Using CAS for the visualization of some basic concepts in calculus of several variables

Ungsana Chundang
iungdang@.cc.kmitt.ac.th
Department of Mathematics
King Mongkut 's Institute of Technology Thonburi

Abstract

In teaching calculus of several variables for engineering students or science students one of the important aims is to create a good comprehension of the basic concepts appearing in these topics. This comprehension should give the students the capability to use their knowledge in solving application problems and/or in their further studies. In many cases the introduction of a new concept can be done by using a good visualization. A CAS with it's rich variety of capabilities, especially graphic capabilities, makes it possible to use dynamical visualization in the introduction of such new concepts.

In this paper an introduction of the concepts of level line, gradient, directional derivative using dynamical visualization is presented. Finally an extremum problem with a constraint is treated by dynamical visualization to get the Lagrange multiplier.

Keywords

Visualization; basic idea; gradient; directional derivative; extremum problem with constraint.

1. Introduction

Teaching calculus to engineering students should promote insight into essential mathematical concepts as well as into important mathematical contents (e.g. main theorems of calculus). The teaching should promote a reflective and deep understanding of the basic concepts of calculus which are important in the further studies of the students. If possible the students should develop geometrically oriented “basic ideas” of such concepts; e.g. they should think of derivatives as rates of change. (For the concept “basic idea” see [7].)

For engineering and also for science students, some mathematical concepts are needed much earlier in other subjects than they appear in a calculus course. So it make sense to introduce sometimes a concept in a preformal way. In such a situation, computers, especially CAS, are very helpful. The graphical possibilities of CAS can be used to generate a good basic idea. This basic idea should provide the students with the knowledge about the properties of the concept they need. Here dynamical visualization is an excellent tool to create good basic ideas in an impressive way. In[3] and [8], there are many comments concerning a visual approach in teaching based on the computer. Let us also quote [1] and [5]. In [2], a great number of animations are considered.

We describe now some examples of such visualizations which allow a vivid introduction into important concepts.

Here the following topics are treated:

1. The concept of gradient for real valued functions of two variables.
2. The concept of directional derivative of such a function and the usual rule how to calculate it.
3. An extremum problem with constraint and the geometrical meaning of such a problem together with the way how to solve it.

Concerning (1) and (2) , it is remarked to the didactical aims among other things that the students should be taught that a real valued function of two variables can be represented in different ways. This can be done, for instance, by its graph (i.e. by a surface in space) and also by the use of level curves. The students should become familiar with different kinds of geometrical representations of a function. Above all, they should be able to interpret a plot of level curves and to translate a simple real situation they have learned in the lessons, into mathematics. They should be able for example to interpret a topographical map of a landscape as a set of level curves of a certain function. They should also understand that different real situations, such as isotherms or isobars lead to the same abstract interpretation. Furthermore they should understand the concept of directional derivative as a local changing rate in a given arbitrary direction. They should also know the meaning of the gradient by its properties and how it is used to calculate a directional derivative, if the direction is given.

According to (1) the concept of level curves is stressed in order to introduce by local geometrical considerations the concepts of *grad f* at a point *P* as a vector. Its direction is the direction of steepest increase of *f* at *P* and it turns out that this vector is orthogonal to the level curve through *P* and can be defined by $(f_x(P),\ f_y(P))$.

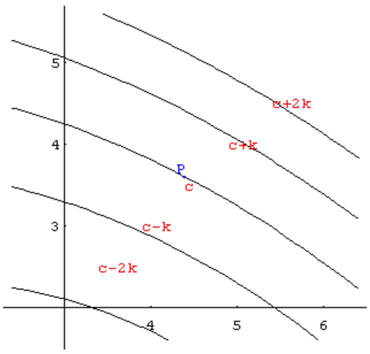
It is known empirically that most of the students have problems to understand the definition of the derivative for functions of several variables (here two) as a linear map from $\mathbb{R}^2$ into $\mathbb{R}$. Even the concept of gradient in the first moment causes difficulties. It is easy to impart the concept of partial derivatives since these are based on the concept of one dimensional derivative which usually is introduced by considering local changing rates. The geometrical introduction of the gradient should be based on both ideas. We can visualize the fact that a C^1-function locally behaves like a linear function. And at the same time we can visualize the fact that the direction of steepest increase of the function is orthogonal to the level lines.

