Euler-Lagrange Equation with SD

P. Reany

June 24, 2025

Abstract

Structured Differentiation (SD) is used to present the treatment of the Euler-Lagrange equation and the Beltrami Identity, with updates from articles by Brownstein and others after him.[1]

1 Introduction

The original point of the development of SD was first to clarify what we mean by total vs partial derivative, and second, to justify when operating across a true equation with a derivative, the result is another true equation. This paper develops the Euler-Lagrange Equation from the perspective of those criteria.

This presentation uses SD to develop the Euler-Lagrange equation to describe the motion of a point particle in a conservative system and in one spacial dimension q. Within any inertial reference system we can separate functional dependence into causes that are (explicitly) time dependent at fixed spacial points, and those that are dependent on motion within the reference system. For example, consider the function f given by

$$f = f(t, \mathbf{x}(t))\,, \tag{1}$$

where $\mathbf{x}$ is a vector-valued function of t. Taking the total derivative of f by t we get

$$\frac{df}{dt} = \frac{\partial f}{\partial t}\frac{dt}{dt} + \frac{\partial f}{\partial \mathbf{x}}\frac{d\mathbf{x}}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{x}}\frac{d\mathbf{x}}{dt}\,. \tag{2}$$

This last equation serves as an example of using the parametric split, which facilites using the chain rule. The parametric split serves to split the total derivative into an explicit and implicit part — the latter often having more than one term. Now there are two ways to form equations of interest to us here. The first is to introduce a new variable as a replacement for an expression of other variables or for a function alone. This is precisely what we did in (1), where the variable f was set equal to the function $f(t, \mathbf{x}(t))$. We will refer to this kind of equation as an identity.

[1] This article was originally published in January 2000.

The second way is to constrain a function or an expression of functions equal to some fixed quantity, which reduces the maximum number of fundamental (independent) variables in the equation. In the first case, we may apply any derivative we like across the equality and still obtain a true equation as a result. But in the second case, unless we know the specific dependence relation that the function(s) have on the variable of differentiation, we must use the total derivative operator to ensure that the resulting equation will be true. It is generally not true that the explicit derivative, say, of an true equation which constrains a function or an expression is itself a true equation. The most important rule to remember is that we may apply any total derivative to any true type 1 or type 2 equation and get a true type 2 equation. Naturally, all type 1 equations are true since they are definitions.

2 Euler-Lagrange Equation Developed

Now to mechanics and beyond: At any time our point particle has (total) energy E (in one-dimensional motion) which relates to the kinetic energy T given by

$$T = \frac{1}{2}m\dot{x}^2\,, \tag{3}$$

(a type 1 equation), and the potential energy V by

$$E = T + V\,, \tag{4}$$

where (4) is an equation of type 1, since E was merely defined in it. Now we derive from it a type 2 equation by using the fact that in a conservative system the total energy is a constant. From this we get that

$$E = \text{const} \tag{5}$$

and differentiating this (totally) by t we get

$$\frac{dE}{dt} = 0\,. \tag{6}$$

Next we introduce a new variable, L, called the *Lagrangian*, defined by

$$L = T - V\,, \tag{7}$$

which is introduced to place the resulting equations of motion in an alternative form than that of Newtonian mechanics. Note that (7) is a type 1 equation, so we may take any derivative we like across it to get a true equation as a result.

And, since we are dealing with a conservative force, the potential V is given by

$$V = V(x)\,. \tag{8}$$

and, with the help of (3), (7) becomes

$$L = \frac{1}{2}m\dot{x}^2 - V(x)\,. \tag{9}$$

Now for a result we'll need shortly. From the last result, we get that

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x}\,,$$

therefore,

$$\frac{\partial L}{\partial \dot{x}}\dot{x} = m\dot{x}^2 = 2T\,. \tag{10}$$

From Equations (4) and (7) we get the algebraic result that

$$E + L = 2T\,. \tag{11}$$

Totally differentiating (11) by time and using (10), and remembering that $dE/dt = 0$, we get

$$\frac{dL}{dt} = \frac{d(2T)}{dt}\,. \tag{12}$$

And with the help of (6), (12) expands to

$$\frac{dL}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\dot{x}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)\dot{x} + \frac{\partial L}{\partial \dot{x}}\ddot{x}\,. \tag{13}$$

