On notational differences in physics and in mathematics, preliminary

Ulrich Mutze *

August 15, 2013

It is surprising how differently classical mechanics looks when treated in a textbook of theoretical physics (e.g. Landau and Lifschitz, vol 1) or in one of mathematics (e.g. [1], chapter 13). The question arises whether one can explain the differences by a few principles. Here is my attempt of such an explanation:

In mathematics everything is formalized as sets and functions, whereas in physics one has variables and expressions (terms).

The connection between these two views is rather simple and intuitive, to the effect that nobody seems yet to have seen a need to make it explicit.

The goal of the present work is to give the objects used in a physics-style exposition a clear mathematical definition that makes them amenable to regular mathematical processing.

In physics one has in mind a *situation* (a state of that part of the world we are interested in) which is either described explicitly or is assumed to be clear from the context. It is a part of reality that the reader is assumed to understand and which is normally not represented by symbols in a physics-style presentation. Understanding a particular part of reality allows us to conceive variations and to ask what kind of reactions between parts such variations will cause. So, in effect, we will not be considering a single situation but a multitude of *relevant situations* selected according to the aspects we have in mind.

To make a connection with mathematical style, we describe the multitude of relevant situations as a set $\mathcal{X}$. The variables that physics uses to represent *physical quantities* then become functions defined on $\mathcal{X}$. This means that given a relevant situation $x \in \mathcal{X}$, the variable φ has a value $\varphi(x)$, which reflects this situation and is determined by it. The data type of the value may differ from variable to variable, so that vectors, tensors, and even more complex quantities are allowed as values for physical quantities.

The collection of relevant situations, as already mentioned, is a part of the informal description in a normal treatise in theoretical physics. The functions such as φ provide the connection between the informal and the formalized part of the concept.

*www.ulrichmutze.de

Today, where for larger and larger parts of reality computer models are established (such as digital terrain models, 3D CAD models in engineering and architecture) and the tools are available to build computer models for complex physical systems by means of object oriented programming languages (such as C++ and Ruby), an interesting intermediate way of thinking about $\mathcal{X}$ suggests itself.

In this view, we consider the informal understanding of the relevant situations as a guide for the definition of an object oriented computational model of the relevant situations which then serves as our space $\mathcal{X}$. This then is a concrete formal system which, however, may use quite different notions then our theoretical description of the physical system for which $\mathcal{X}$ is the space of relevant situations.

These things are easier to explain with a concrete example than in due generality. So let us deal with a scenario as set up in the first chapters of a textbook on classical thermodynamics. We want to deal with *simple systems* which in [2], p. 8, are defined as 'systems that are macroscopically homogeneous, isotropic, uncharged, and chemically inert, that are sufficiently large that surface effects can be neglected, and that are not acted on by electric, magnetic, or gravitational fields'. In our computational model we would simulate an idealized version of reality by letting the thermodynamic systems consist of classical particles with specified conservative interaction potentials. Their motion would be confined by containers which would be realized as potential barriers without a particulate structure. The containers would be provided with movable walls for making their volume variable. Software versions of Maxwell's demon would allow to heat and to cool arbitrarily selected parts of the system. Monitoring the impulsive momentum transfer in particle-to-wall collisions one would have microscopic data from which 'the pressure in the container' can easily be derived. Semipermeable membranes (implemented e.g. as sieves) may be inserted and allow separating species of particles which differ in size. Pipes with valves would connect the containers to allow particles to migrate.

Let this computational model be realized as a computer program together with some read-only files from which it reads data, and some write-only files to which it writes data. The output data may include series of digital images so that we may get visualizations of processes in the form of movies. Of course, during run,the program may show images and status indicators on a computer screen.

As any autonomously running program, it goes through a series of *states*, where a state is simply a distribution of values over all available storage places. When the program runs in debug mode, any involved storage area can be inspected so that any state of the program is fully accessible. For a program of the kind as it was sketched here, the states are too complex, however, that for a human it made sense to study and to interpret it.

The program can serve us as an instructive metaphor for a piece of the real world which contains just the objects for which our program contains the corresponding software constructs. By design, 'object oriented' programming languages allow us to express the correspondence between 'real object to be represented' and its data representation in a suggestive manner.

