Differentials : an introduction with hyperreal numbers and infinite microscopes

Henry, Valérie

Published in:
Primus

Publication date:
2010

Document Version
Early version, also known as pre-print

An introduction to differentials based on hyperreal numbers and infinite microscopes

by Valerie Henry
HEC-Management School of the University of Liege
7, bld du Rectorat, bât B31
4000 Liege
Belgium
Tel: 0032-4-366 27 36
email : V.Henry@ulg.ac.be

Abstract In this paper, we propose to introduce the differential of a function through a non classical way, lying on hyperreals and infinite microscopes. This approach is based on the developments of nonstandard analysis, wants to be more intuitive than the classical one and tries to emphasize the functional and geometric aspects of the differential. In the second part of the work, we analyse the results of an experiment made with undergraduate students who had been taught calculus by a non standard way for nearly two years.

Keywords: differential, hyperreal numbers, infinite microscope, nonstandard analysis

1 Introduction

Differentials are interesting from a didactic point of view because they are fundamental in calculus not only for pure mathematicians but also for applied users like engineers, physicists, economists or even biologists.

It is well known that undergraduate students have difficulties in understanding differentials [3] and their applications. First, differentials are used in two quite different situations: on the one hand, they give local information by approximating non-linear functions by linear ones; on the other hand, they produce global results by solving differential equations. Secondly, mathematicians and applied users consider differentials as distinct objects: they are functions for the former and they represent "small quantities" for the latter. Moreover, studies have shown that students in physics have difficulties in grasping this aspect of linear approximation for differentials ([2],[5]).

In this paper, we propose a new presentation of differentials, working in the context of non-standard analysis with hyperreal numbers and infinite microscopes. We conjectured that this method could help students to imagine differentials more easily as linear functions which allow a local approximation for non-linear functions. We tested our conjecture on a sample of undergraduate students in economics who had had a first course about univariate analysis by a nonstandard method and a three months course in multi-variate calculus by the same method.

2 Hyperreal numbers

Since Leibniz, a fundamental idea in calculus is to study "infinitesimal" changes of functions. However $\mathbb{R}$, being an Archimedean ordered field, has no positive real number which is "infinitesimal", i.e. greater than zero but smaller than every positive real number. It is therefore necessary to work in a numerical system which contains the real numbers, but which is also larger than $\mathbb{R}$ in order to contain the "infinitesimal numbers" considered by Leibniz. Such a system was rigorously defined by Robinson [8]. During the 1960s, this American mathematician developed the non-standard analysis which restored the Leibniz's infinitesimal numbers left aside by the mathematicians of the 19th century.

Like Robinson, we work with hyperreal numbers but use the teaching theory developed by Keisler [7]. Indeed, experimental research has shown that this presentation yielded good results for the students under consideration [6], [9].

To make our text self-contained, we briefly recall Keisler's definitions and notations. The hyperreal numbers are defined by the following axioms [7]:

- **Algebraic axioms.** The set of hyperreal numbers is a field and $\mathbb{R}$ is one of its subfields.

- **Order axioms.** The set of the hyperreal numbers is totally ordered, and its order extends the natural order on $\mathbb{R}$.

- **Infinitesimal axiom.** There exists a positive hyperreal number.

- **Standard part axiom.** Every finite hyperreal number x (i.e. every hyperreal which lies between two real numbers) is infinitely close to exactly one real number which is called the *standard part* of x and denoted by $st(x)$; in this case, we write $x \approx \mathrm{st}(x)$.

- **Function axiom.** Every real function f of one or more variables has a *natural extension*: it is a corresponding hyperreal function f^* of the same number of variables, which has f as restriction and for which the sum, difference, product and reciprocal functions are given by the algebraic axioms.

- **Solution axiom.** If two systems of formulas (equations or inequations) have exactly the same real solutions, then they have exactly the same hyperreal solutions.

Moreover, we denote by H a positive infinite, i.e. a hyperreal number which is greater than every positive real number.

