Université Paris Cité 2022-2023

Basic notes on functional derivatives

(extracted directly from some notes I wrote on a first course on classical mechanics and Lagrangians)

D.A. Steer

APC, 10 rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France

(November 17, 2022)

Contents

1 Calculus of variations a): functionals and the Euler-Lagrange equations

How does one find the trajectory which minimises a functional (also called the principle of least action)? There is an infinite list applications of such a type of calculation outside the domain of classical mechanics. They range from optics (Snell-decartes law and mirages (see exercises below)), to the problem of the brachistochrone, through to finding geodesics of particles in relativity, classical and quantum field theory etc...). We begin by giving some simple examples of applications of this type of calculation.

1.1 Examples of functionals

In general we wish to minimise a functional $F[q_\alpha]$ which depends on a path $q_\alpha(\tau)$. Depending on which calculation we are doing, τ may not necessarily be time. Here are some examples of the type of calculation we might want to do:

1. Consider 2-dimensional Euclidean space, with coordinates (x, y). Let us fix two points $A = (x_A, y_A)$ and $B = (x_B, y_B)$, and consider a path $y(x)$ which goes between these two

points. Clearly the distance L between the points A and B depends on the path $y(x)$ between these two points:

$$\boxed{\text{functional } \ L[y], \ \text{ path } \ y(x)}$$

Which path minimises the functional $L[y]$? From experience we know the answer – a straight line!

2. One can ask the same question, but on the surface of the sphere. Then

$$\boxed{\text{functional } \ L[\theta], \ \text{ path } \ \theta(\phi)}$$

Which path minimises the functional $L[\theta]$? The answer is again known from experience – great circles.

3. The brachistochrone. In the vertical plane with coordinates (x, z) (x-horizontal and z-vertical), consider two given points A and B connected by a wire of shape $z(x)$. A bead of mass m starts at point A with zero velocity, and travels to B under the influence of gravity (we neglect any frictional forces). Which is the shape of the curve $y(x)$ such that the time taken for the bead to go from A to B is minimum?

$$\boxed{\text{functional } \ T[y], \ \text{ path } \ y(x)}$$

(This problem was first posed in 1696, and birth to the calculus of variations)

4. What must be the shape of the curve between A and B such that the area of the solid of revolution is minimum?

$$\boxed{\text{functional } \ A[y], \ \text{ path } \ y(x)}$$

5. A plane takes off from New-York bound for London. The on-board computer must choose the optimal path $\vec{x}(t)$ (a sequence of altitudes, longitudes, and latitudes for all t), such that given all the wind directions etc, the consumption of fuel is minimum:

$$\boxed{\text{functional } \ V_F[\vec{x}] \text{ (volume of fuel)}, \ \text{ path } \ (r(t), \theta(t), \phi(t)) \ \text{ (sequence of long, lat and alt)}}$$

In all these examples on must minimise a functional $F[q_\alpha]$ which depends on a path $q_\alpha(\tau)$: for the different examples we have

1. $\alpha = 1$ (one degree of freedom); path parametrised by $\tau = x$; and $q_1(\tau) = y(x)$. The functional $F = L$.

2. $\alpha = 1$ (one degree of freedom); path parametrised by $\tau = \phi$; and $q_1(\tau) = \theta(\phi)$. The functional $F = L$.

3. $\alpha = 1$ (one degree of freedom); path parametrised by $\tau = x$; and $q_1(\tau) = y(x)$. The functional $F = T$.

4. $\alpha = 1$ (one degree of freedom); path parametrised by $\tau = x$; and $q_1(\tau) = y(x)$. The functional $F = A$.

5. $\alpha = 3$ (three degrees of freedom); path parametrised by $\tau = t$; and $q_1(\tau) = r(t), q_2(\tau) = \theta(t), q_3(\tau) = \phi(t)$. The functional $F = V_F$.

Notice also that the functionals in all these examples are *numbers* whose value depends on the value of a function at all points.