Let us look into this 'real world metaphor' a bit closer. For this purpose we simplify the scenario as much as possible and consider a single container, housing a single species

of particles. A state of the program specifies for each of the particles a position, a momentum, an energy (which may include interaction with neighboring particles and with walls). Further, for all particles that interact with a wall, the program gives a force that the wall 'feels' as a result of this interaction, and a point on the wall to which the force acts. When our program runs, these quantities can be observed readily, whereas for a real thermodynamic system these quantities are hidden to us by their huge number and the very limited control we can have over individual particles. The properties that we can easily access and control in real thermodynamic systems are those treated as *thermodynamic quantities (variables)* in thermodynamics textbooks. In our case, the 'extensive quantities' U of internal energy, V of volume, S of entropy, N of 'mole number', and the 'intensive quantities' p of pressure, T of temperature, and μ of chemical potential. At least in an elementary context these quantities are considered constant in time and referring to a certain kind of states only: the equilibrium states. In our computational model, states change with time but move in the direction of thermodynamical equilibrium automatically, e.g. [3]. By self-suggesting formulas for mean-values over many time steps one computes values for the quantities

$$U, V^1, S, N, p, T, \mu \tag{1}$$

which come out sufficiently constant, to infer equilibrium values from them. Notice that getting these values is very much analogous to making measurements in a lab experiment. What determines the values is a highly complex situation (the program state) which in detail cannot interpreted with reasonable effort. The results of several 'measurements' may give sufficient information on the program state that the results of further 'measurements' can be inferred.

In this example it happens that the computational model of our 'reality' used particle mechanics, whereas the system under actual consideration is a homogeneous one. This reflects the role of statistical mechanics as a theory from which classical thermodynamics can be derived.

Not always the two layers need to differ so profoundly. Since all computational models are finite, discretization will play a role in most cases.

In the following we will develop the framework in which physical quantities are function on the space $\mathcal{X}$ of relevant situations. If we consider dynamical problems it is natural to consider points x and x' that describe situations at different points in time as different. This allows us to consider *time* as a real-valued function on $\mathcal{X}$. So, time is treated on equal footing with physical quantities. Since, as discussed above, it is reasonable to consider for $\mathcal{X}$ rather different options, it makes sense not to assume more than that $\mathcal{X}$ is simply a set. One should be aware that this implies an idealization which, from a physical point of view is very strong: for sets there is a definition of strict equality

$$x = y \quad :\Longleftrightarrow \quad \forall\, z^2\, (z \in x \;\Leftrightarrow\; z \in y)$$

[1] the volume of the gas container may slowly change with time and, particulaly if the geometry is complicated, such as in a Wankel rotary combustion engine, the actual volume may be obtained by Monte Carlo integration

[2] The quantor lets z range 'over all sets', which does not imply that there has to be a 'set of all sets'.

which is an equivalence relation, whereas empirical equality, due to its unsharp character fails to be strictly transitive.

We use the usual Zermelo-Fraenkel set theory which is a monotypic first-order theory with '$\in$' as the only non-logic basic symbol. (Although my primary source, [4], deals with the v. Neumann-Bernays form, in which there is a distinction between sets and classes.) The predicate of being a *relation* is defined as

$$\mathbf{Rel}(r) \quad :\Longleftrightarrow \quad \forall\, a \in r\ \exists\, x, y\ (a = (x, y))\ , \tag{2}$$

where the derived concept of an *ordered pair* is used (e.g. $(x, y) := \{\{x\}, \{x, y\}\}$). Notice that no specification of sets from which the x and y are to be taken is implied. The predicate of being an *equivalence relation* is defined as

$$\begin{aligned} \mathbf{Equ}(r) \quad :\Longleftrightarrow \quad & \mathbf{Rel}(r) \wedge \\ & \forall\, x, y\ ((x, y) \in r\ \Rightarrow\ (y, x) \in r)\ \wedge \\ & \forall\, x, y, z\ ((x, y) \in r\ \wedge\ (y, z) \in r\ \Rightarrow\ (x, z) \in r)\ . \end{aligned} \tag{3}$$

Notice that $\mathbf{Equ}(r)$ implies $(x, y) \in r \Rightarrow (y, x) \in r \Rightarrow (x, x) \in r$. So, an equivalence relation is *reflexive* in addition to being *symmetric* and *transitive*, which is what the definition says. A particular equivalence relation is the *identity*, which is defined as

$$(x, y) \in \mathrm{id} \quad :\Longleftrightarrow \quad x = y\ . \tag{4}$$

The predicate of being a *function* is defined as

$$\mathbf{Fun}(f) \quad :\Longleftrightarrow \quad \mathbf{Rel}(f)\ \wedge\ \forall\, x, y, y'\ \big((x, y) \in f \wedge (x, y') \in f\ \Rightarrow\ y = y'\big)\ . \tag{5}$$