In the following pictures respectively animations it is tacitly assumed that the functions under consideration are sufficiently smooth.

2. Gradient of a function

In the first step, we will show that the direction, in which f increases most rapidly at a point *P*, is orthogonal to the level curve of f, which passes through the point *P*. Here we have assumed that such a point exists. We start the visualization by considering graph f in three dimensions, together with the curves which arise as intersection of the surface $z = f(x, y)$ with the planes $z = c$. Then we project these curves in the xy-plane in order to obtain the level curves of f.

After that, we use the idea of the so-called function-microscope (see [4]), that is zooming level curves in a small neighbourhood of *P* again and again. It turns out that finally all level curves in this neighbourhood appear as straight lines. It can also be seen that finally these straight lines are parallel to each other. The level curve through *P* denotes the value *c* of the function. Here we show only the level curves $c\text{-}2k$, $c\text{-}k$, c, $c+k$, and $c+2k$, where k is real and positive. In the process of zooming the parameter k runs through a sequence of k-values k_p which tends to zero if we would repeat this process infinitely often.



$\Rightarrow \ldots \Rightarrow \ldots \Rightarrow \ldots \Rightarrow \ldots \Rightarrow \ldots$

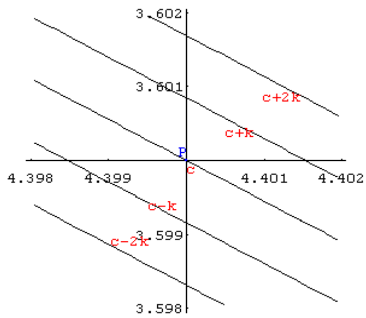
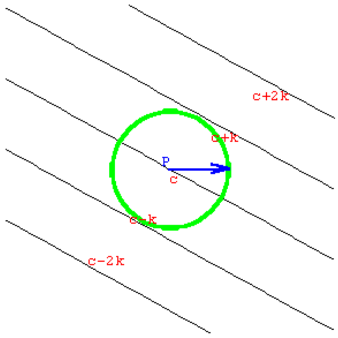


Figure 2.1 Figure 2.2

In the final picture, we use k again instead of k_p. But we note that we have here $0 < k << 1$. The sequence of arrows between the first picture (Figure 2.1) and the last picture (Figure 2.2), indicates that one can see here on the screen, in fact, an animation.

In the last picture, we draw now a circle around the point P which touches the level curve $c+k$ see Figure 2.3. Considering the value of f at all points of this disk, it can be seen immediately that f takes its maximum value at that point where the circle touches the level curve $c+k$ see Figure 2.4. Therefore, it can be seen that the direction of most rapid increase of f at the point P is orthogonal to the level curve c.



$\Rightarrow\ldots\Rightarrow\ldots\Rightarrow\ldots\Rightarrow\ldots\Rightarrow\ldots$

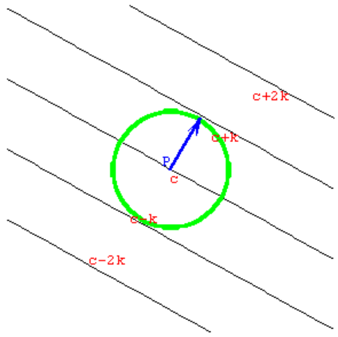


Figure 2.3 Figure 2.4

The sequence of arrows has the same meaning as mentioned earlier.