3 Infinite microscopes

To study local properties of an observed object, it is convenient to "magnify" it. Here, we consider a function f defined on a subset D in the Euclidean space $\mathbb{R}^n$, for a given arbitrary integer n and a point $P = (r_1, \dots, r_n)$ in the interior of D. We denote by $\mathcal{G}_f$ (resp. $\mathcal{G}_{f^*}$) the graph of f (resp. of f^*, the natural extension of f), by $\bar{P}$ the point of $\mathcal{G}_f$ whose coordinates are $(r_1, \dots, r_n, r_{n+1})$, where $r_{n+1} = f(r_1, \dots, r_n)$.

We shall magnify by a number H, which is an arbitrary positive infinite hyperreal number, the graph $\mathcal{G}_{f^*}$ of the natural extension of f in a neighbourhood of the point $\bar{P}$ (in the space $\mathbb{R}^{n+1}$). Hence, we use a "microscope" of infinite power H, applied to the point $\bar{P}$; this is a well-known way to proceed in non-standard analysis [1], [6], [10]. In practice, in the equation which defines the function f,

$$x_{n+1} = f(x_1, \dots, x_n),$$

we replace the coordinates $x_1, \dots, x_n, x_{n+1}$ respectively by $X_1, \dots, X_n, X_{n+1}$ such that, for $i = 1, \dots, n+1$,

$$X_i = H(x_i - r_i) \Longleftrightarrow x_i = r_i + \frac{X_i}{H}.$$

Obviously, we only take into account the points $(X_1, \dots, X_n, X_{n+1})$ of $\mathcal{G}_{f^*}$ for which each coordinate X_j is finite; so, the "old" coordinate $x_j = r_j + \frac{X_j}{H}$ is a hyperreal number which is infinitely close to r_j. Moreover, through the microscope we only see the *standard part* of the graph $\mathcal{G}_{f^*}$, i.e. the set of points whose coordinates are the standard part of the corresponding coordinates.

As an elementary example, consider the function f of two variables defined by

$$(x_1, x_2) \mapsto x_3 = f(x_1, x_2) = x_1^2 - x_2^2. \tag{1}$$

When we apply a microscope of infinite power H at the point $\bar{P} = (2, 1, 3)$, equality (1) becomes

$$3 + \frac{X_3}{H} = \left(2 + \frac{X_1}{H}\right)^2 - \left(1 + \frac{X_2}{H}\right)^2. \tag{2}$$

We easily obtain

$$X_3 = 4X_1 - 2X_2 + \frac{X_1^2}{H} - \frac{X_2^2}{H}. \tag{3}$$

If X_1, X_2, X_3 are finite, we can take the standard part in (3). Since $\frac{X_i}{H}$ is an infinitesimal, we find

$$\mathrm{st}(X_3) = 4\ \mathrm{st}(X_1) - 2\ \mathrm{st}(X_2). \tag{4}$$

So, in the eyepiece of this infinite microscope, we see the plane, through the origin of $\mathbb{R}^3$, whose equation can be written

$$z = 4x - 2y. \tag{5}$$

This plane can be displayed graphically. Indeed, if a mathematical software like *Mathematica* is used to plot the graph of f in 3D, then we obtain a curved surface when the two variables x and y vary in intervals of length 10 centered on 2 and 1, respectively (see Figure 1). However, when the length of the intervals becomes 0.01, the software works like a microscope of power 1000 and the graph appears to be a plane whose equation is given in (5)(see Figure 2).