But we can also expand dL/dt formally, using that

$$L = L(x, \dot{x}, t)\,. \tag{14}$$

to get

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \frac{\partial L}{\partial x}\dot{x} + \frac{\partial L}{\partial \dot{x}}\ddot{x}\,. \tag{15}$$

Now, the equation we seek is the result of equating the two expressions from (15) and (13), which we will do after we determine what to do with the term $\partial L/\partial t$, where the partial derivative is always an explicit derivative in SD.

So, according to convention, which SD agrees with on this point, the expression $\frac{1}{2}m\dot{x}^2$ is not explicitly dependent on time, so $\partial T/\partial t = 0$. Furthermore, since the system is conservative, we know that $\partial V/\partial t = 0$. And, since L in (7) is a definition then we can take any derivative we like on it and get back a true equation. Doing so with a partial derivative we get that

$$\frac{\partial L}{\partial t} = \frac{\partial T}{\partial t} - \frac{\partial V}{\partial t} = 0 - 0 = 0\,. \tag{16}$$

For the sake of accuracy, we must rewrite (14) to drop the explicit dependence of L on time t:

$$L = L(x, \dot{x})\,. \tag{17}$$

We are now ready to equate the two expressions (15) and (13), setting the term $\partial L/\partial t$ to zero, yielding

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)\dot{x} + \frac{\partial L}{\partial \dot{x}}\ddot{x} = \frac{\partial L}{\partial x}\dot{x} + \frac{\partial L}{\partial \dot{x}}\ddot{x}\,, \tag{18}$$

which, after cancellation and regrouping, gives

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x}\right]\dot{x} = 0\,. \tag{19}$$

So, when $\dot{x} \neq 0$, we have that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0\,, \tag{20}$$

which is the Euler-Lagrange equation.

Now, for the case when $\dot{x} = 0$ we have two subcases: 1) $\dot{x} = 0$ at isolated points in time, or 2) $\dot{x} = 0$ over continuous time intervals. In the former case we can claim that (20) still holds by using continuity arguments, and in the latter case, the equation of motion may be determined by elementary means, yielding $x(t) =$ constant.

Thus, except for trivial cases, (20) holds generally. In short, the Euler-Lagrange equation is merely the chain rule expansion of dL/dt constrained by a few physical constraints.

3 Euler-Lagrange Equation Developed Through the Principle of Least (Extremal) Action

I'm going to follow the development of this section consistent with that of Prof. V. Balakrishnan, Department of Physics, IIT Madras, India (as presented in a YouTube video of his class presentation), not because it's original, but because of the following quote (I attempt to present as accurately as I can) from his on-line lecture on the Lagrangian formalism:[2]

> I'm going to give you an answer that may not be very satisfactory, but that is the only answer that there is. This is not an axiomatic subject. We haven't reached a stage where some absolute proof is written at some level and everything is derived axiomatically from that. Physics does not work like that. It starts by simple examples and generalizes; and once you put a whole lot of things together and the same formula applies for all of them (instead of individual formulas), you take that as granted; and then you ask what are the limits of its applicability and generalize it — and further and further and further, and so on, and this is the way it goes. And that's how we found that all these equations of motion can be subsumed under one principle — the Principle of Extremal Action.

[2] This highly insightful explanation was given, apparently extemporaneously, in response to the question from one of his students about the authoritative status of the Principle of Extremal Action and the Lagrangian formalism.

