# eCOMMONS

**Loyola University Chicago**
**Loyola eCommons**

Physics: Faculty Publications & Other Works

Faculty Publications

4-2013

# Unintended Consequences of Imprecise Notation – an Example from Mechanics

Asim Gangopadhyaya
*Loyola University Chicago*, agangop@luc.edu

Gordon Ramsey
*Loyola University Chicago*, gramsey@luc.edu

## Recommended Citation

# Unintended consequences of imprecise Notation
# - an example from mechanics

## Asim Gangopadhyaya[1] and Gordon Ramsey[2]

*Department of Physics, Loyola University Chicago,*

*1032 W. Sheridan Rd., Chicago, IL 60660, U.S.A.*

**Abstract**

We present a conundrum that results from the imprecise use of notation for partial derivatives. Taking an example from mechanics, we show that lack of proper care in representing partial derivatives in Lagrangian and Hamiltonian formulations paradoxically leads to two different values for the time derivative of the canonical momentum. This problem also exists in other areas of physics, such as thermodynamics.

**Key words: Partial Differentiation, Classical Mechanics; Lagrangian Formulation; Hamiltonian Formulation**

**PACS: 45.20.Jj, 02.30Xx**

We generally insist that our students speak and write with precision, but they often do not. One case in point is that of partial differentiation, an operation that has a ubiquitous presence in many areas of physics, especially in areas of thermodynamics and canonical transformations. When we represent the partial differentiation of a function $\phi(x_1, x_2, \cdots, x_n)$ of several independent variables with respect to one of those variables $x_k$ by $\frac{\partial \phi}{\partial x_k}$, we often do not underscore the constraints in this variation; i.e., the variables that are kept constant while $x_k$ changes. As the following example from classical mechanics shows, this omission can frequently lead to serious misunderstanding.

The comparative study of the Lagrangian and Hamiltonian approaches to classical mechanics is a standard topic in undergraduate courses in mechanics. The crux of the difference between the two approaches is that they have

---

[1]agangop@luc.edu

[2]gramsey@luc.edu

two different sets of independent variables. This aspect is not always sufficiently appreciated by students. Here we present a simple paradox that challenges the reader to engage deeply, and thus develop a clearer perspective on each formalism, and their differences. Paradoxes provide many benefits in instruction, namely to motivate deeper thinking and provide a more thorough understanding of the topic. As pointed out by Welch [1] in this journal, their greatest benefit is to address "specific deficiencies in understanding and reasoning." We shall see that the resolution of our paradox lies in the careful handling of partial differentiation. Hence, one of our aims is to urge students to be cognizant of the roles played by various variables in the variation of a function that depends on more than one physical quantity.

In undergraduate classical mechanics courses, students are generally encouraged to express the Lagrangian, $\mathcal{L}$ and Hamiltonian, $\mathcal{H}$ in various coordinate systems. This not only helps them build very important expertise in choosing an efficient set of generalized coordinates for a system, often using symmetries, it also helps them explore how $\mathcal{L}$ and $\mathcal{H}$ could look very different in one coordinate system compared to another, and yet carry exactly the same physical information.

Let us consider a Lagrangian of a free system; i.e., the potential $V(q) = 0$. The Lagrangian is given by

$$\mathcal{L}(q, \dot{q}) = T(q, \dot{q}) - V(q) = T(q, \dot{q}) \tag{1}$$

The canonical momentum is defined by

$$p \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}} \ . \tag{2}$$

Then, the Lagrange's equation of motion $\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial \mathcal{L}}{\partial q}$ implies

$$\dot{p} = \frac{\partial T}{\partial q} \ . \tag{3}$$

Let us do a similar analysis using Hamiltonian formulation. Assuming the kinetic energy to be a homogeneous quadratic function of velocities, and since the potential $V$ is identically zero, the Hamiltonian is $T + V = T$ [2, 3]. Hamilton's equation of motion then implies,

$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial q} = -\frac{\partial T}{\partial q}. \tag{4}$$