Let f be a function and $(x, y) \in f$, then one also writes $f(x)$ for y. Any relation r defines a domain $\mathrm{dom}(r)$ and a range $\mathrm{ran}(r)$ by

$$\mathrm{dom}(r) := \{x|\ \exists\, y\,((x, y) \in r)\}\ , \quad \mathrm{ran}(r) := \{y|\ \exists\, x\,((x, y) \in r)\}\ . \tag{6}$$

A co-domain Y as it appears in the definition of *maps* $f : \mathrm{dom}(f) \to Y$ is here not part of the definition of a function. Here the most important feature of relations will be the non-commutative, though associative, product operation

$$r \circ s := \{(x, z)|\ \exists\, y\ ((x, y) \in s \wedge (y, z) \in r)\}\ . \tag{7}$$

The order of r and s on the right-hand side is chosen such that for functions one gets the conventional definition of composition. For any relation r the *reverse* relation is defined by

$$r^{-1} := \{(y, x)|(x, y) \in r\}\ , \tag{8}$$

for which we obviously have $(r^{-1})^{-1} = r$ and $(r \circ s)^{-1} = s^{-1} \circ r^{-1}$. Let r, s, r', s' relations such that $r \subset {}^{3}r'$ and $s \subset s'$, then $r \circ s \subset r' \circ s'$. These operations of *relation algebra* are the key to an elegant characterization of relation properties:

[3] this includes the case $r = r'$

Theorem 1 *Let r be a relation. Then*

$$\mathbf{Fun}(r) \iff r \circ r^{-1} \subset \mathrm{id}\ , \quad \mathbf{Equ}(r) \iff r^{-1} \subset r \ \wedge\ r \circ r \subset r\ .$$

Proof: Straightforward, [4] (14.2), (17.2), (17.5). ■

Theorem 2 *For any function f the relation $r := f^{-1} \circ f$ is an equivalence relation. It satisfies: $(x, x') \in r \ \Leftrightarrow\ f(x) = f(x')$.*

Proof: $(x,z) \in r \ \Leftrightarrow\ \exists\, y\left((x,y) \in f \wedge (y,z) \in f^{-1}\right) \ \Leftrightarrow\ \exists\, y\,((x,y) \in f \wedge (z,y) \in f) \ \Leftrightarrow$ $\exists\, y\,(y = f(x) \wedge y = f(z)) \ \Leftrightarrow\ f(x) = f(z)$. From this form of r, the properties of an equivalence relation follow directly from the corresponding properties of the equality relation. ■

The converse of the Theorem is not true, however. For some relations r, the relation $r^{-1} \circ r$ may be an equivalence relation although r is not a function. A simple example is $r := \{(1,1),(1,-1),(-1,1),(-1,1)\}$ for which $r^{-1} \circ r = r$ and, since $(1,1) \in r \wedge (1,-1) \in r$, r is not a function.

The equivalence relations associated with functions indicate whether their quotient relation is again a function:

Theorem 3 *Let f and g be functions such that* $\mathrm{dom}(f) = \mathrm{dom}(g)$ *and let the relation h be defined by $h := g \circ f^{-1}$. Then*

$$f^{-1} \circ f \ \subset\ g^{-1} \circ g \quad \Longleftrightarrow \quad \mathbf{Fun}(h) \ \wedge\ g = h \circ f\ .$$

Proof: $\Leftarrow$: $(x,y) \in f^{-1} \circ f \Leftrightarrow f(x) = f(y) \Rightarrow g(x) = h(f(x)) = h(f(y)) = g(y) \Rightarrow (x,y) \in g^{-1} \circ g$. Hence $f^{-1} \circ f \subset g^{-1} \circ g$. $\Rightarrow$: $h \circ h^{-1} = g \circ f^{-1} \circ f \circ g^{-1} \subset g \circ g^{-1} \circ g \circ g^{-1} \subset \mathrm{id} \circ \mathrm{id} = \mathrm{id}$. Hence, by the first part of Theorem 1, h is a function. Now consider the function $h \circ f : h \circ f = g \circ f^{-1} \circ f \subset g \circ g^{-1} \circ g \subset \mathrm{id} \circ g = g$, thus $h \circ f \subset g$. Since both functions have the same domain, namely $\mathrm{dom}(f)$, they are equal. ■