> And so we believe it is true, until somebody comes along with a better alternative. So, believe me, this [the Lagrangian formalism] has been tested for several hundred years, and at this level it's completely established.
>
> It's just that the question is, Where are you going to write the Lagrangian from? Especially when you go to quantum mechanics? When you go to quantum fields, where are you going to write the Lagrangian down from? You don't have the experience of Newton's equations. You don't have this. You're not in the non-relativistic regime; you're not in the classical regime. Then, invariance principles play a big role. It turns out that Nature is guided to a large extent by symmetries and invariance principles, which [...] are hidden here. But I'll bring them out as we go along.
>
> You would write the Lagrangian down based on considerations of simplicity — simplest possible choices, subject to the invariances: one of them is the one he [one of the students in the class] has pointed out, that the Lagrangian must be a scalar. But *not* a scalar in the sense that you and I understand in *this* course, whereby by "scalar" we simply mean something that is not a vector; something doesn't change under rotations of the coordinate system. But by "scalar" I'll mean, not only that something does not change under rotation of the coordinate system, but also Lorentz transformations: shifts from one inertial frame to another. I will use what is called a "four-dimensional scalar," or a "Lorentz scalar" — a scalar under Special Relativity — that would guide me in writing down the equations of motion in more complicated situations.

So, my problem now is to motivate how to arrive at the Euler-Lagrange equations of motion by introducing an integral. Well, which equations of motion? For starters, Newton's equations of motion, $\mathbf{F} = m\mathbf{a}$. Let's look at an example. Suppose I toss a soft ball to a friend standing some meters away from me. Ignoring air resistance, what does Newton's Second Law say about the path of the ball from me to my friend? It claims it to be a parabola. Now I ask, of all possible pathways between the start and finish points, given fixed initial conditions, how does nature choose the parabola as the actual path? In the context of Newton's framework of solving this problem, all one can say is that the parabola comes out of solving Newton's Second Law of motion.

Now, consider a similar problem of finding the correct pathway of a light ray moving from point A in a medium whose index of refraction is n_1 to a point B in another medium whose index of refraction is $n_2 \neq n_1$. The two mediums touch each other along a flat, smooth surface of contact. We must ask how it is that out of the infinite number of possible pathways the light could travel between A and B, how does it consistently travel just one unique path, day after day, experiment after experiment? This question has been answered in terms of a minimization principle, stated this way: Light will travel from A to B so as to minimize its total time of flight. This minimization rule is known as *Fermat's*

Principle. Thus, in moving form point x_1 to point x_2, the total time can be expressed as the integral of an infinite number of infinitesimal time intervals by

$$T = \int_{x_1}^{x_2} dt = \int_{x_1}^{x_2} \frac{ds}{v} \,. \tag{21}$$

Naturally, we can ask if there is not a similar approach to classical Newtonian mechanics to find a minimization approach to finding the Euler-Lagrange equations? If so, then the reason the projectile moves on a parabola is because such a path, chosen out of all possible pathways, minimizes something.

The integrand of this integral will be referred to as the Lagrangian, and represented by the symbol L. We'll represent the integral itself by the symbol φ (varphi), to get

$$\varphi = \int_{t_1}^{t_2} L\, dt \,. \tag{22}$$

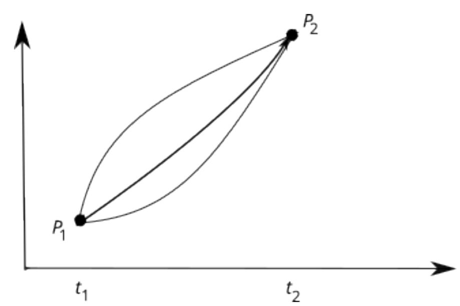

Figure 1. In Nature, the Action provides for a unique extremal path among all possible paths from any two distinct points in time.

