{1}{x^2} + (\ldots)$. This gives us the graph of f *outside* the region represented by the dotted square in (Fig 1), its largest part in spite of appearances.

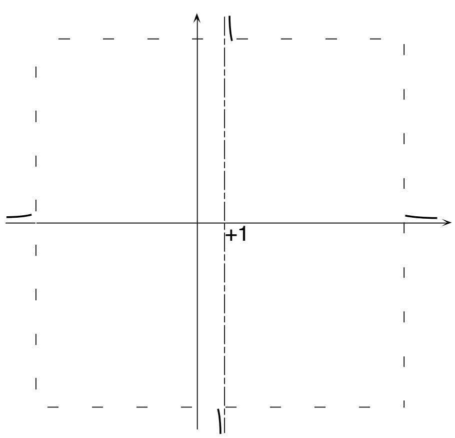

Figure 1.

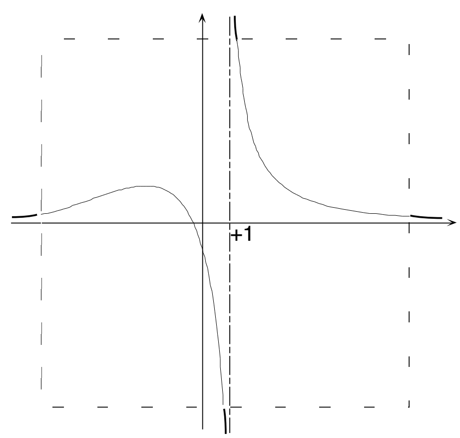

Figure 2.

The proximate graph in Fig. 2. is then obtained by joining the local graphs smoothly. While there could be "fluctuations", this is the graph that would be seen from far away and it shows the existence of at least one maximum, at least one inflection and at least one zero.

To validate this, we must show that the function f is smooth everywhere except at the pole. Near a point x_0 other than the pole 1, say 2, $f(2+h) = \dfrac{2+h}{(2+h)^3-1} = \dfrac{2+\mathrm{o}[1]}{7+\mathrm{o}[1]} = \dfrac{2}{7} + \dfrac{\mathrm{o}[1]}{7+\mathrm{o}[1]} = \dfrac{2}{7} + \mathrm{o}[1]$ by division in *ascending* powers. That the remainder is o[1] shows that BCA $f(2+h) = \dfrac{2}{7} = f(2)$ is the best constant approximation of $f(2+h)$ and that f does not jump at 2.

Similarly, $f(2+h) = \dfrac{2+h}{(2+h)^3-1} = \dfrac{2\ +\ h}{7\ +12h\ +\ o[h]} = \dfrac{2}{7} - \dfrac{17}{49}\ h\ +\ o[h]$ by division in *ascending* powers. That the remainder is $o[h]$ shows that BAA $f(2+h) = \dfrac{2}{7} - \dfrac{17}{49}\ h$ is the best affine approximation of $f(2+h)$ and that f has a tangent at 2: $T_2 f(x) = \dfrac{2}{7} - \dfrac{17}{49}(x-2)$. This is precisely how the tangent plane is defined in higher dimensions [**12**].

Moreover, the coefficient of h, $-\dfrac{17}{49}$, is the instant rate of change of f at 2 since $-\dfrac{17}{49} = \dfrac{f(2+h) - f(2) - o[1]}{h}$ and thus $-\dfrac{17}{49} = \lim_{h\emptyset 0} \dfrac{f(2+h) - f(2)}{h}$. In other words, the rate of change of f is that of its BAA.

3. Elementary Differential Calculus. Qualitatively, the sign of the coefficients of h^0, h^1 and h^2 in the principal part of $f(x_0+h)$ give, respectively, the sign, the variation and the concavity of f near x_0. More generally, we define a function to be continuous iff it is locally approximately constant[2] and to be differentiable iff it is locally approximately affine. Thus, calculus notions, continuity and differentiability, appear here as systematic generalizations of precalculus notions, constancy and affinity.

Quantitatively, the coefficients give the value of f and of the successive rates of change: We define f', the derivative of f, to be the function whose value at x_0 is the coefficient of h in $f(x_0+h)$. More generally, for a function that is n-differentiable (locally approximately polynomial of degree n), we define the n^{th} derivative of f as the function $f^{(n)}(x)$ whose value at x_0 is $n!$ times the coefficient of h^n in $f(x_0+h)$ (Peano derivatives [**7**]). Alternatively, we could define f' as before and then, recursively, $f^{(n+1)} = [f^{(n)}]'$. Even though the existence of an osculating polynomial of degree n does not quite insure the existence of any recursive derivative of order >1 [**8**], for all practical purposes, the two notions are equivalent.

Most of the usual theorems of the differential calculus, e.g. the derivation rules, the inverse function theorem, can be proved much more easily because, systematically, the proofs involve computing $f(x_0+h)$ and looking at the coefficient of h. The "derivative tests" become trivial and L'Hôpital's rule superfluous. And even if defining continuity at x_0 by the local existence of a best constant approximation does not help to prove that a continuous function on a closed bounded interval is bounded, it points out very clearly where the difficulty lies. Because f is continuous on an interval, say $[a,b]$, $\forall x_0 \in [a,b]$, $f(x_0+h) = f(x_0) + o[1]$. Suppose h is in a neighborhood of 0, whose size depends on x_0, such that $o(1) < \dfrac{1}{10}$ for example. If we knew that we could cover $[a,b]$ using finitely many of these intervals, say N, then $\left|f(x) - f(a)\right|$ would be bounded by $\dfrac{N}{10}$ and the theorem would be proved. This then raises the question as to whether, from any open covering of a closed bounded interval, we can extract a finite one. The Mean Value Theorem is seen here as a remainder theorem, that is as providing bounds on the error made when we approximate $f(x_0+h)$ by $f(x_0)$. Finally, following Picard [**20**], the Fundamental Theorem is proven by solving by finite differences the following initial value problem: given a function $f(x)$, find the value at x_1 of a function $F(x)$ such that $F(x_0) = y_0$ and $F'(x) = f(x)$. Stressing the "antiderivative aspect" is more important, *at this point* and for most calculus students, than stressing the "measure theoretic aspect" which, in practice, devolves into teaching integration techniques.