In the second step, we consider now the local rate of change of f in the direction of the x-axis and in the direction of the y-axis. Suppose P has the co-ordinates (x_0, y_0). We draw a line from the point P parallel to the x-axis until it meets the level curve $c+k$ and a second line parallel to the y-axis until it meets the level curve $c+k$. Suppose that these both segments have the length a and b respectively (see Figure 2.5).

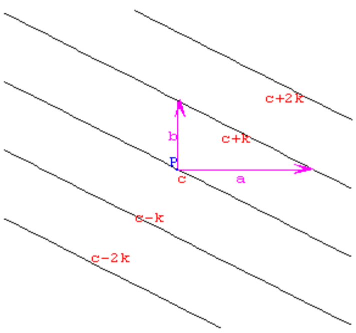
Considering the partial difference quotients, we get

$$\frac{f(x_0+a, y_0) - f(x_0, y_0)}{a} = \frac{k}{a} = f_x(x_0, y_0) + o(1) \text{ for } k \to 0$$

and

$$\frac{f(x_0, y_0+b) - f(x_0, y_0)}{b} = \frac{k}{b} = f_y(x_0, y_0) + o(1) \text{ for } k \to 0.$$

Note here that $k \to 0$ implies that $a \to 0$ and $b \to 0$. That means the vector (f_x, f_y) is the limit of the vector $(\frac{k}{a}, \frac{k}{b})$, if $k \to 0$. (See also Figure 2.6).



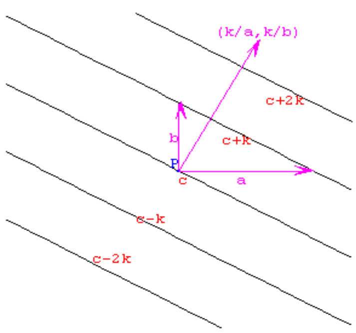


Figure 2.5 Figure 2.6

Now we draw the vector $(a, -b)$ from the point P (this vector is parallel to the level curve c) (see Figure 2.7). Since the scalar product of $(\frac{k}{a}, \frac{k}{b})$ and $(a, -b)$ equals zero, the same is true for the scalar product of $(\frac{k}{a}, \frac{k}{b})$ and a vector lying in the level line c. So in the limit we have the desired result that the vector (f_x, f_y) is perpendicular to the level curve c.

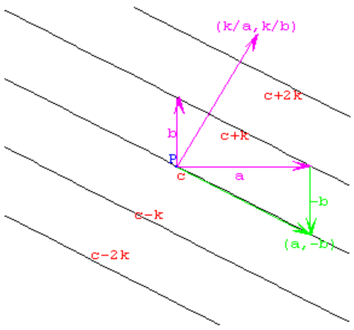


Figure 2.7

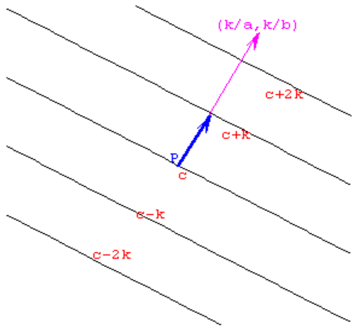


Figure 2.8

Figure 2.8, which shows the vector $(\frac{k}{a}, \frac{k}{b})$ together with the plot of the direction of most rapid increase of f, emphasizes the fact that this direction is the direction of (f_x, f_y).

Now one can define the gradient of f as *grad f*(P): = $(f_x(P), f_y(P))$ and we have the essential properties that *grad f*(P) is orthogonal to the level curve through P and that its direction is the direction of most rapid increase of f.

Note: What we have done now is nothing else as to consider the behaviour of a linear function. But the method of zooming, which leads in a very impressive way to this special case, also allows some insight into the local behaviour of a more general function. So this elementary way of introducing the concept of gradient by simple geometric arguments maybe gives some help to an engineering student not only to use formal definitions and rules.

