The Partial Derivative

A misleading application of this notion, differing from that which is commonly accepted in mathematics, is in widespread use in physics

Contents

1. Overview

In physics, quite frequently functions of the form $L(t, q(t), \dot{q}(t))$ are encountered. It's the characteristic feature of this type of functions, that some variables (in this example t) are showing up both explicitly and implicitly (in this example as variable of q and of $\dot{q}$). Such mixed functions are not compatible with the tenets of pure mathematics, because they don't comply with the standard mathematical rules for computing derivatives. But as many physicists would not like to do without mixed functions, they extended the rules of calculus such, that the incompatibility of mixed functions and calculus is remedied.

For that purpose, physicists discern between a "total derivative"

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{\partial L}{\partial t} + \frac{\partial L}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial L}{\partial \dot{q}}\frac{\mathrm{d}\dot{q}}{\mathrm{d}t} \ , \tag{1}$$

in which the chain-rule is to be applied, and an "explicit derivative"

$$\frac{\partial L}{\partial t} \ , \tag{2}$$

in which the chain rule must *not* be applied. Instead the derivative is to be performed only with respect to the explicit variable. The quite misleading notion "partial derivative" is commonly used for the explicit derivative.

The method of physicists will be called convention P in the sequel. For the tenets of pure mathematics, we will use in this article the notion convention M. Furthermore we will define a convention M+ , whose nomenclature is similar to the nomenclature of convention M, while M+ in content is identical to convention P. Thus convention M+ will be helpful to clearly identify the differences in-between the conventions.

In the following section we will shortly compile the rules of computing derivatives according to convention M. After that we will introduce the differing rules according to convention M+ and according to convention P. In particular we will clarify the notion "partial derivative", which is used in a misleading manner in convention P. In the last section, a hint for the correct application of convention P is given.

2. The Rules for the Computation of Derivatives

2.1. Convention M

We use the notion convention M for those rules for the computation of derivatives, which are customary in the mathematical literature.

Definitions:

f is the explicit, and x is the implicit variable of the function $j(f(x))$. The function $m(n(x), f(x), y)$ has the explicit variables n, f, y, and the implicit variable x. x is an explicit variable of n and of f.

The derivative of a function f, which is depending on exactly one explicit variable x (with x possibly, but not necessarily, depending on further variables), is defined by

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} \equiv \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \quad . \tag{3a}$$

If a function $g(x, y)$ is depending on several explicit variables (which again possibly, but not necessarily, might depend on further variables), then the partial derivative of $g(x, y)$ with respect to x is defined by

$$\frac{\partial g(x,y)}{\partial x} \equiv \lim_{\Delta x \to 0} \frac{g(x+\Delta x, y) - g(x,y)}{\Delta x} \quad . \tag{3b}$$

The partial differential of a function with respect to x is defined by

$$\mathrm{d}_x f(x) \equiv \frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x \tag{4a}$$

$$\mathrm{d}_x g(x,y) \equiv \frac{\partial g}{\partial x}\,\mathrm{d}x \ . \tag{4b}$$

The total differential of a function is by definition the sum of it's partial differentials with respect to all of the function's explicit variables:

$$\mathrm{d}f(x) \equiv \mathrm{d}_x f = \frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x \tag{5a}$$

$$\mathrm{d}g(x,y) \equiv \mathrm{d}_x g + \mathrm{d}_y g = \frac{\partial g}{\partial x}\,\mathrm{d}x + \frac{\partial g}{\partial y}\,\mathrm{d}y \tag{5b}$$

"The differential $\mathrm{d}x$ of an independent variable x is equal to it's increment, to which an arbitrary value may be assigned".[1] [1]

No partial derivative is defined for functions, which depend on only one explicit variable.

For functions which depend on several explicit variables, no derivative (without the word "partial" and marked by the sign d) with respect to one of these variables (i. e. something like $\frac{\mathrm{d}g(x,y)}{\mathrm{d}x}$) is defined.

Thus convention M is using the sign ∂ and the notion "partial derivative", if the derivative is computed for a function, which does depend on several explicit variables. Convention M in contrast is using the notion "derivative", if the derivative is computed for a function which does depend on one single explicit variable only.

There are not any consequences caused by the nominal discrimination in-between derivative and partial derivative in convention M. In particular, the chain rule (see below) is applied both in derivatives and in partial derivatives. Differentials however must always be marked with the sign d, and a partial differential of a function with several variables is something different from the function's total differential.