Thus, the conundrum is that $\dot{p}$ is given by $\frac{\partial T}{\partial q}$ in the Lagrangian formulation, and by $-\frac{\partial T}{\partial q}$ in the Hamiltonian

formalism; i.e.,

$$\dot{p} = \left.\frac{\partial T}{\partial q}\right|_{\text{Lagrangian Formulation}} = -\left.\frac{\partial T}{\partial q}\right|_{\text{Hamiltonian Formulation}} . \tag{5}$$

This suggests that $\dot{p} = 0$, as one would expect in a case of a free particle. Is this necessarily true? For example, a particle described by cylindrical coordinates $(r, \phi, z)$, the kinetic energy is given by $T = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 + \frac{1}{2}m\dot{z}^2$, the derivative $\frac{\partial T}{\partial r} \neq 0$, and hence $\dot{p}_r \neq 0$. Thus, for this case, it becomes difficult to explain how a non-zero $\dot{p}_r$ be equal to $\frac{\partial T}{\partial r}$ in Lagrangian formulation and $-\frac{\partial T}{\partial r}$ in Hamiltonan formulation.

Similarly, for a particle described by spherical coordinates $(r, \theta, \phi)$, the kinetic energy is given by $T = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}mr^2\sin^2\theta \ \dot{\phi}^2$, neither the $\dot{p}_r$, nor the $\dot{p}_\theta$ are equal to zero. So, how do we explain that $\dot{p}_\theta$ is equal to $\frac{\partial T}{\partial \theta}$ in the Lagrangian formulation and $-\frac{\partial T}{\partial \theta}$ in the Hamiltonian approach, and yet $\dot{p}_\theta \neq 0$? As we will show later, the solution for this apparent problem lies in the fact that $\frac{\partial T}{\partial q_i}$ calculated in two different formulations are different. It is crucial to know which variables are being kept constant in these partial differentiations. For the Lagrangian formulation, the derivative $\frac{\partial T}{\partial q_i}$ implies that $T$ is being differentiated, keeping coordinates $q_j$ with $j \neq i$ constant, and all $\dot{q}_j$'s are held constant. For the Hamiltonian formulation, the derivative $\frac{\partial T}{\partial q_i}$ implies that $T$ is being differentiated keeping coordinates $q_j$ with $j \neq i$ constant, and all $p_j$'s are held constant. Since the constraints under which these two derivatives are being taken are different, the derivatives themselves may be exactly equal in magnitude and opposite in sign, and hence may make Eq. (5) consistent with a non-zero $\dot{p}$.

Let us check this possibility with a free particle in spherical coordinates. The Lagrangian and the Hamiltonian of the system are given by $\mathcal{L} = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}mr^2\sin^2\theta \ \dot{\phi}^2$ and $\mathcal{H} = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2\sin^2\theta}$ respectively [?]. From Lagrange's equation, we get

$$\dot{p}_\theta = \left.\frac{\partial T}{\partial \theta}\right|_{\{r,\phi,\dot{r},\dot{\theta},\dot{\phi}\}} = mr^2\sin\theta \ \cos\theta \ \dot{\phi}^2 . \tag{6}$$

From Hamilton's method, we get

$$\dot{\phi} = \frac{p_\phi}{mr^2\sin^2\theta} \ ; \ \text{and} \ \dot{p}_\theta = -\left.\frac{\partial T}{\partial \theta}\right|_{\{r,\phi,p_r,p_\theta,p_\phi\}} = \frac{p_\phi^2\cos\theta}{mr^2\sin^3\theta}. \tag{7}$$

Substituting for $p_\phi$ in Eq. (7), we see that $\dot{p}_\theta$ is indeed equal to $mr^2\sin\theta \ \cos\theta \ \dot{\phi}^2$ in both cases.

Thus, in spherical coordinates, Eq. (5) does not lead to any inconsistency for $\dot{p}_\theta$. We now show that this is indeed true in all coordinate systems, provided that the kinetic energy is a homogeneous quadratic function of the generalized velocities.