3. Directional derivative

Based on the introduction of the gradient, we want to define now the directional derivative of f with respect to a direction given by a unit vector $\vec{e} = (\cos\alpha, \sin\alpha)$ as the local changing rate

$$\partial f_{\vec{e}}(p) = \lim_{t \to 0} \frac{f(P + t\vec{e}) - f(P)}{t} .$$

Let us come back to the previous picture, in which we defined the gradient of f at P. We represented the function f by level lines. Zooming, in a neighbourhood of P, we got finally parallel level lines (see Figure 3.1).

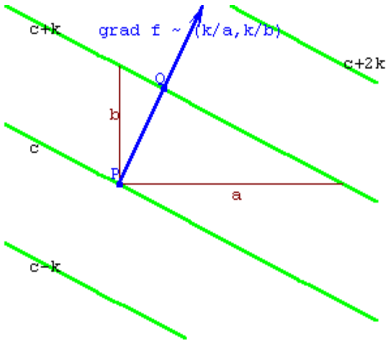


Figure 3.1

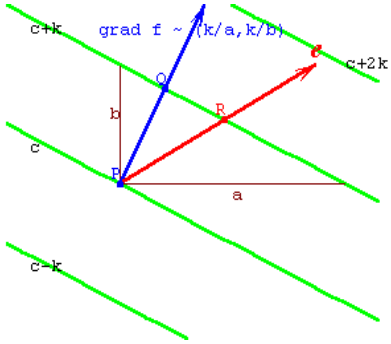


Figure 3.2

We defined *grad f*(*P*) = $\lim\limits_{k\to 0}(\frac{k}{a},\frac{k}{b})$, and we know that in this simplified situation the angle at Q equals $\frac{\pi}{2}$. If we choose P as the origin in a coordinate system, then the level curve $c+k$ satisfies the equation $\frac{x}{a}+\frac{y}{b}=1$.

Consider now, additionally, the unit vector $\vec{e}=(\cos\alpha,\sin\alpha)$, and let R be the point where the line with the direction $\vec{e}$ meets the level line $c+k$ (see Figure 3.2).

Let τ be the distance $|PR|$. Then

$$\frac{f(P+\tau\vec{e})-f(P)}{\tau}=\frac{f(R)-f(P)}{\tau}=\frac{k}{\tau}.$$

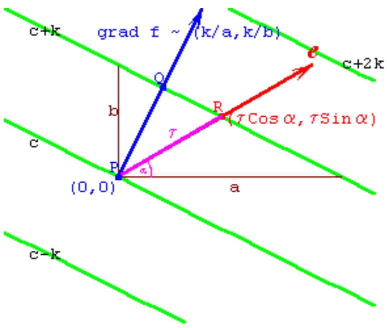


Figure 3.3

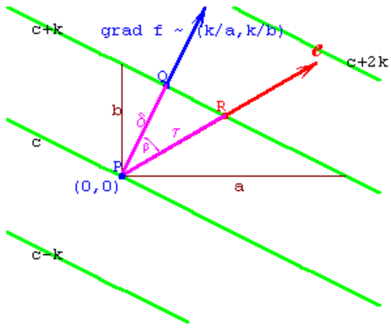


Figure 3.4

From Figure 3.3, we see that the co-ordinates of f at R are $\tau\cos\alpha$ and $\tau\sin\alpha$ respectively. But R is on the level line $c+k$, so we have $\frac{\tau\cos\alpha}{a}+\frac{\tau\sin\alpha}{b}=1$.

Multiplication by $\frac{k}{\tau}$ gives

$$\frac{k}{a}\cos\alpha+\frac{k}{b}\sin\alpha=\frac{k}{\tau}\quad\text{which is}\quad(\frac{k}{a},\frac{k}{b})\bullet\vec{e}=\frac{k}{\tau},$$

where $\bullet$ stands for the scalar product. $k\to 0$ implies $\tau\to 0$.
So for $k\to 0$, we find

$$grad\ f(P)\bullet\vec{e}=\partial f_{\vec{e}}(p).$$

Hence our result is:
The directional derivative of f at P in the direction $\vec{e}$ can be calculated as the scalar product of the gradient of f at P and the vector $\vec{e}$.