This integral is referred to as the *Action*. So, how to get a differential equation out of the action? We begin by taking the differential of the action and setting it equal to zero, and then hope to find that the resulting integrand can be set equal to zero. So, let's get started.

$$\delta\varphi = \int_{t_1}^{t_2} \delta L\, dt = 0 \,. \tag{23}$$

We assume that the minimal pathway between points P_1 and P_2 exists and that this delta represents all possible arbitrary displacements from this minimal path, while maintaining the continuity of pathways and the fixity of the endpoints. So, what functional form should L have? How about

$$L = L(q, \dot{q}) \,, \tag{24}$$

(where we assume that q and $\dot{q}$ are mutually independent of each other)? Using this, we get

$$\delta L = \frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\,. \tag{25}$$

This last equation comes from the standard application of the chain rule in SD.[3] The partial derivatives are explicit, as always in the application of the chain rule, and will treat q and $\dot{q}$ as independent, even if they are not. Anyway, substituting (25) into (23), we get

$$\int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\right] dt = 0\,. \tag{27}$$

Comparing the integrand above to the known Euler-Lagrange equations, we must paradoxically factor out both δq and $\delta \dot{q}$, and somehow use the fact that the variation in path at the endpoints is zero. We can make a start to both goals by noting that, since the time derivative of q and the variation in pathway are independent of each other, then we can change their ordering, to get

$$\delta \dot{q} = \delta \frac{dq}{dt} = \delta \frac{d}{dt} q = \frac{d}{dt}\delta q\,. \tag{28}$$

Substituting this into (27), we get

$$\int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\frac{d}{dt}\delta q\right] dt = 0\,. \tag{29}$$

Now, to factor out the δq, all we have to do is to remove the derivative operator from it by contrafluxing the derivative[4], like this

$$\int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q}\delta q + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\delta q\right) - \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right)\delta q\right] dt = 0\,. \tag{30}$$

We can write this last equation as the sum of two integrals, as follows

$$\int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right] \delta q\, dt + \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\delta q\right) dt = 0\,. \tag{31}$$

By use of the Fundamental Theorem of Integral Calculus, the second integral above becomes

$$\int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\delta q\right) dt = \left(\frac{\partial L}{\partial \dot{q}}\delta q\right)\Bigg|_{t_1}^{t_2}\,. \tag{32}$$

[3]This is basically just a gradient of L through its variant dotted by a vector variation,

$$\delta L = [\partial L/\partial q, \partial L/\partial \dot{q}] \begin{bmatrix}\delta q \\ \delta \dot{q}\end{bmatrix}\,. \tag{26}$$

[4]What most people refer to as "integrating by parts."

But since, by construction, $\delta q(t_1) = \delta q(t_2) = 0$, this integral is zero, leaving us to conclude that

$$\int_{t_1}^{t_2} \left[\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} \right] \delta q \, dt = 0 \,. \tag{33}$$

And the only way to guarantee that this integral is identically zero is to set its integrand indentically to zero, yielding

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0 \,. \tag{34}$$

Or, algebraically changing the order of terms, we have

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \,. \tag{35}$$

Finally, the proof that the Euler-Lagrange equations are valid whenever Newton's equations of motion are valid. Let $L = T - V$, then,

$$L = \frac{1}{2} m \sum_i \dot{q}_i^2 - V(\mathbf{r}) \,, \tag{36}$$

and

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_k} - \frac{\partial L}{\partial x^k} = m\ddot{q}_k - \frac{\partial V}{\partial x_k} = 0 \,. \tag{37}$$

And, since $F_k = -\partial V/\partial x_k$, then we can write the final form as Newton's Second Law in component form

$$m\ddot{q}_k = F_k \,. \tag{38}$$

4 The Beltrami Identity

In the 19th century, Eugenio Beltrami discovered the following identity, a corollary to the Euler-Lagrange equation, which has been since named after him. Let L be the Lagrangian satisfying the Euler-Lagrange equation with $\partial L/\partial t = 0$, then

$$L - \dot{q}\frac{\partial L}{\partial \dot{q}} = c \,, \tag{39}$$

where c is a constant.