1.2 Euler-Lagrange equations: derivation

From section 1.1, the functional, paths and coordinates chosen change from one example to the next. We derive the EL equations for a functional

$$F[\vec{x}] \qquad \text{with path } \ \vec{x}(t) \tag{1}$$

It is very easy to translate the results we will obtain from $\vec{x}(t) \to q_\alpha(\tau)$ and hence to all the examples considered above.

Given a path $\vec{x}(t)$, the simplest functionals are integrals along the path of $\vec{x}(t)$ and its derivatives with respect to t:

$$\begin{aligned} F_1[\vec{x}] &= \int_{t_i}^{t_f} |\vec{x}(t)|^2 dt \\ F_2[\vec{x}] &= \int_{t_i}^{t_f} |\dot{\vec{x}}(t)|^2 dt \\ F_3[\vec{x}] &= \int_{t_i}^{t_f} (\vec{x}(t)\cdot\dot{\vec{x}}(t)) dt \\ \ldots\ldots\ldots & \end{aligned}$$

Note that a functional is a scalar: on the RHS one must integrate over a scalar.

A general functional takes the form

$$\boxed{F[\vec{x}] = \int_{t_i}^{t_f} f(\vec{x}, \dot{\vec{x}}, t) dt} \tag{2}$$

[1]How do we find the path $\hat{\vec{x}}$ which minimises the functional $F[\vec{x}]$ subject to the conditions $\vec{x}(t_i) = \vec{x}_i$ and $\vec{x}(t_f) = \vec{x}_f$? The answer is that the path $\hat{\vec{x}}$ is the solution of the Euler-Lagrange equations.

Note that throughout this course we will use the 'Einstein summation convention':

$$\vec{y} \cdot \frac{\partial}{\partial \vec{x}} = \sum_i y_i \frac{\partial}{\partial x_i} = y_i \frac{\partial}{\partial x_i} \tag{3}$$

Let $\hat{\vec{x}}(t)$ be the minimising path which we are trying to determine, and $\vec{\eta}(t)$ a small variation about that path. Thus we construct the new path

$$\vec{x}(t) = \hat{\vec{x}}(t) + \vec{\eta}(t) \sim \hat{\vec{x}}(t) \tag{4}$$

and also impose

$$\vec{\eta}(t_i) = \vec{\eta}(t_f) = 0 \tag{5}$$

Hence $\vec{x}(t)$ is a path which is infinitesimally close to $\hat{\vec{x}}(t)$, and which also starts at $\vec{x}_i$ and finishes at $\vec{x}_f$. Notice that there is no variation of the independent variable t but only variation of the functions $\vec{x}(t)$.

[1]You can ask why we don't consider functionals containing up to n derivatives of $\vec{x}$, that is $F = \int_{t_i}^{t_f} f(\vec{x}, \dot{\vec{x}}, \ldots \vec{x}^{(n)}, t) dt$ where $\vec{x}^{(n)} = d^n\vec{x}/dt^n$. One reason is pragmatic: none of the examples we study is of this form. The second is fundamental: in the case of the action $F[\vec{x}] = S[\vec{x}]$, when there is a dependence on third or higher order derivatives of the generalised coordinates, the theory suffers from the so-called Ostrogradski instability. The theory has no stable ground state. See for example `http://arxiv.org/pdf/astro-ph/0601672`, page 4, for a very readable introduction to this instability.