In a similar way in which a functions determines function values $f(x)$ for any $x \in \mathrm{dom}(f)$, a general relation r defines function values $r[a]$ for any set a

$$r[a] := \{y |\ \exists\, x\ (x \in a \wedge (x,y) \in r)\}\ . \tag{9}$$

For $a \cap \mathrm{dom}(r) = \emptyset$ one thus has $r[a] = \emptyset$. Of course,

$$\mathbf{Fun}(f) \ \wedge\ x \in \mathrm{dom}(f) \quad \Longrightarrow \quad \{f(x)\} = f[\{x\}]\ . \tag{10}$$

Lemma 1 *Let r be a relation and ϱ an equivalence relation. Then*

$$(x,y) \in r \circ \varrho \ \wedge\ (x,x') \in \varrho \quad \Longrightarrow \quad (x',y) \in r \circ \varrho\ .$$

Proof: $(x,y) \in r \circ \varrho \Rightarrow (x,y') \in \varrho \wedge (y',y) \in r$ for some y'. Since ϱ is an equivalence relation, also $(x',y') \in \varrho \wedge (y',y) \in r$ and, therefore, $(x',y) \in r \circ \varrho$. ■

Theorem 4 *Let g be a function and ϱ an equivalence relation such that* $\mathrm{dom}(g) = \mathrm{dom}(\varrho)$. *Then*

$$\varrho \subset g^{-1} \circ g \quad \Longleftrightarrow \quad g = g \circ \varrho \,.$$

Proof: $\Rightarrow$: $(x,y) \in \varrho \Rightarrow (x,y) \in g^{-1} \circ g \Leftrightarrow g(x) = g(y)$. This allows us to compute $g \circ \varrho$. Let $x \in \mathrm{dom}(\varrho)$ be arbitrarily selected. Then the elements x' of $\varrho[\{x\}]$ satisfy $(x,x') \in \varrho$ and, according to our previous result, $g(x) = g(x')$. Therefore $g[\varrho[\{x\}]] = \{g(x)\}$ and $(g \circ \varrho)[\{x\}] = \{g(x)\}$. Since $g[\varrho[\{x\}]]$ has a single element, $g \circ \varrho$ is a function, and since this single element equals $g(x)$, this function equals g. $\Leftarrow$: according to Lemma 1 function g is constant on equivalence classes of ϱ. This is exactly what $\varrho \subset g^{-1} \circ g$ says. ■

Corollary 1 *Let f and g be functions such that* $\mathrm{dom}(f) = \mathrm{dom}(g)$. *Then*

$$f^{-1} \circ f \subset g^{-1} \circ g \quad \Longleftrightarrow \quad g = g \circ (f^{-1} \circ f) \,.$$

Proof: Apply Theorem 4 to the case that ϱ is of the form $f^{-1} \circ f$. ■

The set of all the physical quantities under consideration , is a finite set F of functions on $\mathcal{X}$. Thus

$$f \in F \quad \Longrightarrow \quad \mathbf{Fun}(f) \wedge \mathrm{dom}(f) = \mathcal{X} \tag{11}$$

For any subset $G \subset F$ we combine the finitely many functions $g \in G$ into a single one, denoted by $\mathcal{F}(G)$ in a natural way:

$$\mathcal{F}(G) := \{(a,b) |\, \exists\, x, g\, (x \in \mathcal{X} \wedge g \in G \wedge a = (x,g) \wedge b = g(x))\} \tag{12}$$

One should note here that defining a function simply as a set of ordered pairs is cleaner and simpler as the usage of mappings, where one would need Cartesian products to define the co-domains. As any function, it defines an equivalence relation on its domain, which we introduce here together with an abbreviating name

$$\mathcal{R}(G) := \mathcal{F}(G)^{-1} \circ \mathcal{F}(G) \,. \tag{13}$$

Of course,

$$(x,x') \in \mathcal{R}(G) \quad \Longleftrightarrow \quad \forall g \in G \left(g(x) = g(x')\right) \,. \tag{14}$$

Let us now study how the natural partial order of equivalence relations and of subsets fits together. First the trivial consequence of the subset relation of the function sets

$$G \subset G' \quad \Longrightarrow \quad \mathcal{R}(G') \subset \mathcal{R}(G) \tag{15}$$

and the less trivial consequence of the subset relation of equivalence relations: Consider $G, G' \subset F$ and assume that

$$\mathcal{R}(G) \subset \mathcal{R}(G') \tag{16}$$

then, by Theorem 3, the relation $\varphi := \mathcal{F}(G') \circ \mathcal{F}(G)^{-1}$ is a function and we have

