Version of March 23, 2020

Variations and the Noether Theorem

Matthias Blau

Contents

1 Introduction

The aim of this short note is to show how various important aspects of analytical mechanics, like the most useful non-fancy version of Noether's theorem, Hamilton's (or the variational) principle, and aspects of Hamilton-Jacobi theory can all be derived and explained in a straightforward way from a simple "Variational Master Equation" (which itself has a 2-line derivation), without having to first go through variational calculus or other contortions.

The only prerequisites are a basic familiarity with the basic framework of Lagrangian mechanics. Variational calculus and other fun things can then (if desired) be developed subsequently as a simple reformulation of what has already been achieved.

I make no claim to originality in these notes (all this is standard material), but I do find that many common text-books deal with these matters in an unnecessarily complicated and misleading way.

2 Variations

Let Q be the configuration space of a mechanical system, with (generalised) coordinates q^a. A path / trajectory in Q is then a curve $q^a = q^a(t)$. A *variation* of a curve $q^a(t)$ is then nothing other than an infinitesimal deformation

$$q^a(t) \to q^a(t) + \delta q^a(t) \tag{2.1}$$

of this curve, that's it. Concretely one can think of such a deformation as

$$\delta q^a(t) = \epsilon \eta^a(t) \quad , \tag{2.2}$$

say, with ϵ a constant infinitesimal parameter, and $\eta^a(t)$ an ordinary function of the coordinate t (subject to some suitable differentiability conditions, C^2 is usually good enough). The characteristic property of such variations is that the (generalised) velocities

$$\dot{q}^a(t) = \tfrac{d}{dt} q^a(t) \tag{2.3}$$

evidently vary according to

$$\tfrac{d}{dt} q^a(t) \to \tfrac{d}{dt} q^a(t) + \tfrac{d}{dt} \delta q^a(t) \quad . \tag{2.4}$$

Thus one has

$$\delta \dot{q}^a(t) \equiv \delta(\tfrac{d}{dt} q^a(t)) = \tfrac{d}{dt} \delta q^a(t) \tag{2.5}$$

("δ and d/dt commute"), as well as

$$\delta \ddot{q}^a(t) = \tfrac{d^2}{dt^2} \delta q^a(t) \tag{2.6}$$

etc.

Now, given any function $\mathcal{F}$ of $q^a(t)$ and $\dot{q}^a(t)$, and perhaps other variables, (below we will consider e.g. a Lagrangian $\mathcal{L}(q, \dot{q}, t)$), its variation can simply be found by using the standard chain rule, i.e.

$$\delta \mathcal{F}(q^a(t), \dot{q}^a(t), \ldots, t) = \frac{\partial \mathcal{F}}{\partial q^a(t)} \delta q^a(t) + \frac{\partial \mathcal{F}}{\partial \dot{q}^a(t)} \delta \dot{q}^a(t) + \ldots \tag{2.7}$$

(a summation over the index ^{a} is implied in these and subsequent equations). Thus the upshot of this is that if you know how to partially differentiate (and you should if you are reading this) then you also know how to vary.

<u>Remarks:</u>

1. Note that in the above t is not varied (variations vary the paths $q^a(t)$, *not* the parameter t - we will come back to issues related to this on various occasions below).

2. For some obscure reason, students are often taught that variations are to be calculated by first calculating something like
$$\mathcal{F}(q^a + \delta q^a, \dot{q}^a + \delta \dot{q}^a, \ldots) - \mathcal{F}(q^a, \dot{q}^a, \ldots) \ , \tag{2.8}$$
then expanding and keeping only the terms linear in the variations, to obtain $\delta \mathcal{F}$.

   Well, that is correct, but it is also *totally stupid* as a means of calculating the variation. It is like saying that if you want to calculate the differential of a function $f(x, y)$ you need to take the difference $f(x+dx, y+dy) - f(x, y)$, and keep the linear terms in the differentials. Yes, that is correct, but once you know how to differentiate, this is not the way you will actually calculate the differential. Rather, you just use the rules of partial differentiation,
$$f(x + dx, y + dy) - f(x, y) \quad \Rightarrow \quad df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \ . \tag{2.9}$$
By exactly the same reasoning one has
$$\mathcal{F}(q^a + \delta q^a, \dot{q}^a + \delta \dot{q}^a, \ldots) - \mathcal{F}(q^a, \dot{q}^a, \ldots) \quad \Rightarrow \quad \delta \mathcal{F} = \frac{\partial \mathcal{F}}{\partial q^a} \delta q^a + \frac{\partial \mathcal{F}}{\partial \dot{q}^a} \delta \dot{q}^a + \ldots \tag{2.10}$$
which is precisely the rule given above.