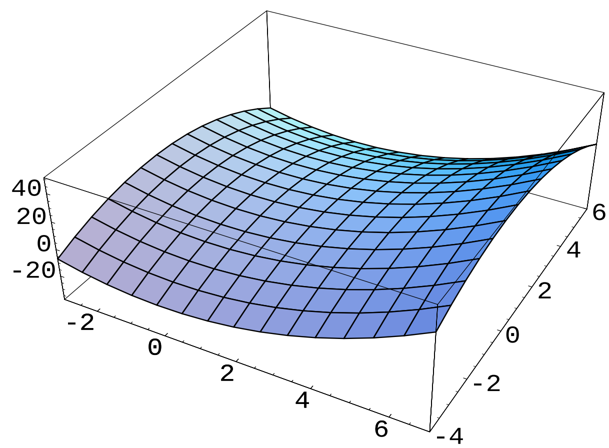

Figure 1: Surface defined by $z = x^2 - y^2$ for $-3 < x < 7$ and $-4 < y < 6$

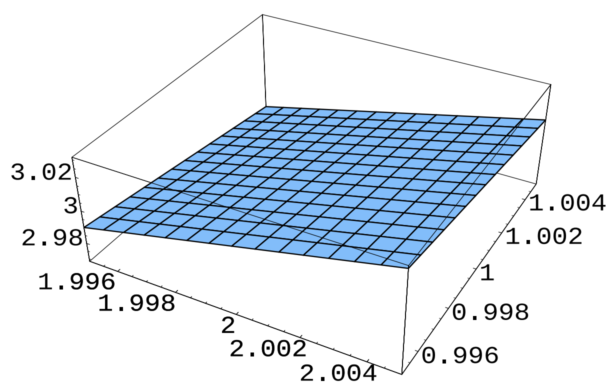

Figure 2: Surface defined by $z = x^2 - y^2$ for $1.995 < x < 2.005$ and $0.995 < y < 1.005$

4 Differentials and infinite microscopes

4.1 Definitions

Let f be a function defined on a domain D of $\mathbb{R}^n$, $P = (r_1, r_2, \ldots, r_n)$ a point of D and $\bar{P} = (r_1, r_2, \ldots, r_n, r_{n+1})$ a point of $\mathbb{R}^{n+1}$ with $r_{n+1} = f(P)$. For any positive infinite hyperreal H, we denote by $\mathcal{M}_H^{\bar{P}}$ the microscope of power H centered on $\bar{P}$.

Definition 1. *A function f is said to be differentiable at P if, for any positive infinite hyperreal H, the standard part of the graph of f in the eyepiece of the infinite microscope $\mathcal{M}_H^{\bar{P}}$ is the graph of a linear function, i.e. a function defined by the following equality:*

$$X_{n+1} = \sum_{j=1}^{n} m_j X_j. \tag{6}$$

For a differentiable function, we can give the definition of the differential.

Definition 2. *If f is differentiable at P, the subset of $\mathbb{R}^{n+1}$ defined by the equation*

$$X_{n+1} = \sum_{j=1}^{n} m_j X_j$$

and which is seen in the eyepiece of the infinite microscope $\mathcal{M}_H^{\bar{P}}$, is the graph of a linear function called the differential of f in P. Then $m_j = \frac{\partial f}{\partial x_j}(P)$, $\forall j = 1 \ldots, n$ and $df(P) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(P) dx_j$.

Looking back at the example of the previous section, we simply get

$$m_1 = \frac{\partial f}{\partial x_1}(2,1) = 4, \; m_2 = \frac{\partial f}{\partial x_2}(2,1) = -2 \text{ and } df(2,1) = 4dx - 2dy.$$

Following these definitions, classical properties of differentials can be given and proved. The comprehensive theory is not the subject of this paper. We rather want to study how undergaduate students can manage these concepts and definitions.

4.2 Didactic thoughts

We now give some points of the new approach that can be seen as advantages for students as compared to a more classical method.

- It is possible to “actually see” the differential at a certain point through an infinite microscope. The principle can be illustrated using software by making successive zooms. This emphasizes the geometrical aspect of the differential ([2], p.19).

- The above definitions also emphasize on the functional aspect of differentials ([2], p.49).
- What is seen in the eyepiece of an infinite microscope only concerns points which are infinitely close to the point under study. This is linked with the idea that the differential approximation is only valid in the neighbourhood of this point. The idea is also consistent with the way practical users tend to use differentials, namely as infinitesimal quantities.
- The developments do not require the calculations of partial derivatives.
- The presentation is similar for univariate or multivariate functions.