Another way to conclude this result from the picture is as follows:
Consider Figure 3.4, since the angle at Q is $\frac{\pi}{2}$, we see from the triangle PRQ that $\cos\beta = \frac{\delta}{\tau}$, where δ denotes the distance $|PQ|$. And Q lies on the level line $c+k$, that is on the line $\frac{x}{a} + \frac{y}{b} = 1$.
The well-known formula for the distance of a point from a straight line (Hesse's formula) gives

$$\delta = \frac{1}{\sqrt{\frac{1}{a^2} + \frac{1}{b^2}}}.$$

That is

$$\left\| \left(\frac{1}{a}, \frac{1}{b} \right) \right\| = \frac{1}{\delta}.$$

So finally,

$$\left\| \left(\frac{k}{a}, \frac{k}{b} \right) \right\| \; \|\vec{e}\| \cos\beta = \frac{k}{\delta} \frac{\delta}{\tau} = \frac{k}{\tau}.$$

From $k \to 0$ we get

$$\|grad\ f(P)\| \; \|\vec{e}\| \cos\beta = \partial f_{\vec{e}}(p).$$

That is

$$\partial f_{\vec{e}}(p) = grad\ f(P) \bullet \vec{e}.$$

Note: The linear way to show here preformally the formula for the directional derivative avoids at this place the chain rule and it shows the simple geometric background of this formula.

4. An extremum problem with constraint

Finally, as an application of the part about gradient, we consider an extremum problem with constraint (maximize or minimize the function f under the constraint $g = 0$).
Here, the geometrical background of the approach to the solution of the problem can be studied. We visualize in detail the relation between gradients of the function f and the gradients of the constraint function g.

The function f to be considered here is

$$f(x, y) = xy \;; \qquad \text{the constraint function } g \text{ is} \qquad g(x, y) = x^2 + 4y^2 - 8.$$

First, the problem is visualized in $\mathbb{R}^3$. The graph f is plotted (see Figure 4.1) and then the graph f is plotted together with the constraint (see Figure 4.2), that means the curve $g = 0$ is lifted on to the graph of f. Then using an animation, we move a point on the surface along the constraint (that means along the lifted curve). This dynamical visualization explains clearly this kind of extremum value problem and moreover it indicates immediately where the maximum (minimum respectively) of f is attained.

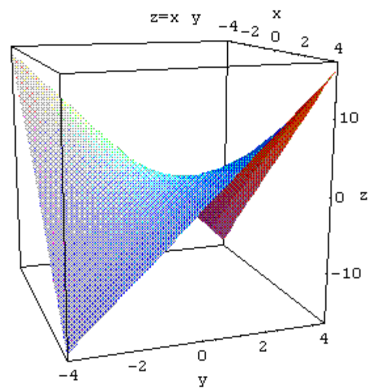


Figure 4.1

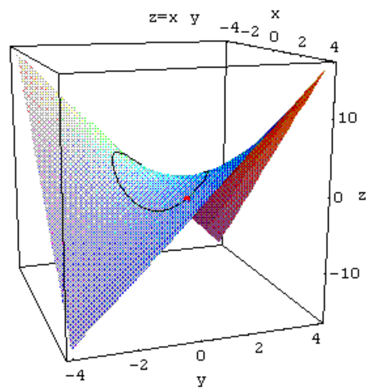


Figure 4.2

To visualize the problem in two dimensions, and to investigate how to find the solution in a geometrical way, we consider level curves of f together with the (level) curve $g = 0$ (see Figure 4.3). Here, using the picture, we point out the fact that extremum points of f cannot occur at points where the curve $g = 0$ crosses a level curve of f. That is an extreme point of f can occur only at a point P where the curve $g = 0$ is tangential to the level curve of f through P. See also [6] where "this principle of touching level lines" is beautifully described.