Chain-Rule:
The derivative resp. the partial derivative of a function with respect to a variable x is equal to the sum of the derivatives resp. partial derivatives of the function with respect to each of it's explicit variables, where each summand must be multiplied by the derivative

[1] „Das Differential $\mathrm{d}x$ einer unabhängigen Variablen x ist gleich ihrem Zuwachs, dem man einen beliebigen Wert beimessen kann.“ The meaning of this sentence is in german as hard to grasp as in my english translation.

resp. partial derivative of this explicit variable with respect to x:

$$\frac{\partial m(n(x),f(h(x)),y)}{\partial x}=\frac{\partial m}{\partial n}\frac{\mathrm{d}n}{\mathrm{d}x}+\frac{\partial m}{\partial f}\underbrace{\frac{\mathrm{d}f}{\mathrm{d}x}}_{\frac{\mathrm{d}f}{\mathrm{d}h}\frac{\mathrm{d}h}{\mathrm{d}x}}+\frac{\partial m}{\partial y}\underbrace{\frac{\mathrm{d}y}{\mathrm{d}x}}_{=0} \tag{6}$$

A function's partial differential with respect to a variable x is equal to the sum of the function's partial derivatives resp. the derivative with respect to each of it's explicit variables, where each summand must be multiplied by the derivative resp. partial derivative of this explicit variable with respect to x, and must be multiplied with the differential $\mathrm{d}x$:

$$\mathrm{d}_x p(f(x),g(x,y))=\frac{\partial p}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x+\frac{\partial p}{\partial g}\frac{\partial g}{\partial x}\,\mathrm{d}x \tag{7}$$

A function's total differential is equal to the sum of it's partial differentials with respect to the explicit variables:

$$\mathrm{d}p(f(x),g(x,y))=\mathrm{d}_f p+\mathrm{d}_g p \tag{8}$$

Mixed Functions:
We use the notion "mixed functions" for functions of the type $q(f(x),x,y)$, in which minimum one variable is showing up both implicitly and explicitly. Let's try to compute the partial derivative of q with respect to x:

$$\frac{\partial q(f(x),x,y)}{\partial x}=\frac{\partial q}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}x}+\frac{\partial q}{\partial x}\underbrace{\frac{\mathrm{d}x}{\mathrm{d}x}}_{=1}+\frac{\partial q}{\partial y}\underbrace{\frac{\mathrm{d}y}{\mathrm{d}x}}_{=0}\quad\text{is wrong!} \tag{9}$$

On both sides of the equation the term $\frac{\partial q}{\partial x}$ is showing up, although $\frac{\partial q}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}x}$ in general is different from zero. The same problem is encountered, if one tries to compute the partial differential of a mixed function with respect to that variable which is both explicit and implicit (and of course the failure propagates to the total differential):

$$\mathrm{d}_x q(f(x),x,y) = \frac{\partial q}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x + \frac{\partial q}{\partial x}\underbrace{\frac{\mathrm{d}x}{\mathrm{d}x}}_{=1}\mathrm{d}x + \frac{\partial q}{\partial y}\underbrace{\frac{\mathrm{d}y}{\mathrm{d}x}}_{=0}\mathrm{d}x \tag{10}$$

Again the ominous $\frac{\partial q}{\partial x}$ is showing up, which results into nonsense according to (9), because it is ambiguous. This is the reason, why mixed functions are incompatible with convention M, and therefore are *not permitted*. According to convention M, each variable of a function must either be exclusively explicit or exclusively implicit.

This requirement of convention M can easily be met due to the definition of auxiliary functions like

$$X(x) \equiv x \tag{11}$$

and the substitution

$$q(f(x),x,y) \longrightarrow Q(f(x),X(x),y)\ . \tag{12}$$

X is exclusively an explicit variable and x is exclusively an implicit variable of Q. Thus the partial derivative of Q with respect to x becomes

$$\frac{\partial Q(f(x),X(x),y)}{\partial x} = \frac{\partial Q}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}x} + \frac{\partial Q}{\partial X}\underbrace{\frac{\mathrm{d}x}{\mathrm{d}x}}_{=1}\quad , \tag{13}$$

and the partial differential becomes

$$\mathrm{d}_x Q(f(x),X(x),y) = \frac{\partial Q}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x + \frac{\partial Q}{\partial X}\frac{\mathrm{d}X}{\mathrm{d}x}\,\mathrm{d}x \quad . \tag{14}$$

Not all physicists were satisfied by the alternative, to eliminate mixed functions by means of appropriate auxiliary functions. Therefore they invented another method to make mixed functions acceptable, which we are calling "convention P". As an intermediate step, we will describe in the next section "convention M+". M+ is differing from P only by nomenclature, but there is no difference in content.