Then

$$
\begin{aligned}
F[\vec{x}] &= \int_{t_i}^{t_f} f(\hat{\vec{x}} + \vec{\eta}, \dot{\hat{\vec{x}}} + \dot{\vec{\eta}}, t) dt \\
&= \int_{t_i}^{t_f} dt \left[f(\hat{\vec{x}}, \dot{\hat{\vec{x}}}, t) + \frac{\partial f}{\partial \vec{x}} \cdot \vec{\eta} + \frac{\partial f}{\partial \dot{\vec{x}}} \cdot \dot{\vec{\eta}} + \ldots \right] \\
&= F[\hat{\vec{x}}] + \int_{t_i}^{t_f} dt \left[\frac{\partial f}{\partial \vec{x}} \cdot \vec{\eta} + \frac{\partial f}{\partial \dot{\vec{x}}} \cdot \dot{\vec{\eta}} + \ldots \right]
\end{aligned} \tag{6}
$$

Now integrate by parts the second term in the last line:

$$
\int_{t_i}^{t_f} dt \frac{\partial f}{\partial \dot{\vec{x}}} \cdot \dot{\vec{\eta}} = \left[\frac{\partial f}{\partial \dot{\vec{x}}} \cdot \vec{\eta} \right]_{t_i}^{t_f} - \int_{t_i}^{t_f} dt \frac{d}{dt} \left(\frac{\partial f}{\partial \dot{\vec{x}}} \right) \cdot \vec{\eta} \tag{7}
$$

The first term vanishes by the boundary conditions. Hence,

$$
\delta F = F[\vec{x}] - F[\hat{\vec{x}}] = \int_{t_i}^{t_f} dt \left[\left(\frac{\partial f}{\partial \vec{x}} - \frac{d}{dt} \frac{\partial f}{\partial \dot{\vec{x}}} \right) \cdot \vec{\eta} + \ldots \right] \tag{8}
$$

The functional is minimum (technically an extremum) if $\delta F = 0$. This relation must hold for any infinitessimal $\vec{\eta}$, no matter how small. Hence we can first neglect the higher order terms in the Taylor expansion, noted with $\ldots$. The remaining term is proportional to $\vec{\eta}$. In the discussion presented so far we are working under the assumption that the generalised coordinates q_α (which in this case are $\vec{x}$) are all **independent**. Hence the $\delta x_i = \eta_i$ $(i = 1, 2, 3)$ are also independent.[2] This is only possible if the coefficient of $\vec{\eta}$ vanishes for all $t_i \le t \le t_f$. Thus $\hat{\vec{x}}(t)$ minimises F if and only if

$$
\boxed{\frac{\partial f}{\partial \vec{x}} - \frac{d}{dt} \frac{\partial f}{\partial \dot{\vec{x}}} = 0} \tag{9}
$$

along the path $\hat{\vec{x}}(t)$. This is the Euler-Lagrange equation, and the theory that underlies it the calculus of variations. Notice that the above equation is a compact form for writing three equations, one for each component of $\vec{x}$:

$$
\frac{\partial f}{\partial x_i} - \frac{d}{dt} \frac{\partial f}{\partial \dot{x}_i} = 0 \qquad (i = 1, 2, 3) \tag{10}
$$

1.2.1 General EL equations

The argument is straightforwardly generalised to functionals

$$
\boxed{F[q_\alpha] = \int_{\tau_i}^{\tau_f} d\tau f(q_\alpha, \dot{q}_\alpha, \tau)} \tag{11}
$$

where the path is $q_\alpha(\tau)$ and $\dot{q_\alpha} = dq_\alpha / d\tau$. In that case

$$
\delta F = F[q_\alpha] - F[\hat{q}_\alpha] = \int_{\tau_i}^{\tau_f} d\tau \left[\frac{\partial f}{\partial q_\alpha} - \frac{d}{d\tau} \frac{\partial f}{\partial \dot{q}_\alpha} \right] \delta q_\alpha \tag{12}
$$

[2] This will not be the case when the system is subject to constraints: see below

Since the q_α are assumed independent, the path which minimises $F[q_\alpha]$ subject to the conditions $q_\alpha(\tau_i)$ and $q_\alpha(\tau_f)$ fixed satisfies

$$\boxed{\frac{\partial f}{\partial q_\alpha} - \frac{d}{d\tau}\frac{\partial f}{\partial \dot{q}_\alpha} = 0} \tag{13}$$

Note that these are N equations, one for each value of α.