Proof:

We begin with the total derivative of L by t, to use later on.

$$\frac{dL}{dt} = \frac{\partial L}{\partial q}\dot{q} + \frac{\partial L}{\partial \dot{q}}\ddot{q} \,. \tag{40}$$

Now we take the Euler-Lagrange equation in the following form

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0 \,, \tag{41}$$

and multiply it through by $\dot{q}$:

$$\dot{q}\frac{\partial L}{\partial q} - \dot{q}\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0\,. \tag{42}$$

The next obvious step is to contraflux, to get

$$\dot{q}\frac{\partial L}{\partial q} - \left[\frac{d}{dt}\big(\dot{q}\frac{\partial L}{\partial \dot{q}}\big) - \ddot{q}\frac{\partial L}{\partial \dot{q}}\right] = 0\,. \tag{43a}$$

or

$$\dot{q}\frac{\partial L}{\partial q} - \frac{d}{dt}\big(\dot{q}\frac{\partial L}{\partial \dot{q}}\big) + \ddot{q}\frac{\partial L}{\partial \dot{q}} = 0\,. \tag{43b}$$

Now we use (40) in (43b) to get

$$\dot{q}\frac{\partial L}{\partial q} - \frac{d}{dt}\big(\dot{q}\frac{\partial L}{\partial \dot{q}}\big) + \big[\frac{dL}{dt} - \frac{\partial L}{\partial q}\dot{q}\big] = 0\,. \tag{44a}$$

After the obvious cancellation of terms, we get

$$\frac{d}{dt}[L - \dot{q}\frac{\partial L}{\partial \dot{q}}] = 0\,. \tag{44b}$$

And on integrating this, we get

$$L - \dot{q}\frac{\partial L}{\partial \dot{q}} = \text{const}\,. \tag{44c}$$

Practical uses of the Beltrami Identity include finding solutions to many problems that the Euler-Lagrange Equations could be used, such as the solution to the Brachistochrone problem.

5 Change of Coordinates

In this section we prove one of the most important aspects of the Lagrangian formalism, which is that we can choose the coordinate system at convenience to solve any given problem and the Euler-Lagrange equations will still apply. However, unlike most proofs I've seen of this theorem, we will here provide a rigorous justification of each step of the proof.[5,6]

Let's begin with (35) and generalize it to n coordinates, yielding

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0\,, \quad 1 \le i \le n\,, \tag{45}$$

where $L = L(q_i, \dot{q}_i, t)$. Now, we make the change of coordinates

$$q_i = q_i(Q_j, t)\,, \quad 1 \le i, j \le n\,, \tag{46}$$

[5] Well, as rigorous as I can make it at this point.

[6] For a reference, see online notes "Lagrange's and Hamilton's Equations" at http://www.physics.rutgers.edu/ shapiro/507/book3.pdf by Prof. Joel Shapiro.

where we assume that the function q_i are of order C^2, so that we may change the order of differentiation as needed. We also assume that the Q_j's are C^2 functions of the q_i's and time. First we prove a needed unintuitive lemma.

Lemma:

$$\frac{\partial \dot{q}_i}{\partial \dot{Q}_j} = \frac{\partial q_i}{\partial Q_j}\,, \quad 1 \le i, j \le n\,, \tag{47}$$

Proof: First, we differentiate (46) totally by t to get

$$\frac{dq_i}{dt} = \dot{q}_i = \frac{\partial q_i}{\partial Q_j}\dot{Q}_j + \frac{\partial q_i}{\partial t}\,, \quad 1 \le i \le n\,, \tag{48}$$

where we use the Einstein convention of summing on repeated indices in a given expression, unless stated to the contrary.