In the case of algebraic (resp. transcendental) functions defined as solutions of functional (resp. differential) equations, the method of undetermined coefficients gives, near the initial point, an *approximate polynomial solution* whose properties announce those of the exact solution. After we obtain an (approximate) addition formula, we can obtain approximate polynomial solutions near other points but this involves passing from the local to the global. This gives a good approximate local study of transcendental functions except near ∞. Once again, the behaviour of these functions at ∞ is new and makes it necessary to include some of them as gauges. For instance, after we observe that, near ∞, e^x cannot be approximated by any power function, we include it as a gauge to study further functions. Such a study of the approximate solutions serves as an excellent introduction to the study, assuming their existence [**22**], of the exact solutions carried out as in, for instance, [**17**], [**11**]; the approximate solutions can then be shown to be the Taylor approximations of the exact solutions.

Of particular interest for students intending to pursue a career in sciences other than mathematics, physics or engineering, is the fact that the basic attitudes developed by the above treatment are precisely those needed in the study of Dynamical Systems, the first place where real applications become feasible. Until Poincaré, the emphasis was ex-

---

[2] If nothing else, this makes it immediately plausible why a function continuous over a closed interval should be Riemann integrable.

clusively on finding solutions in closed form and/or series solutions but the solutions are usually in so complicated a form as to necessitate difficult qualitative methods of investigation to describe their behaviour. In contrast, a direct study from the equation and the phase portrait is often quite feasible [**3**] and can therefore follow immediately the differential calculus sketched above.

4. Conclusion. From Newton and Leibniz to Robinson, the desire has always been to extend calculations beyond algebra but, in the conventional calculus, we do not calculate anymore. Lagrange wanted to free calculus from "*any consideration of infinitesimals, vanishing quantities, limits and fluxions and reduce it to the algebraic study of finite quantities*" [**15**], [**16**]. While he did not entirely succeed as the algebra of little ohs rests on the notion of limit at 0, this algebra—explicit or implicit—provides a treatment of elementary calculus that is incomparably transparent.

REFERENCES.

1. H. Anton. "Calculus with analytic geometry." 1988 John Wiley & Sons. New York.
2. H. Anton. In Defense of the Fat Calculus Text. *UME Trends*. **2**(6) (1991) 1.
3. M. Artigue and V. Gautheron. "Systèmes Differentiels, Etude Graphique." 1983 CEDIC/F. Nathan. Paris.
4. I. C. Bivins. What a Tangent Line is When it isn't a Limit. *The College Mathematics Journal*. **17**(2) (1986) 133-143.
5. W. Boyce and R. DiPrima. "Calculus." 1988 Wiley. New York.
6. T. Bröcker and L. Lander. "Differentiable germs and catastrophes." London Mathematical Society Lecture Notes. 1975 Cambridge University Press. London.
7. A. M. Bruckner and J. J. Leonard. "Derivatives." Papers in Analysis. Mathematical Association of America. Washington D. C.
8. A. Denjoy. Sur l'intégration des coefficients différentiels d'ordre supérieur. *Fundamenta Mathematicae*. **25** (1916) 273-326.
9. J. Dieudonné. "Foundations of Modern Analysis." Pure and Applied Mathematics. Eilenberg ed. 1960 Academic Press. New York.
10. J. Dieudonné. "Infinitesimal Calculus." 1968/1971 Hermann/Houghton Mifflin Company. Boston.
11. R. L. Finney and D. R. Ostbey. "Elementary Differential Equations with Linear Algebra." 1984 Addison Wesley. Reading.
12. F. J. Flanigan and J. L. Kazdan. "Calculus Two: Linear and Non-Linear Functions." 1971 Prentice-Hall. Englewood Cliffs.
13. A. Gleason. The geometric content of advanced calculus. *CUPM Geometry Conference*. (1967). Durst Ed. Mathematical Association of America. 2. 53-119.
14. S. Kuhn. The Derivative à la Carathéodory. *The American Mathematical Monthly*. **98**(1) (1991) 40-44.
15. J. L. Lagrange. "Théorie des fonctions analytiques." Oeuvres de Lagrange. Serret ed. 1881 Gauthier-Villars. Paris.
16. J. L. Lagrange. "Leçons sur le calcul des fonctions." Oeuvres de Lagrange. Serret ed. 1884 Gauthier-Villars. Paris.
17. S. Lang. "Analysis I." 1976 Addison Wesley. Reading.
18. H. Levi. "Polynomials, Power Series and Calculus." 1968 Van Nostrand. Princeton.
19. R. Lopez and J. Mathews. Using a Computer Algebra System to Solve for Maxima and Minima. *College Mathematics Journal*. **21**(5) (1990) 410-414.
20. E. Picard. "Traité d'Analyse." Cours de la Faculté des Sciences de Paris. 1901 Gauthier–Villars. Paris.
21. M. H. Protter and C. B. Morrey. "A First Course in Real Analysis." 1977 Springer-Verlag. New York.
22. P. Renz. "Steps toward a rethinking of the foundations and purposes of introductory calculus." Toward a lean and lively calculus. Douglas ed. 1986 The Mathematical Association of America. Washington, D.C.
23. P. Zorn. Content and Discontent in San Antonio. *UME Trends*. **3**(1) (1991) 1.