2.2. Convention M+

The meaningless equation (9) could be transformed into the correct equation (13) due to the auxiliary function $X(x) \equiv x$ and the substitution $q(f(x), x, y) \rightarrow Q(f(x), X(x), y)$. Now we transform X back to x and Q back to q. Clearly we must not simply replace X by x and Q by q, because that would directly lead back to the meaningless equation (9). The essential point in (13) is, that the term $\frac{\partial Q}{\partial X}$ is uniquely specifying, that the derivative must be computed with respect to the explicit X, but not with respect to the implicit x, which is hidden in f. This distinct specification must not get lost in the process of back-transformation. Therefore the correctly back-transformed equation (13) is

$$\frac{\partial q(f(x), x, y)}{\partial x} = \frac{\partial q}{\partial f} \frac{\mathrm{d} f}{\mathrm{d} x} + \left. \frac{\partial q}{\partial x} \right|_e \tag{15}$$

$$\Big|_e \equiv \text{derivative with respect to the explicit variable only}$$

The back-transformation of the correct partial differential (14) results into

$$\mathrm{d}_x q(f(x), x, y) = \frac{\partial q}{\partial f} \frac{\mathrm{d} f}{\mathrm{d} x} \mathrm{d} x + \left. \frac{\partial q}{\partial x} \right|_e \mathrm{d} x \quad . \tag{16}$$

To cast this into a simple rule, we apply the mark $|_e$ *always*, if in application of the chain rule the derivative with respect to an explicit variable is performed, and not only in those cases where it is needed for the discrimination versus the derivatives resp. the partial derivatives of mixed functions with respect to their implicit variables:

$$\frac{\partial q(f(h(x)),x,y)}{\partial x} = \left.\frac{\partial q}{\partial f}\right|_e \underbrace{\frac{\mathrm{d}f}{\mathrm{d}x}}_{\left.\frac{\mathrm{d}f}{\mathrm{d}h}\right|_e \frac{\mathrm{d}h}{\mathrm{d}x}} + \left.\frac{\partial q}{\partial x}\right|_e \tag{17}$$

$$\mathrm{d}_x q(f(x),x,y) = \left.\frac{\partial q}{\partial f}\right|_e \frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x + \left.\frac{\partial q}{\partial x}\right|_e \mathrm{d}x \tag{18}$$

That's all. The following new formulation of the rules for the computation of derivatives resp. partial derivatives and partial differentials will obviously always give the correct results, even in case of mixed functions:

The derivative resp. the partial derivative of a function with respect to a variable x is equal to the sum of the derivatives resp. partial derivatives — which must be marked by the sign $|_e$ — of the function with respect to each of it's explicit variables, where each summand must be multiplied by the derivative resp. partial derivative of this explicit variable with respect to x. The mark $|_e$ is indicating, that only the derivative with respect to the explicit variable shall be computed, i.e. that the chain-rule shall not be applied.

A function's partial differential with respect to a variable x is equal to the sum of the function's partial derivatives resp. the derivative — which must be marked by the sign $|_e$ — with respect to each of it's explicit variables, where each summand must be multiplied by the derivative resp. partial derivative of this explicit variable with respect to x, and must be multiplied with the differential $\mathrm{d}x$. The mark $|_e$ is indicating, that only the derivative with respect to the explicit variable shall be computed, i.e. that the chain-rule shall not be applied.

This extension of the rules is the only difference in-between convention M+ and convention M.

2.3. Convention P

Convention P if differing from convention M+ only by nomenclature and notation. There is no difference in content.

Convention P eliminates the annoying and useless discrimination between "derivative" with the mark d and "partial derivative" with the mark ∂. Both are named total derivative according to convention P, and are marked by the letter d. Same as in case of the derivative and in case of the partial derivative of convention M, the chain rule is applied in case of the total derivative of convention P.

Thus the letter ∂ isn't needed any more, and a new use can be assigned to it: Convention P utilizes ∂ to replace the mark $|_e$ of convention M+. Thus ∂ in convention P is indicating the explicit derivative resp. the explicit partial derivative, to which the chain rule must *not* be applied.