Finally, observe the following three points

1. The functions f and g related by

$$g(q_\alpha, \dot{q}_\alpha, \tau) = f(q_\alpha, \dot{q}_\alpha, \tau) + \text{const} \tag{14}$$

   satisfy the same EL equations (though the value of the functionals $F[g]$ and $F[f]$ differ).

2. The functions f and g related by

$$g = f + \frac{d\Lambda(q_\alpha, \tau)}{d\tau} \tag{15}$$

   where $\Lambda(q_\alpha, \tau)$ is an arbitrary scalar function of τ and q_α (but not $\dot{q}_\alpha$) also satisfy the same EL equations. To see this, either plug straight into the EL equations (which is a mess), or recall that the EL equations are derived from F which itself only changes by a constant under the transformation.

3. In (13), apart from their independence, the q_α are arbitrary. Suppose we'd decided to work with different independent generalised coordinates Q_α. Then in terms of these the EL equations must take the form

$$\frac{\partial f}{\partial Q_\alpha} - \frac{d}{d\tau}\frac{\partial f}{\partial \dot{Q}_\alpha} = 0. \tag{16}$$

1.3 A note on Functionals and functional derivatives

Functions f are maps from (for example) the reals to the reals

$$f : \mathbb{R} \to \mathbb{R} \qquad\qquad \text{(function).} \tag{17}$$

That is, given an $x \in \mathbb{R}$, $f(x) \in \mathbb{R}$. Of course functions may also be vectors: an example is the magnetic field $\vec{B}$, which associates the magnetic field to every point of 3-space. In this case the function is a mapping from $\mathbb{R}^3 \to \mathbb{R}^3$. One can also have scalar function, such as $\phi(\vec{x})$ (the Higgs field is a possible example) which is a mapping from $\mathbb{R}^3 \to \mathbb{R}$, as well as complex functions.

Let $\mathcal{F}$ denote the space of functions. In physics one generally deals with functions which are infinitely differentiable so that the underlying coordinate space is a manifold M and the space of functions is denoted by $C^\infty(M)$. A functional F is a map

$$F : \mathcal{F} \to \mathbb{R} \qquad\qquad \text{(functional).} \tag{18}$$

That is, for $f \in \mathcal{F}$, $F[f] \in \mathbb{R}$. A functional depends on the value of the function $f(\vec{x})$ at all points. (It should now be clear that a functional is *not* a function of a function – which is nothing other than a function!)

Note that the arguments of a function and a functional are labelled differently: $f(\bullet)$ and $F[\bullet]$ respectively.

1.3.1 Functional derivative

The Functional derivative

$$\frac{\delta F}{\delta f(x)}$$

of a functional $F[f]$ is defined as follows: for any infinitesimal $\delta f(x)$,

$$\delta F[f] = F[f+\delta f] - F[f] \equiv \int dx \frac{\delta F}{\delta f(x)} \delta f(x) \tag{19}$$

This is in analogy for an ordinary function of n variables,

$$\delta f(x_1,\ldots,x_n) = f(x_1+\delta x_1,\ldots,x_n+\delta x_n) - f(x_1,\ldots,x_n) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\delta x_j \tag{20}$$

Similarly the functional taylor series is:

$$F[f_0+f] = F[f_0] + \int dx \frac{\delta F[f_0]}{\delta f(x)} \delta f(x) + \frac{1}{2}\int dx_1 dx_2 \frac{\delta^2 F[f_0]}{\delta f(x_1)\delta f(x_2)} \delta f(x_1)\delta f(x_2) + \ldots \tag{21}$$

<u>*Specific examples*</u>:

1. Consider

$$F[f] = \int dx f(x) \qquad \Rightarrow \qquad \delta F = \int dx \delta f(x) \tag{22}$$

so that comparing with (19) gives

$$\frac{\delta F}{\delta f(x)} = 1 \tag{23}$$

2. Consider

$$F[f] = \int dx \delta(x-y) f(x) \qquad \Rightarrow \qquad \delta F = \int dx \delta(x-y)\delta f(x) \tag{24}$$