$$\mathcal{F}(G') = \varphi \circ \mathcal{F}(G) \,, \tag{17}$$

and by Corollary 1 we have

$$\mathcal{F}(G') \;=\; \mathcal{F}(G') \circ \mathcal{R}(G)\;. \tag{18}$$

This shows that the equivalence relations introduced here capture the concept of functional dependence between the functions collected in F: Consider two disjoint subsets G, G' of F. From (15) we have

$$\mathcal{R}(G \cup G') \subset \mathcal{R}(G) \tag{19}$$

and

$$\mathcal{R}(G \cup G') \subset \mathcal{R}(G')\;. \tag{20}$$

Lets see what it means when we have equality in (19). Then we have due to (20)

$$\mathcal{R}(G \cup G') = \mathcal{R}(G) \;\Rightarrow\; \mathcal{R}(G') \supset \mathcal{R}(G) \tag{21}$$

and thus (16) and the consequences (17) and (18).

This means that the functions $f' \in G'$ [4] can be expressed by a function term in functions $f \in G$:

$$\forall\, x \in \mathcal{X} \quad \big(f'(x)\big)_{f' \in G'} = \varphi\big((f(x))_{f \in G}\big)\;. \tag{22}$$

Normally this kind of expressions is derived from criteria concerning non-vanishing Jacobians. Here, however, we did not even assume differentiability of the functions $f \in F$.

In physics texts on thermodynamics we may read equations like

$$S = S(U, V, n) \tag{23}$$

for physical quantities S, U, V, and n.

In the syntax accepted today in mathematics this is nonsense: A function (which S has to be in order that $S(U, V, n)$ makes sense) cannot equal a function value (which $S(U, V, n)$ is). In the syntax of physics and the syntax of mathematical logic this is perfectly right. Here, S is understood here as a term and U, V, n as variables (which are terms too). Equation (23) is simply understood as an abbreviation of the statement

$$\mathrm{free}(S) = \mathrm{free}(S) \cup \{U, V, n\}\;, \tag{24}$$

where $\mathrm{free}(T)$ is the set of free variables of the term T. Of course, one could write (24) with less formal similarity to (23) as $\{U, V, n\} \subset \mathrm{free}(S)$.

With the function formalism developed above there are some equivalent ways to formulate what (23) is intended to express. For the function sets $F = \{S, U, V, n\}$, $G = \{U, V, n\}$, $G' = \{S\}$ (22) says

$$S(x) = \varphi(U(x), V(x), n(x)) \tag{25}$$

[4] the apostrophe to f is a diacritical mark and does not denote the derivative

We have prepared three formulations of the premises that lead to this conclusion: From (16)

$$\mathcal{R}(S) \supset \mathcal{R}(U,V,n) \ , \tag{26}$$

and from the first equation of (21)

$$\mathcal{R}(S,U,V,n) = \mathcal{R}(U,V,n) \ , \tag{27}$$

and by (18)

$$S = S \circ \mathcal{R}(U,V,n) \ . \tag{28}$$

The last of these formulations shows the highest degree of formal resemblance to (23). The meaning is in all three cases: Let x and x' be elements of $\mathcal{X}$ such that $U(x) = U(x')$, $V(x) = V(x')$, and $n(x) = n(x')$. Then one has $S(x) = S(x')$. Notice that if $\mathcal{X}$ would be a manifold and (U,V,n) system of coordinates (a global chart) for it, then one would conclude $x = x'$ and $S(x) = S(x')$ a fortiori. For the function φ in (25) one has

$$\varphi = S \circ \mathcal{F}(U,V,n)^{-1} \ . \tag{29}$$

In this representation the name S appears, which resembles the physics (and engineering) naming convention according to which the nature as a physical quantity determines the name, and not the argument structure of a representing function.

Equation (25) says that in any relevant situation, the value of S allows a functional representation in terms of the values of the quantities U, V, and n. In a treatise written in normal physical style, where situations are treated implicitly (without being associated with symbols), this is interpreted simply as 'S is a function of U, V, and n' and (29) allows us to express this function by well-defined operations from the 'coordinate-free' state functions S, U, V, n.

References

[1] Loomis, Sternberg: Advanced Calculus, Addison-Wesley 1968

[2] H. B. Callen: Thermodynamics, Wiley, 1963

[3] "Approach of a System of Particles towards Thermal Equilibrium" from The Wolfram Demonstrations Project.
`http://demonstrations.wolfram.com/ApproachOfASystemOfParticlesTowardsThermalEquilibrium/`
Contributed by: Ulrich Mutze and Stephan Leibbrandt (published 2012-03-08)

[4] Jürgen Schmidt: Mengenlehre I, Bibliographisches Institut, BI 56/56a, 1966