Let us start with a system with zero potential [3] and with $N$-degrees of freedom, described by the Lagrangian

$$\mathcal{L}(q_i, \dot{q}_i, t) = \frac{1}{2} \sum_{ij}^{N} F_{ij}(q) \dot{q}_i \dot{q}_j \, , \tag{8}$$

where the <u>symmetric</u> coefficients $F_{ij}$ are, in general, functions of $N$-independent coordinates $q_i$. From here on, we will suppress the $q$-dependence of $F_{ij}$. The canonical momenta $p_i$ are given by

$$p_i = \frac{\partial \mathcal{L}(q_i, \dot{q}_i, t)}{\partial \dot{q}_i} = \sum_j F_{ij} \, \dot{q}_j \equiv F_{ij} \, \dot{q}_j, \tag{9}$$

where summation over repeated indices is assumed. Since coordinates $q_i$ are assumed to be independent, the inverse of the matrix $F_{ij}$ exists, and allows us to invert Eq. (9) to find velocities in terms of momenta. We denote the elements of this inverse matrix by $\left(F^{-1}\right)_{ij}$. From Eq. (9), solving for the velocities, we get

$$\dot{q}_i = \left(F^{-1}\right)_{ij} \, p_j \, . \tag{10}$$

Since the Hamiltonian must be written as a function of $q_i, p_i$ and $t$, the Hamiltonian for the system is then given by

$$\mathcal{H}(q_i, p_i, t) = p_i \dot{q}_i - \mathcal{L} = p_i \dot{q}_i - \frac{1}{2} F_{ij} \dot{q}_i \dot{q}_j \tag{11}$$

$$= p_i \left(F^{-1}\right)_{ij} \, p_j - \frac{1}{2} F_{ij} \left(\left(F^{-1}\right)_{ik} \, p_k\right) \left(\left(F^{-1}\right)_{j\ell} \, p_\ell\right) \, , \tag{12}$$

where we have substituted velocities from Eq. (10) into Eq. (11). Now, using the associativity of matrix multiplication, we get

$$\mathcal{H}(q_i, p_i, t) = p_i \left(F^{-1}\right)_{ij} \, p_j - \frac{1}{2} \underbrace{\left(F_{ij} \left(F^{-1}\right)_{ik}\right)}_{\delta_{jk}} p_k \left(\left(F^{-1}\right)_{j\ell} \, p_\ell\right)$$

$$= p_i \left(F^{-1}\right)_{ij} \, p_j - \frac{1}{2} p_k \left(\left(F^{-1}\right)_{k\ell} \, p_\ell\right)$$

$$= \frac{1}{2} p_i \left(F^{-1}\right)_{ij} \, p_j \, , \tag{13}$$

where we have relabeled the dummy indices in the second term of the second line. Since the matrix $F_{ij}$ is a function of $q_i$'s, the Hamiltonian has been properly represented in terms of $q_i$, $p_i$ and t. Now, let us compute $\dot{p}_i$ for this general case using the two different formalisms. In the Lagrangian formalism, from Eq. (8), we have

$$\dot{p}_i = \frac{\partial \mathcal{L}}{\partial q_i} = \frac{\partial T}{\partial q_i} = \frac{1}{2} \left[\frac{\partial}{\partial q_i} F_{jk}\right] \dot{q}_j \dot{q}_k \, . \tag{14}$$

---

[3]It is worth noting that this discrepancy in the values for $\dot{p}$ returned by the two formalisms is present even for the case with a non-zero potential $V(r, \theta, \phi)$; it just appears more dramatic when the potential is zero.

On the other hand, from Eq. (13), the Hamiltonian formalism gives

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} = -\frac{\partial T}{\partial q_i} = -\frac{1}{2} \left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{\ell m} \right) p_\ell \, p_m \tag{15}$$

Now, substituting for momenta from Eq. (9), we get

$$
\begin{aligned}
\dot{p}_i &= -\frac{\partial T}{\partial q_i} = -\frac{1}{2} \left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{\ell m} \right) \left( F_{\ell j} \, \dot{q}_j \right) \left( F_{mk} \, \dot{q}_k \right) \\
&= \frac{1}{2} \left[ -F_{j\ell} \left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{\ell m} \right) F_{mk} \right] \dot{q}_j \, \dot{q}_k \ .
\end{aligned} \tag{16}
$$