According to convention P, the derivative with respect to x and the partial differential with respect to x of the mixed function q are written as follows:

$$\frac{\mathrm{d}q(f(h(x)),x,y)}{\mathrm{d}x} = \frac{\partial q}{\partial f}\frac{\mathrm{d}f}{\mathrm{d}x} + \frac{\partial q}{\partial x} \tag{19}$$

$$\mathrm{d}_x q(f(h(x)),x,y) = \frac{\partial q}{\partial f}\underbrace{\frac{\mathrm{d}f}{\mathrm{d}x}}_{\frac{\partial f}{\partial h}\frac{\mathrm{d}h}{\mathrm{d}x}}\,\mathrm{d}x + \frac{\partial q}{\partial x}\,\mathrm{d}x \tag{20}$$

In cases where both types of derivatives give the same results, it's left to the taste of the user, to chose one of both. For example,

$$\frac{\mathrm{d}\,q(f(h(x)),x,y)}{\mathrm{d}y} = \frac{\partial q}{\partial y}\ . \tag{21}$$

Therefore either form may be chosen.

Up to this point, convention P is logical and convenient. But unfortunately, convention P absurdly kept the name “partial derivative”, when the mark ∂ of the partial derivative of convention M was assigned to mark the explicit derivative. In the explicit derivative $\frac{\partial m}{\partial f}$, there is absolutely nothing more “partial” than in the total derivative $\frac{\mathrm{d}m}{\mathrm{d}f}$. But by today the misleading notion regrettably is so deeply rooted in convention P, that we probably will not get rid of it any more. Thus also future generations of physics students will agonize about the notion “partial derivative” and try in vain to grasp any reasonable meaning behind that name.

In the following table, a comparision of convention M+ and convention P is compiled:

	Convention M+		Convention P	
Derivative	Mark	Name	Mark	Name
with chain-rule	d	derivative	d	total deriv.
with chain-rule	∂	partial deriv.	d	total deriv.
without chain-rule	$\vert_e$	explicit deriv.	∂	partial deriv.

3. Beware of Incorrect Use!

The “partial” derivative of convention P is — in spite of it’s delusive name — nothing other than the explicit derivative of convention M+ . Therefore it’s use is sensible only (a) as a factor in a product, which is computed in application of the chain rule, or (b) in cases where both types of derivatives anyway give identical results. Improper application in other places may lead to ambiguous results. To explicate this warning, let’s consider for example the following functions:

$$
\begin{aligned}
a(x,y) &\equiv x+y \\
b\big(a(x),x\big) &\equiv a\cdot x \qquad (22a)\\
X(x) &\equiv x \\
B\big(a(x),X(x)\big) &\equiv a\cdot X \qquad (22b)\\
\beta(x) &\equiv (x+y)\cdot x = x^2+xy \qquad (22c)
\end{aligned}
$$

Obviously

$$b = B = \beta \quad \text{for arbitrary } x,y\ .$$

According to convention P, in the partial (i. e. in the explicit) derivative only explicit variables are considered, and the chain rule is *not* applied:

$$
\begin{aligned}
\frac{\partial b}{\partial x} &= a = x+y \qquad (23a)\\
\frac{\partial B}{\partial x} &= 0 \qquad (23b)\\
\frac{\partial \beta}{\partial x} &= 2x+y \qquad (23c)
\end{aligned}
$$

Therefore according to convention P we find that $\frac{\partial b}{\partial x} \neq \frac{\partial B}{\partial x} \neq \frac{\partial \beta}{\partial x}$, even though $b=B=\beta$ for arbitrary x,y.

There is however never a sensible reason, to apply the partial derivative in convention P with such contradictory results. In practically all cases it is the total derivative, which fits to the physically meant circumstances. Let's consider a typical application: If B is some observable quantity, which depends on several quantities x,y,z,..., then one frequently would like to know how B changes, if for example x changes by a certain amount, while y,z,... remain unchanged. If conventions M or M+ are used, then one finds the answer by means of the partial derivative $\frac{\partial B}{\partial x}$. If convention P is used, then the same answer is found by computing the total derivative $\frac{\mathrm{d}B}{\mathrm{d}x}$. In contrast to M or M+, with convention P the partial derivative $\frac{\partial B}{\partial x}$ will lead to the correct answer only if the dependence on x is exclusively explicit (as for example in the function $\beta(x, y)$ above).