Note that the left hand side is not only a functional, but also a function of y, since y is not integrated over on the right hand side. Comparing with (19) gives

$$\frac{\delta F}{\delta f(x)} = \delta(x-y) \tag{25}$$

But by definition $F[f] = \int dx \delta(x-y) f(x) = f(y)$. So we get to the very useful identity

$$\frac{\delta f(y)}{\delta f(x)} = \delta(x-y). \tag{26}$$

3. Consider

$$F[f] = \int G(x,y) f(y) dy \tag{27}$$

Note that the left hand side is not only a functional, but also a function of x, since x is not integrated over on the right hand side. It therefore makes no sense to calculate $\delta F/\delta f(x)$, but we can calculate $\delta F/\delta f(z)$. From above

$$\delta F = \int dy G(x,y)\delta f(y) = \int dz G(x,z) \delta f(z) \tag{28}$$

since y is the dummy integration variable. Hence comparing with (19)

$$\frac{\delta F}{\delta f(z)} = G(x,z) \tag{29}$$

4. Consider

$$F[f] = a + \int dx b(x) f(x) + \frac{1}{2}\int dx dy c(x,y) f(x) f(y) \tag{30}$$

where, without loss of generality one can assume $c(x,y) = c(y,x)$. Then

$$\delta F[f] = \int dx b(x)\delta f(x) + \frac{1}{2}\int dx dy c(x,y)[f(x)\delta f(y) + f(y)\delta f(x)] \tag{31}$$

so that from (19):

$$\frac{\delta F}{\delta f(x)} = b(x) + \int dy c(x,y) f(y) \tag{32}$$

5. Consider

$$F[f] = \int dx \delta(x-y)(f(x))^n = f(y)^n \qquad \Rightarrow \qquad \delta F = \int dx \delta(x-y) n f(x)^{n-1}\delta f(x) \tag{33}$$

(using that $(f+\delta f)^n = f^n + nf^{n-1}\delta f + \ldots$). Thus

$$\frac{\delta(f(y)^n)}{\delta f(x)} = \delta(x-y) n f(x)^{n-1} \tag{34}$$

1.3.2 Functional derivatives and the EL equation

Finally, now return to Eq. (12), which we can rewrite as

$$\delta F = F[q_\alpha] - F[\hat{q}_\alpha] = \int_{\tau_i}^{\tau_f} d\tau \left[\frac{\partial f}{\partial q_\alpha} - \frac{d}{d\tau}\frac{\partial f}{\partial \dot{q}_\alpha}\right]\delta q_\alpha \equiv \int_{\tau_i}^{\tau_f} d\tau \frac{\delta F}{\delta q_\alpha}\delta q_\alpha. \tag{35}$$

Thus the EL equations are equivalent to setting

$$\frac{\delta F}{\delta q_\alpha} = 0 \tag{36}$$

exactly as one would expect at an extremum.

1.3.3 Functional Integration

Functional integration, or path integration, is a huge subject in itself, but provides one of the most powerful methods of modern theoretical physics. The functional integration approach is (essentially by definition) central to systems with an infinite number of degrees of freedom, and is very suitable for the introduction and formulation of diagrammatic perturbation theory of quantum field theory and statistical physics. Any beginning course on quantum field theory starts with functional integrals: this is well outside the scope of the present course.

For a curious reader, the path integral (denoted by $\int \mathcal{D}f(x)$) was first introduced by Feyman: for example in quantum mechanics, the transition amplitude between an initial quantum state $|q_i t_i\rangle$ and a final quantum state $|q_f t_f\rangle$ is written as

$$\langle q_f t_f | q_i t_i \rangle = \int \mathcal{D}q e^{\frac{i}{\hbar}S[q]} \tag{37}$$

where for simplicity we have assumed one generalised coordinate q, and S is the action mentioned above. The classical path is $q(t)$. Rather loosely, a preliminary way to think of $\mathcal{D}q$ is as

$$\mathcal{D}q \propto \prod_t dq(t) \tag{38}$$

2 Calculus of variations b): Symmetries and conservation laws

In nearly all physical phenomena there exist quantities which are conserved during the evolution of the system — depending on the problem, these may be the angular momentum, the total energy, the total momentum etc. Knowing the existence of such conserved quantities is crucial in determining the dynamics of the system under consideration (think of the collision of billiard balls for example). The aim of this short section is to see how such conservation laws appear in the functional approach.