In Eqs. (14) and (16), we have computed the time derivatives of momentum $p_i$ from Lagrangian and Hamiltonian formalisms respectively. Consistency of these two expressions requires that factors within square brackets of these two equations be equal; i.e.,

$$\left[ \frac{\partial}{\partial q_i} F_{jk} \right] = \left[ -F_{j\ell} \left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{\ell m} \right) F_{mk} \right] \ . \tag{17}$$

Multiplying the above equation by $\left( F^{-1} \right)_{sj}$ and summing over the index $j$, we get

$$
\begin{aligned}
\left( F^{-1} \right)_{sj} \left[ \frac{\partial}{\partial q_i} F_{jk} \right] &= \left( F^{-1} \right)_{sj} \left[ -F_{j\ell} \left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{\ell m} \right) F_{mk} \right] \\
&= -\underbrace{\left( F^{-1} \right)_{sj} F_{j\ell}}_{\delta_{s\ell}} \left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{\ell m} \right) F_{mk} \\
&= -\left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{sm} \right) F_{mk} \\
&= -\left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{sj} \right) F_{jk} \ .
\end{aligned} \tag{18}
$$

Thus, consistency of the two results requires

$$\left( F^{-1} \right)_{sj} \left[ \frac{\partial}{\partial q_i} F_{jk} \right] + \left( \frac{\partial}{\partial q_i} \left( F^{-1} \right)_{sj} \right) F_{jk} = 0 \ . \tag{19}$$

However, the left-hand side of Eq. (19) is a perfect differential that can be written as

$$\frac{\partial}{\partial q_i} \left( \left( F^{-1} \right)_{sj} F_{jk} \right) = \frac{\partial}{\partial q_i} \delta_{sk} \ , \tag{20}$$

which is identically zero. Thus, for a system where the kinetic energy is a homogeneous quadratic function of the velocities, the two seemingly different expressions for $\dot{p}_i$, one derived from the Lagrangian and the other from the Hamiltonian formulation, are thus proved to be identical.

**Conclusion:** In this note, we present a conundrum that appears in coordinate systems where the kinetic energy is a quadratic homogeneous function of velocities and a function of generalized coordinates. The conundrum is most clear in the absence of an interaction, as both the Lagrangian and the Hamiltonian are equal to the kinetic energy, and the expressions for $\dot{p}_i$ in two formalisms are $\frac{\partial T}{\partial q_i}$ and $-\frac{\partial T}{\partial q_i}$ respectively. The purpose of this note is to show that if $\mathcal{L}$ and $\mathcal{H}$ are expressed correctly in their respective sets of independent variables, and the differentiations are done carefully, then there is no discrepancy. This type of puzzle goes a long way toward instilling a deeper understanding of the partial differentiations involved in these two different formulations of classical mechanics. We would like to stress that this problem is not localized to the area of mechanics alone. There are many examples in other fields as well. For example, in thermodynamics, the molar specific heat of a gas $\frac{\Delta Q}{\Delta T}$ could refer to either $C_p$ or $C_v$ depending on whether pressure or the volume is kept constant during the variation of temperature. Similarly, as pointed out by one of the referees, for an ideal gas we have $dU = \frac{3}{2} n\, R\, dT = \frac{3}{2} p\, dV$. Thus, while $\left.\frac{\partial U}{\partial V}\right|_p = \frac{3}{2} p$, $\left.\frac{\partial U}{\partial V}\right|_T = 0$. These examples drawn from different areas of physics clearly show that for partial differentiation, we need to pay great deal of attention to the variables that are constrained to a constant value.

# References

[1] Daniel W. Welch, *Using paradoxes*, Am. J. Phys., **48**, 629 (1980).

[2] Stephen Thornton and Jerry B. Marion, *Classical Dynamics of Particles and Systems* , (5th edition), Brooks/Cole, United States, section 7.8, p 258.

[3] Herbert Goldstein, Charles P. Poole and John L. Safko, *Classical Mechanics* (3rd edition), Addison Wesley.