The vector-operators grad, div, rot are defined only for functions, which depend explicitly on the space-time coordinates. It is conventional, to use also in convention P (same as in conventions M and M+) the partial derivatives for their definitions:

∗ Cartesian coordinates with unit vectors $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, $\boldsymbol{e}_z$:

$$\operatorname{grad} f \equiv \frac{\partial f}{\partial x}\,\boldsymbol{e}_x + \frac{\partial f}{\partial y}\,\boldsymbol{e}_y + \frac{\partial f}{\partial z}\,\boldsymbol{e}_z \tag{24a}$$

$$\operatorname{div} g \equiv \frac{\partial g_x}{\partial x} + \frac{\partial g_y}{\partial y} + \frac{\partial g_z}{\partial z} \tag{24b}$$

$$\operatorname{rot} h \equiv \Big(\frac{\partial h_z}{\partial y} - \frac{\partial h_y}{\partial z}\Big)\,\boldsymbol{e}_x + \Big(\frac{\partial h_x}{\partial z} - \frac{\partial h_z}{\partial x}\Big)\,\boldsymbol{e}_y + \\ + \Big(\frac{\partial h_y}{\partial x} - \frac{\partial h_x}{\partial y}\Big)\,\boldsymbol{e}_z \tag{24c}$$

∗ Cylinder coordinates with unit vectors $\boldsymbol{e}_\rho$, $\boldsymbol{e}_\varphi$, $\boldsymbol{e}_z$:

$$\operatorname{grad} f \equiv \frac{\partial f}{\partial \rho}\,\boldsymbol{e}_\rho + \frac{1}{\rho}\frac{\partial f}{\partial \varphi}\,\boldsymbol{e}_\varphi + \frac{\partial f}{\partial z}\,\boldsymbol{e}_z \tag{25a}$$

$$\operatorname{div} g \equiv \frac{1}{\rho}\frac{\partial(\rho\, g_\rho)}{\partial \rho} + \frac{1}{\rho}\frac{\partial g_\varphi}{\partial \varphi} + \frac{\partial g_z}{\partial z} \tag{25b}$$

$$\operatorname{rot} h \equiv \Big(\frac{1}{\rho}\frac{\partial h_z}{\partial \varphi} - \frac{\partial h_\varphi}{\partial z}\Big)\,\boldsymbol{e}_\rho + \Big(\frac{\partial h_\rho}{\partial z} - \frac{\partial h_z}{\partial \rho}\Big)\,\boldsymbol{e}_\varphi + \\ + \Big(\frac{1}{\rho}\frac{\partial(\rho\, h_\varphi)}{\partial \rho} - \frac{1}{\rho}\frac{\partial h_\rho}{\partial \varphi}\Big)\,\boldsymbol{e}_z \tag{25c}$$

∗ Spherical coordinates with unit vectors $\boldsymbol{e}_r$, $\boldsymbol{e}_\varphi$, $\boldsymbol{e}_\vartheta$:

$$\operatorname{grad} f \equiv \frac{\partial f}{\partial r}\,\boldsymbol{e}_r + \frac{1}{r\sin\vartheta}\frac{\partial f}{\partial\varphi}\,\boldsymbol{e}_\varphi + \frac{1}{r}\frac{\partial f}{\partial\vartheta}\,\boldsymbol{e}_\vartheta \tag{26a}$$

$$\operatorname{div} g \equiv \frac{1}{r^2}\frac{\partial(r^2 g_r)}{\partial r} + \frac{1}{r\sin\vartheta}\frac{\partial g_\varphi}{\partial\varphi} + \frac{1}{r\sin\vartheta}\frac{\partial(\sin\vartheta\, g_\vartheta)}{\partial\vartheta} \tag{26b}$$

$$\begin{aligned}\operatorname{rot} h \equiv{}& \frac{1}{r\sin\vartheta}\Big(\frac{\partial(\sin\vartheta\, h_\varphi)}{\partial\vartheta} - \frac{\partial h_\vartheta}{\partial\varphi}\Big)\,\boldsymbol{e}_r + \\ &+ \frac{1}{r}\Big(\frac{\partial(r\, h_\vartheta)}{\partial r} - \frac{\partial h_r}{\partial\vartheta}\Big)\,\boldsymbol{e}_\varphi + \\ &+ \Big(\frac{1}{r\sin\vartheta}\frac{\partial h_r}{\partial\varphi} - \frac{1}{r}\frac{\partial(r\, h_\varphi)}{\partial r}\Big)\,\boldsymbol{e}_\vartheta\end{aligned} \tag{26c}$$

The use of total derivatives would lead to the same results, but conventionally the definition with partial derivatives is chosen.

Partial derivatives, which would give ambiguous results under convention P, are superfluous. Under convention P, partial derivatives should be used only as factors in chain-rule products. They should not be used in other cases, unless the variables, with respect to which the derivatives are performed, are explicit anyway (the vector operators have been mentioned as examples). If this rule is respected, one will always get reasonable and correct results with convention P.

References

[1] I. N. Bronstein, K. A. Semendjajew:
Taschenbuch der Mathematik
(Harry Deutsch, Zürich und Frankfurt/Main, 1973).