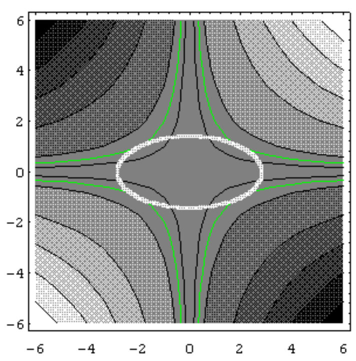


Figure 4.3

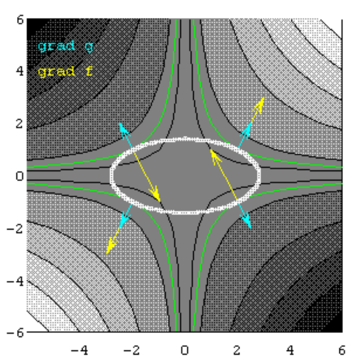


Figure 4.4

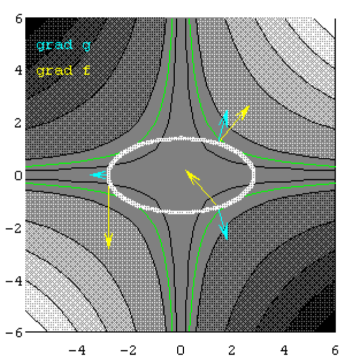


Figure 4.5

From this, together with the property that the gradient of a function at a point is orthogonal to the level curve of the function which passes through that point, we can conclude that *grad f* and *grad g* are parallel to each other at an extreme point P. That is, we have

$$\mathit{grad}\, f = \lambda \cdot \mathit{grad}\, g \quad \text{and} \quad g = 0$$

as the necessary conditions for an extreme point.

As is indicated in Figure 4.4 and Figure 4.5, we show on the screen an animation which moves a point Q together with *grad f* (Q) and *grad g* (Q) along the curve g=0. Here one can see exactly at which points the two gradients become parallel.

These kinds of visualization, as described in section 2 and in section 3, presumably help the students to get a better understanding of the concepts "gradient" and "directional derivative". And the visualization, as described in section 4, will support the students' understanding about the geometrical background of a method they learn and apply formally. Furthermore, the visualization of the graph of a function appearing in a problem will help the students to understand the problems better. They can also acquire self-confidence in analysing or solving problems.

References

[1] Braden, B., D K. Krug, P. W. McCartney, S. Wilkinson: “Discovering Calculus with Mathematica,” John Wiley & Sons, New York, 1992.

[2] U. Chundang: “On the Use of Computer Algebra Systems in a Calculus Course for Thai Engineering Students -Developing and Testing Modules for Visualization-,” Thesis, Universitaet Kassel, 1996.

[3] S. Cunningham: “Some strategies for using visualization in mathematics teaching,” *ZDM*, 3, 1994, pp 83-85.

[4] A. Kirsch: “Ein Vorschlag zur visuellen Vermittlung einer Grundvorstellung zum Ableitungsbegriff,” *MU* 25, Heft 3, 1979, pp 25-41.

[5] E. Packel, S. Wagon: “Animating Calculus. Mathematics Notebooks for the Laboratory,” W. H. Freeman and company, New York, 1994.

[6] G. Polya: “Mathematics and Plausible Reasoning,” Introduction and analogy in Mathematics, Volume 1, Princeton University Press, 1954.

[7] R. vom Hofe: ”Concepts about the Generation of Adequate Images and Basic Ideas-Normative, Descriptive and Constructive Aspects,” What is Research in Mathematics Education, and What are Results, ICMI-publication, Washington DC, 1994.

[8] W. Zimmermann, S. Cunningham (eds.): “Visualization in Teaching and learning Mathematics,” MAA Notes, Number 19, Mathematical Association of America, USA, 1991.