But how do the q_i's depend on the Q_j's and on time? This way:

$$\dot{q}_i = \dot{q}_i(Q_j, \dot{Q}_j, t)\,, \quad 1 \le i, j \le n\,, \tag{49}$$

where we assume that the new positions and velocity coordinates are independent of each other, as we did with the old position and velocity coordinates, i.e., the new system vs the old system.[7] Therefore, for the upcoming differentiation, we need to use that

$$\frac{\delta \dot{q}_i}{\delta \dot{Q}_j} = \frac{\partial \dot{q}_i}{\partial \dot{Q}_j}\,, \quad 1 \le i, j \le n\,, \tag{50}$$

since there is no implicit dependence of $\dot{q}_i$ on $\dot{Q}_j$ through q_i. Now we can totally differentiate (48) by $\dot{Q}_k$ to get[8]

$$\frac{\partial \dot{q}_i}{\partial \dot{Q}_k} = \frac{\delta}{\delta \dot{Q}_k}\Big[\frac{\partial q_i}{\partial Q_j}\dot{Q}_j + \frac{\partial q_i}{\partial t}\Big]\,, \quad 1 \le i, j, k \le n\,. \tag{51}$$

Now, the RHS of this last equation can be written as

$$\Big[\frac{\delta}{\delta \dot{Q}_k}\frac{\partial q_i}{\partial Q_j}\Big]\dot{Q}_j + \frac{\partial q_i}{\partial Q_j}\frac{\delta \dot{Q}_j}{\delta \dot{Q}_k} + \frac{\delta}{\delta \dot{Q}_k}\frac{\partial q_i}{\partial t}\,, \quad 1 \le i, j, k \le n\,, \tag{52}$$

On remembering that velocities in the new coordinates are independent of each other, we write

$$\frac{\delta \dot{Q}_j}{\delta \dot{Q}_k} = \delta_{jk}\,, \quad 1 \le j, k \le n\,, \tag{53}$$

where δ_{jk} is the Kronecker delta, which is zero when the indices are not equal, and unity when they are equal. Thus, (52) becomes

$$\Big[\frac{\delta}{\delta \dot{Q}_k}\frac{\partial q_i}{\partial Q_j}\Big]\dot{Q}_j + \frac{\partial q_i}{\partial Q_k} + \frac{\delta}{\delta \dot{Q}_k}\frac{\partial q_i}{\partial t}\,, \quad 1 \le i, j, k \le n\,. \tag{54}$$

[7]Equations (46) and (49) together are apparently referred to as a *point transformation.*

[8]We need to introduce a new index letter or we won't get the correct result.

And, because we can change the order of differentiation, we get

$$\Big[\frac{\partial}{\partial Q_j}\frac{\delta q_i}{\delta \dot{Q}_k}\Big]\dot{Q}_j + \frac{\partial q_i}{\partial Q_k} + \frac{\partial}{\partial t}\frac{\delta q_i}{\delta \dot{Q}_k}\,, \quad 1 \le i, j, k \le n\,. \tag{55}$$

However, the position coordinates in the original system are not functions of the velocities in the second system, hence $\dfrac{\delta q_i}{\delta \dot{Q}_k} = 0$ for all i, k. Thus, (51) reduces to

$$\frac{\partial \dot{q}_i}{\partial \dot{Q}_k} = \frac{\partial q_i}{\partial Q_k}\,, \quad 1 \le i, k \le n\,, \tag{56}$$

which is what we were to show.

Now, let the Lagrangian in the new system be $\mathcal{L} = \mathcal{L}(Q_i, \dot{Q}_i, t)$, given by

$$\mathcal{L}(Q_i, \dot{Q}_i, t) = L(q_j(Q_i, t), \dot{q}_j(Q_i, \dot{Q}_i, t), t)\,, \quad 1 \le i, j \le n\,, \tag{57}$$