3. In particular, one therefore has simple familiar rules like
$$\delta(q^a q^b) = (\delta q^a) q^b + q^a (\delta q^b) = q^b (\delta q^a) + q^a (\delta q^b) \quad , \quad \delta((q^1)^2) = 2q^1 (\delta q^1) \tag{2.11}$$
and
$$\delta(\delta_{ab} \dot{q}^a \dot{q}^b) = 2\delta_{ab} \dot{q}^a \delta \dot{q}^b \ . \tag{2.12}$$

4. For some obscure reason, this simple variation is often confused with the *functional derivative*, which appears in similar contexts, but which is a more complicated object, with more involved (distributional) calculational rules like
$$\text{Functional Derivative :} \quad \frac{\delta q^a(t)}{\delta q^b(t')} = \delta^a_b \delta(t - t') \ . \tag{2.13}$$
This functional derivative is really like an infinite-dimensional differential. It can be used to differentiate *functionals*, i.e. functions of functions (like an action), and variational calculus can be conveniently phrased in this framework (Frechet calculus on infinite-dimensional topological vector spaces ...), but we are most certainly not doing this here. At the level we are working at, the naive and simple variations (as we have defined them) are completely sufficient.

5. The crucial difference between variations of a function of a path $q^a(t)$ and its velocity $\dot{q}^a(t)$, say, and the differential of a function of two finite-dimensional variables x and y, is that in the latter case the differentials dx and dy are independent, whereas the characteristic and defining feature of variations is that $\delta\dot{q} = (d/dt)\delta q$. This has no impact on the calculational rules so far, but it does allow us to rewrite the variation in a different way. This will directly lead us to the crucial *Variational Master Equation* (3.2) in section 3 below.

3 Variational Master Equation (VME)

We saw above, that the variation of a function $\mathcal{L}(q,\dot{q},t)$ (now a Lagrangian for us) is simply

$$\delta\mathcal{L}(q^a(t),\dot{q}^a(t),t) = \frac{\partial\mathcal{L}}{\partial q^a(t)}\delta q^a(t) + \frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)}\delta\dot{q}^a(t) \tag{3.1}$$

Using the one piece of information we have at the moment, namely the identity (2.5), we can rewrite this as

$$\delta\mathcal{L}(q^a(t),\dot{q}^a(t),t) = \left(\frac{\partial\mathcal{L}}{\partial q^a(t)} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)}\right)\delta q^a(t) + \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)}\delta q^a(t)\right) \quad . \tag{3.2}$$

This is what I will refer to as the *Variational Master Equation* (VME).

It will be useful to introduce (and occasionally use) the abbreviation

$$\mathcal{E}_a \equiv \frac{\delta\mathcal{L}}{\delta q^a} := \frac{\partial\mathcal{L}}{\partial q^a} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}^a} \tag{3.3}$$

for the Euler-Lagrange equations for the a'th "field" $q^a(t)$, or for the Euler-Lagrange (or variational) derivative of the Lagrangian $\mathcal{L}$ with respect to $q^a(t)$. With this, and the usual definition of the (generalised) conjugate momenta

$$p_a = \frac{\partial\mathcal{L}}{\partial\dot{q}^a} \quad , \tag{3.4}$$

the VME (still valid for an arbitrary variation) can also compactly be written as

$$\delta\mathcal{L} = \mathcal{E}_a\delta q^a + \frac{d}{dt}(p_a\delta q^a) \quad . \tag{3.5}$$

What makes this equation so useful is that it relates 3 apparently quite different objects. On the lhs, one has a variation, the Euler-Lagrange equations appear in the 1st term on the rhs, and the 2nd term on the rhs is a total time-derivative, so that structurally the equation looks like

$$\text{Variation} = \text{Euler-Lagrange Equations} + \text{Total Time-Derivative} \quad . \tag{3.6}$$

Thus, if we can eliminate or constrain one of the terms in these equations, then we obtain a potentially non-trivial and interesting relation between the other two. This can be achieved by selecting appropriate variations or classes of variations and/or by integrating the VME.

Concretely,

1. by integrating and choosing the variations to preserve the end-points of the path, one eliminates the 2nd term on the right-hand side and obtains a 1-line proof of Hamilton's principle that Lagrangian dynamics is such that the path is a stationary point of the action (section 4);

2. by choosing special variations $\delta_s q^a$ that leave the Lagrangian invariant ($\delta_s \mathcal{L} = 0$, infinitesimal symmetries) or invariant up to a total time-derivative, one constrains the left-hand side and obtains a 1-line proof of Noether's (first) theorem (section 5);

3. by restricting to solutions of the Euler-Lagrange equations and variations among them one eliminates the 1st term on the right-hand side and obtains a simple proof of the Hamilton-Jacobi relations which relate the time- and space-derivatives of the "classical" action to energy and momentum respectively (section D).