In the following we work with $F[q_\alpha] = \int_{\tau_i}^{\tau_f} f(q_\alpha, \dot{q}_\alpha, \tau)$.

Definition A function $C(q_\alpha, \dot{q}_\alpha, \tau)$ is a constant of motion (= conserved quantity) along the path $\hat{q}_\alpha$ solution to the EL equations, if its total derivative wrt τ vanishes along the path:

$$\boxed{\frac{dC}{d\tau} = \frac{\partial C}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial C}{\partial \dot{q}_\alpha}\ddot{q}_\alpha + \frac{\partial C}{\partial \tau} = 0} \tag{39}$$

Note the use of the summation convention here. The total derivative takes into account the τ-dependence due to the evolution in τ of the $q_\alpha(\tau)$ and $\dot{q}_\alpha(\tau)$.

When do we find conserved quantities?

1. *Suppose f is explicitly τ independent, $\partial f/\partial\tau = 0$, so that $f = f(q_\alpha, \dot{q}_\alpha)$.*

   Along the trajectory satisfying the EL equations, it follows that

$$\begin{aligned} \frac{df}{d\tau} &= \frac{\partial f}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial f}{\partial \dot{q}_\alpha}\ddot{q}_\alpha \\ &= \frac{d}{d\tau}\left(\frac{\partial f}{\partial \dot{q}_\alpha}\right)\dot{q}_\alpha + \frac{\partial f}{\partial \dot{q}_\alpha}\ddot{q}_\alpha \\ &= \frac{d}{d\tau}\left(\frac{\partial f}{\partial \dot{q}_\alpha}\dot{q}_\alpha\right) \end{aligned} \tag{40}$$

   where in going from the 1st to 2nd line, we've used the EL equations. Therefore, putting the lhs on the rhs, it follows that

$$h = \frac{\partial f}{\partial \dot{q}_\alpha}\dot{q}_\alpha - f = \text{constant} \tag{41}$$

   Namely h is conserved if $\partial f/\partial\tau = 0$. (Théorème de Beltrami)

   If such a conserved quantity exists, use it! It provides a different way of encoding the EL equations and has the advantage of being first order in time, making it often easier to find the path which minimises the functional F.

2. *Suppose f does not depend explicitly on one of the generalised coordinates, say q_1, i.e. $\partial f/\partial q_1 = 0$, so that $f = f(q_2, q_3, \ldots, q_N, \dot{q}_1, \ldots, \dot{q}_N, t)$.*

   Then from the EL equations, it follows that

$$\frac{d}{d\tau}\left(\frac{\partial f}{\partial \dot{q}_1}\right) = 0 \qquad \Rightarrow \qquad \frac{\partial f}{\partial \dot{q}_1} = \text{constant} \tag{42}$$

   There are numerous examples of this. One is that of a particle moving in a central potential $V(r)$, say in two dimensions. Then the obvious generalised coordinates are (r, θ) and the dynamics of the particle is determined by minimising a functional $F[r, \theta] = \int dt f(\dot{r}, \dot{\theta}, r)$. The corresponding conserved quantity $\partial f / \partial \dot{\theta}$ is nothing other than the angular momentum of the particle.

3. *Noether's theorem and symmetry*

   What happens when f is invariant (or symmetric) under a change of generalised coordinates?