5 A student sample experiment

The experiment took place in June 2003. We considered 45 undergraduate university students in Economics. They had followed a first course in multi-variate calculus for 3 months. We asked them the following question inspired by an exercise proposed by Artigue - Menigaux - Viennot ([2], p.12)[1]

> *Let f be a function from $\mathbb{R}^2$ to $\mathbb{R}$ defined by*
>
> $$f(x,y) = x + 3y + |x|^2\left(\sqrt{|\sin x|} + y^3\right).$$
>
> 1. *Find the image of the graph of f in the eyepiece of an infinite microscope pointed on the origin of the space.*
> 2. *Prove then that f is differentiable at (0,0).*
> 3. *Give the differential df of f at (0,0).*

The answers that we hoped for are summarized below:

Solution :

Applying an infinite microscope pointed on (0,0) to the function f gives us the following equation for the image of the graph of f :

$$\frac{Z}{H} = \frac{X}{H} + 3\frac{Y}{H} + \left|\frac{X}{H}\right|^2\left(\sqrt{\left|\sin\left(\frac{X}{H}\right)\right|} + \left(\frac{Y}{H}\right)^3\right)$$

Multiplying by H and taking the standard part of the equation gives us

$$Z = X + 3Y$$

[1] The question was asked in French

as the other terms contain a factor $\frac{1}{H}$ which is infinitely small. The function f is then differentiable at $(0,0)$ since its image through the microscope is a plane and its differential is $df = dx + 3dy$.

With this exercise, we wanted to find out whether the microscope method could facilitate the search for a linear approximation of this kind of functions and avoid the recourse to partial derivatives. The experiment also allowed us to compare our results with those of Artigue *et al.* who had used a more classical method.

The analysis was split into three parts: a first one concerning the image of the graph of f in the microscope, a second one related to the differentiability of f, and a third one about the differential of f.

5.1 About the image in the microscope

The first two steps of the search of the image of f in the microscope were not a problem. Although all students managed to make the change of variables and then take the standard part of the obtained equality, only 26 could interpret their work correctly to give the image of f as $Z = X+3Y$. Of the other students, 8 wrote $f(X,Y) = X+3Y$, making a confusion between the function f and its image by the variables change. The 11 remaining students gave $Y = -\frac{X}{3}$ as an answer. Apparently, they didn't link $f(x,y)$ with the third variable z. They then simply replaced $f(x,y)$ by 0 and obtained an answer which looked like the one they got in the univariate case. Although we did not ask then to interpret the answer, two thirds of the participants explained their result. Among these, 13 said that the image in the eyepiece of the microscope was a plane containing the origin and 17 that it was a straight line also containing the origin. It should be added that among these 17 students, 11 were already wrong with the equation of the image and the 6 others had found the right equation of a plane but could not interpret it correctly.

5.2 About the differentiability

In the second part of the question, the students were asked to prove that the function was differentiable at (0,0). The answers of nineteen of them were based on the results obtained at the first part of the question with the microscope, 16 tried to use the classical definitions and 10 couldn't give any answer to the question. Nobody could give a correct explanation referring to classical methods :

- 7 tried to justify the differentiability by the continuity at (0,0);
- 2 put forward the fact that the function was derivable at (0,0);
- 6 gave the two reasons above together;
- 1 invoked that the gradient of f was different from $\overrightarrow{0}$.

Among the students who chose the microscope, only 11 proofs were correct justifying the differentiability by the fact that the image of the function in the eyepiece of the microscope was a plane. The 18 others gave a wrong explanation, saying that the function was differentiable because its image in the microscope was a straight line. Here again, we find a confusion between the univariate and multi-variate cases. The results about the differentiability of the function are summarized in the following table:

Microscope		Classical method	No answer
Right	Wrong	Wrong	
11 (24.4%)	8 (17.8%)	16 (35.6%)	10 (22.2%)

5.3 About the calculation of the differential

At first sight, we might think that the results are satisfactory. Indeed, only 7 students did not find $df = dx + 3dy$. A closer look however revealed that the interpretation was not so clear. Actually, 29 students chose to calculate the partial derivatives in (0,0), none of them found the right answer for the two derivatives but, putting $x = 0$ and $y = 0$ at the end of the calculation, gave them the correct final result. Here are, for example, the two answers obtained for $\frac{\partial f}{\partial x}$ (the first was given 18 times, the second 11 times):

$$\frac{\partial f}{\partial x} = 1 + 2x\left(\sqrt{|\sin x|} + y^3\right) + \frac{x^2 \cos x}{\sqrt{|\sin x|}}$$

$$\frac{\partial f}{\partial x} = 1 + 2x\left(\sqrt{|\sin x|} + y^3\right) + \frac{x^2 |\cos x|}{\sqrt{|\sin x|}} \; .$$