4 Hamilton's Action Principle from the VME: 2-Line Proof

Our interest here will mainly be in the second option, but just for completeness we start with the argument for the first item. Thus we turn to the VME (3.2) and seek to eliminate the total-derivative term on the rhs, by integrating this equation over a time-interval $[t_1, t_2]$,

$$\int_{t_1}^{t_2} dt\ \delta\mathcal{L}(q^a(t), \dot{q}^a(t), t) = \int_{t_1}^{t_2} dt\ \left(\frac{\partial\mathcal{L}}{\partial q^a(t)} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)}\right)\delta q^a(t) + \int_{t_1}^{t_2} dt\ \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)}\delta q^a(t)\right) \ . \tag{4.1}$$

Since t is not varied (this is important and we will come back to this at length in our discussion of Noether's theorem below: we are only considering variations of the path $q^a(t)$, not explicit variations of the time coordinate t), we can exchange the variation and the integral on the lhs,

$$\int_{t_1}^{t_2} dt\ \delta\mathcal{L}(q^a(t), \dot{q}^a(t), t) = \delta\int_{t_1}^{t_2} dt\ \mathcal{L}(q^a(t), \dot{q}^a(t), t) \ . \tag{4.2}$$

The 2nd term on the rhs is the integral of a total derivative, and thus a boundary term. If we now restrict to variations that leave the end-points $q(t_1)$ and $q(t_2)$ fixed, i.e. $\delta q^a(t_1) = \delta q^a(t_2) = 0$, then the 2nd term on the rhs is zero,

$$\int_{t_1}^{t_2} dt\ \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)}\delta q^a(t)\right) = p_a(t_2)\delta q^a(t_2) - p_a(t_1)\delta q^a(t_1) = 0 \ . \tag{4.3}$$

Introducing the *action*

$$S[q] = \int_{t_1}^{t_2} dt\ \mathcal{L}(q^a(t), \dot{q}^a(t), t) \ , \tag{4.4}$$

the integrated version of the VME can thus be written as

$$\delta S[q] = \int_{t_1}^{t_2} dt\ \left(\frac{\partial\mathcal{L}}{\partial q^a(t)} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)}\right)\delta q^a(t) \ . \tag{4.5}$$

Since the integral on the rhs is zero for all allowed variations $\delta q^a(t)$ iff the term in brackets in the integrand is zero, we conclude

$$\delta S[q] = 0 \quad \forall\delta q \quad \Leftrightarrow \quad \frac{\partial\mathcal{L}}{\partial q^a(t)} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}^a(t)} = 0 \ . \tag{4.6}$$

These are the Euler-Lagrange equations.

Since the condition $\delta S[q] = 0$ means that the path $q(t)$ is a stationary "point" of the action functional, we arrive at *Hamilton's principle* or the action principle, that the motion of a dynamical system between given initial and end points and times is such that the action is stationary.

Remarks:

1. The action $S[q]$ is a *functional*, a function on the space of paths, assigning to the path $q^a(t)$ the real number $S[q]$. It also depends on the initial values t_i and $q^a(t_i)$, $i = 1, 2$, but if these are fixed (as in the present case), then it is not necessary to indicate this dependence explicitly.

2. As in the case of an ordinary function, a stationary point can be a (local) minimum or maximum (an extremum), or something like a saddle-point, this is not relevant for the action principle. As one typically finds either a minimum or a maximum (so that $-S$ would have a minimum), common terminology includes "extremising the action" and "the principle of least action".

5 Noether Theorem from the VME: 1-Line Proof

Here is another obvious consequence of the VME (3.2):

Let $q^a \to q^a + \delta_s q^a$ be a variation that leaves the Lagrangian invariant for all paths $q^a(t)$,

$$\delta_s \mathcal{L}(q, \dot{q}, t) = \frac{\partial \mathcal{L}}{\partial q^a(t)} \delta_s q^a(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}^a(t)} \delta_s \dot{q}^a(t) = 0 \quad . \tag{5.1}$$

We will refer to such a transformation as an infinitesimal symmetry of the Lagrangian. Then there is a corresponding conserved quantity, namely

$$P_\delta = \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \delta_s q^a = p_a \delta_s q^a \quad , \tag{5.2}$$

i.e. P_δ is constant along any solution to the Euler-Lagrange equations.

Proof: $\delta_s \mathcal{L} = 0$ implies

$$0 = \left(\frac{\partial \mathcal{L}}{\partial q^a(t)} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^a(t)} \right) \delta_s q^a(t) + \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^a(t)} \delta_s q^a(t) \right) \quad . \tag{5.3}$$

Thus

$$\frac{\partial \mathcal{L}}{\partial q^a(t)} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}^a(t)} = 0 \quad \Rightarrow \quad \frac{d}{dt} P_\delta = 0 \quad . \tag{5.4}$$

That's it. It really is as simple as that!