   Consider a one-parameter family of maps

$$q_\alpha(\tau) \to Q_\alpha(s, \tau) \qquad (s \in R) \tag{43}$$

   such that $Q_\alpha(0, t) = q_\alpha$. If this transformation leaves the functional form of $f(q_\alpha, \dot{q}_\alpha, t)$ invariant *to linear order in* s, it is said to be a *continuous symmetry* of f. Mathematically, a continuous symmetry requires

$$\left.\frac{\partial f}{\partial Q_\alpha} \frac{\partial Q_\alpha}{\partial s}\right|_{s=0} + \left.\frac{\partial f}{\partial \dot{Q}_\alpha} \frac{\partial \dot{Q}_\alpha}{\partial s}\right|_{s=0} = 0 \tag{44}$$

   Why? Since we work to linear order (43) can be written as

$$q_\alpha(t) \to Q_\alpha(s, \tau) = q_\alpha(\tau) + s h_\alpha(\tau) + \dots . \tag{45}$$

   Of course, for any given Lagrangian, the aim is to find exactly *which* h_α (or equivalently which $Q_\alpha(s, \tau)$) leave the Lagrangian invariant. Then, on doing a Taylor expansion to linear order,

$$f(Q_\alpha(s, \tau), Q_\alpha(s, \tau), \tau) = f(q_\alpha, \dot{q}_\alpha, \tau) + s \left(\left.\frac{\partial f}{\partial Q_\alpha}\right|_{s=0} h_\alpha + \left.\frac{\partial f}{\partial \dot{Q}_\alpha}\right|_{s=0} \dot{h}_\alpha \right) + \dots \tag{46}$$

   For a symmetry, the term proportional to s must vanish. This is nothing other than the above condition, since once you have found $Q_\alpha(s, \tau)$ then h_α is obtained from (45) through

$$h_\alpha(\tau) = \left.\frac{\partial Q_\alpha(s, \tau)}{\partial s}\right|_{s=0} \tag{47}$$

   Statement of theorem:

   Noether's theorem states that for each such symmetry there is a conserved quantity given by

$$\boxed{\left.\frac{\partial f}{\partial \dot{q}_\alpha} \frac{\partial Q_\alpha}{\partial s}\right|_{s=0} = \text{constant}} \tag{48}$$

   (Again recall that there's an implicit summation over α here.)

   Proof:

By definition

$$
\begin{aligned}
0 = \left.\frac{\partial f}{\partial s}\right|_{s=0} &= \left.\frac{\partial f}{\partial Q_\alpha}\frac{\partial Q_\alpha}{\partial s}\right|_{s=0} + \left.\frac{\partial f}{\partial \dot{Q}_\alpha}\frac{\partial \dot{Q}_\alpha}{\partial s}\right|_{s=0} \\
&= \left.\frac{\partial Q_\alpha}{\partial s}\right|_{s=0}\frac{\partial f}{\partial q_\alpha} + \left.\frac{\partial \dot{Q}_\alpha}{\partial s}\right|_{s=0}\frac{\partial f}{\partial \dot{q}_\alpha} \\
&= \left.\frac{\partial Q_\alpha}{\partial s}\right|_{s=0}\frac{d}{dt}\frac{\partial f}{\partial \dot{q}_\alpha} + \left.\frac{\partial \dot{Q}_\alpha}{\partial s}\right|_{s=0}\frac{\partial f}{\partial \dot{q}_\alpha} \\
&= \frac{d}{dt}\left[\frac{\partial f}{\partial \dot{q}_\alpha}\left.\frac{\partial Q_\alpha}{\partial s}\right|_{s=0}\right]
\end{aligned}
\tag{49}
$$

But since the LHS vanishes for a continuous symmetry, it follows that (48) is conserved.

We will see numerous applications of Noethers theorem and conserved quantities once we have defined the action for classical mechanics. (For example, many systems are invariant under spatial translations $\vec{x}(t) \to \vec{x}(t) + s\vec{n}$. The quantity conserved by Noether's theorem will turn out to be the total linear momentum).