Now we take the total derivative of (57) by Q_k, noting that $\delta\mathcal{L}/\delta Q_k = \partial\mathcal{L}/\partial Q_k$,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial Q_k} &= \frac{\partial L}{\partial q_j}\frac{\delta q_j}{\delta Q_k} + \frac{\partial L}{\partial \dot{q}_j}\frac{\delta \dot{q}_j}{\delta Q_k} \\ &= \frac{\partial L}{\partial q_j}\frac{\partial q_j}{\partial Q_k} + \frac{\partial L}{\partial \dot{q}_j}\frac{d}{dt}\frac{\delta q_j}{\delta Q_k} \\ &= \frac{\partial L}{\partial q_j}\frac{\partial q_j}{\partial Q_k} + \frac{\partial L}{\partial \dot{q}_j}\frac{d}{dt}\frac{\partial q_j}{\partial Q_k}\,, \quad 1 \le j, k \le n\,. \end{aligned} \tag{58}$$

Now we take the total derivative of (57) by $\dot{Q}_k$, noting that $\delta\mathcal{L}/\delta \dot{Q}_k = \partial\mathcal{L}/\partial \dot{Q}_k$,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \dot{Q}_k} &= \frac{\partial L}{\partial q_j}\frac{\delta q_j}{\delta \dot{Q}_k} + \frac{\partial L}{\partial \dot{q}_j}\frac{\delta \dot{q}_j}{\delta \dot{Q}_k} \\ &= \frac{\partial L}{\partial q_j}\cdot 0 + \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial \dot{Q}_k} \\ &= \frac{\partial L}{\partial \dot{q}_j}\frac{\partial q_j}{\partial Q_k}\,, \quad 1 \le j, k \le n\,, \end{aligned} \tag{59}$$

where we used (47) in the last step. Now, on forming

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{Q}_k} - \frac{\partial \mathcal{L}}{\partial Q_k}\,, \quad 1 \le k \le n\,, \tag{60}$$

and substituting in from (58) and (59), we get

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}_j}\frac{\partial q_j}{\partial Q_k}\right] - \left[\frac{\partial L}{\partial q_j}\frac{\partial q_j}{\partial Q_k} + \frac{\partial L}{\partial \dot{q}_j}\frac{d}{dt}\frac{\partial q_j}{\partial Q_k}\right]\,, \quad 1 \le j, k \le n\,. \tag{61}$$

On contrafluxing the third term on the LHS and then simplifying, we get

$$\left[\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j}\right]\frac{\partial q_j}{\partial Q_k}\,, \quad 1 \le j, k \le n\,. \tag{62}$$

But since we know that the expression in the square brackets is identically zero for all j, then the whole expression is zero, leaving us to conclude that

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{Q}_k} - \frac{\partial \mathcal{L}}{\partial Q_k} = 0\,, \quad 1 \leq k \leq n\,. \tag{63}$$

6 Post SD Challenges to Conventional Partial Differentiation

I first engaged myself in the development of SD in 1982 as a response to the confusion I experienced in both advanced calculus and in physics to the use of so-called ‘partial’ differentiation. I immediately began the long-term project of writing many papers on the subject, and distributed them to friends and colleagues, as well as publishing them on the Internet on my personal website.

I have long been amazed that I am the only person on earth who seems to be dismayed enough by the problems of so-called ‘partial’ differentiation to want to do something about it. But now I am happy to report that others found a need for some form of revision to the conventions out there. Although some of the references I found out there go back to 1999, I discovered the existence of these challengers to the old ways of doing things only recently.

In 1999, the *American Journal of Physics* published an article by K. R. Brownstein, titled, “The whole-partial derivative” [Vol. 67 (7), pp. 639–641], in which the author illustrated the need to distinguish between total and partial derivatives. He suggested that we adopt the symbol $\acute{\partial}$ for the whole-partial derivative, which acts on both the explicit and implicit dependence of functions, which makes it the equivalent of δ in SD.

For a reference to a paper claiming to use Brownstein’s ideas, see “BROWNSTEIN’S WHOLE-PARTIAL derivative 2005 Lorentz gauge”[9] by A. Chubykalo, A. Espinoza and R. Flores-Alvarado who claim that the standard Lagrangian formalism does not compute the Lienard-Wiechert potentials correctly. Perhaps one day I can get back to this topic.

[9] Found at https://arxiv.org/pdf/physics/0503207.pdf.