Using the microscope method, 9 students answered correctly $df = dx + 3dy$; five of them inferred from their results the values of the partial derivatives, i.e. $\frac{\partial f(0,0)}{\partial x} = 1$ and $\frac{\partial f(0,0)}{\partial y} = 3$. Moreover 3 persons wrote the differential as a linear function of the variables X and Y; it would have been interesting to know if these students understood the fact that the values taken by the two variables were infinitesimal.

It is worth mentioning that 4 students, who had chosen to calculate the partial derivatives, found out that their results were consistent with those obtained for the first question with the infinite microscope.

5.4 Conclusions

Considering the results of this experiment, we believe that the use of an infinite microscope and the calculations linked to it were assimilated quite well by the students. The novel approach offers the advantage of insisting on the intuitive aspects of differentials by defining the differential as the linear function whose graph is seen in the eyepiece of an infinite microscope. The algorithmic aspect of the calculations can be considered from different standpoints. It can be seen as an advantage by some students who are often pleased with this kind of exercise. On the other hand, some would say this leads to a systematisation of the calculations and hides the real meaning of the search for the differential. We think that these two aspects are not necessarily opposite but can be complementary and that, for some students, *practice can give meaning* but this would require too long of a debate to be discussed here. Further, we must state that the link between the image of the graph of a curve in the eyepiece of an infinite microscope and the differential is still not clear for students. About this fact, we should add that students had first been confronted with the classical method and that the microscope method applied to the differentials was new. This might explain why many of them answered the first question correctly but did not think to use their results to answer the second and third questions. Further work would really be necessary to complete this study. We would like to teach this method to students who do not know anything about differentials and see whether the results would be the same or significantly different from those obtained in this limited experiment.

References

[1] Antibi, A., Bair J. and Henry V. 2004. Limites de courbes: théorie et applications en analyse. *Mathématique et Pédagogie*. 147: 65-87.

[2] Artigue M., Menigaux J. and Viennot L. 1989. *Questionnaires de travail sur les différentielles*. Irem et Université de Paris VII. 76.

[3] Artigue M., Menigaux J. and Viennot L. 1990. Some aspects of students' conceptions and difficulties about differentials. *European Journal of Physics*. 11: 262-267.

[4] Briand J. and Chevalier M.C. 1995. *Les enjeux didactiques dans l'enseignement des mathématiques*. Paris: Hatier.

[5] Greco. 1989. Didactique et acquisition des connaissances scientifiques. Groupe Mathématiques et Physique - Enseignement supérieur. Irem et Université de Paris VII. 74.

[6] Henry V. 2004. *Questions de didactique soulevées par un enseignement de l'analyse non standard à de futurs économistes*. Toulouse: Doctoral thesis.

[7] Keisler H.J. 1976. *Elementary Calculus*. Boston: Prindle, Weber & Schmidt Inc.

[8] Robinson A. 1961. *Non-standard Analysis*. Koninkl. Ned. Akad. Wetensch. Proc. Ser. A 64: 432-440.

[9] Sullivan K.A. 1976. The Teaching of Elementary Calculus Using Non-standard Approach. *American Mathematical Monthly*. 370-375.

[10] Tall D.O. 1980. Looking at graphs through infinitesimal microscopes, windows and telescopes. *Mathematical Gazette*. 64: 22-49.

Short Biography

Valerie Henry was born in Liege (Belgium) in 1977. She got her Degree in Mathematics at the University of Liege in 1999 and her PhD in Didactic of Sciences at the University Paul Sabatier in Toulouse (France) in 2004. Working since 1999 as an assistant at the service of applied mathematics in the Faculty of Economics, Business and Social Sciences at the University of Liege, she is principally interested in Didactics of Mathematics and Statistics, especially concerning the teaching to students in Economics and Business. Her thesis was moreover dedicated to several didactic questions about the teaching of nonstandard analysis to this kind of students.