Here is a minor (and obvious, but quite useful) variant and generalisation:

Let $q^a \to q^a + \delta_s q^a$ be a variation that leaves the Lagrangian *quasi-invariant* for all paths $q^a(t)$, i.e. invariant up to a total time-derivative,

$$\delta_s \mathcal{L}(q,\dot{q},t) = \frac{\partial \mathcal{L}}{\partial q^a(t)}\delta_s q^a(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}^a(t)}\delta_s \dot{q}^a(t) = \frac{d}{dt}F_\delta \tag{5.5}$$

(quasi-symmetry or also simply just symmetry of the Lagrangian). Then (by the same reasoning as above and by simply replacing 0 by $(d/dt)F_\delta$ on the lhs of (5.3)) there is a corresponding conserved quantity, namely

$$P_\delta = p_a \delta_s q^a - F_\delta \quad . \tag{5.6}$$

Remarks:

1. Note that in the above we considered only variations of the paths $q^a(t)$, not variations of the independent variable t. This does not mean that we cannot deal with symmetries associated with transformations of t. What it means is that we should reinterpret them as transformations acting on the $q^a(t)$ alone. This avoids many completely unnecessary complications and pitfalls that invariably arise when one tries to formulate the Noether theorem directly for symmetries that involve explicit transformations of t. It is therefore surprising, that most textbook treatments of Noether's theorem actually take this latter, more complicated, approach.

2. In this more traditional approach, one considers infinitesimal transformations

$$\bar{t} = t + \epsilon X(q^a,t) \quad , \quad \bar{q}^a = q^a + \epsilon Y^a(q^a,t) \quad . \tag{5.7}$$

which are such that under the substitution

$$t \to \bar{t} \quad , \quad q^a(t) \to \bar{q}^a(\bar{t}) \tag{5.8}$$

the action $\int dtL$ is invariant to order ϵ (and up to boundary terms). However, one can think of this combined transformation as defining a true variation (only $q^a(t)$ is varied, not t) via (retaining only the linear term in ϵ)

$$\delta q^a(t) = \bar{q}^a(t) - q^a(t) = \epsilon\left(Y^a(q(t),t) - X(q(t),t)\dot{q}^a(t)\right) \quad . \tag{5.9}$$

Then the above invariance condition for the *action* (including the transformation of the integration measure dt) under (5.7) turns out to be completely equivalent to quasi-invariance of the *Lagrangian* under this variation (5.9). The proof of the traditional version of the Noether theorem and its equivalence with the variational version will be provided in sections A and B.

Once one has convinced oneself that there is no loss of generality involved in only considering true variations (that act only on the paths $q^a(t)$ and not on the parameter t), it is much more convenient to phrase the Noether theorem entirely in terms of these variatonal terms, as done above, because the Noether theorem for these variations has the simple 1-line proof given above.

3. In particular, we can think of infinitesimal time-translations $t \to \bar{t} = t + \epsilon$ alternatively as defining new (translated) paths $\bar{q}^a(t)$ by

$$\bar{q}^a(t) = q^a(t-\epsilon) \ . \tag{5.10}$$

Taylor expanding this, we have

$$\bar{q}^a(t) = q^a(t) - \epsilon \dot{q}^a(t) + \dots \tag{5.11}$$

The difference between the left-hand side and the first term on the right-hand side is now an infinitesimal difference between two different paths at the same point, and therefore this defines a variation

$$\delta q^a(t) = -\epsilon \dot{q}^a(t) \quad , \quad \delta \dot{q}^a(t) = -\epsilon \ddot{q}^a(t) \ . \tag{5.12}$$

This is just the special case of the variation (5.9) introduced above, with $X = 1, Y^a = 0$. Acting with this variation on the Lagrangian, one finds

$$\delta \mathcal{L} = -\frac{d}{dt}(\epsilon \mathcal{L}) + \frac{\partial}{\partial t}(\epsilon \mathcal{L}) \ , \tag{5.13}$$

so the Lagrangian is quasi-invariant if $\mathcal{L}$ does not depend explicitly on t, and we are now entitled to call this variation $\delta_s q^a(t)$. The corresponding conserved quantity is then essentially the Hamiltonian function (energy) $\mathcal{H}$,

$$p_a \delta_s q^a - \mathcal{F}_\delta = -\epsilon(p_a \dot{q}^a - \mathcal{L}) = -\epsilon \mathcal{H} \ . \tag{5.14}$$

6 Example: Galilean Symmetry Group of a Closed System

As a warm-up we consider a single free particle with Lagrangian $\mathcal{L} = T = m\dot{\vec{x}}^2/2$, and the Galilei group consisting of space- and time-translations, rotations and Galilean velocity transformations (boosts),

$$t \to \bar{t} = t + c \quad , \quad x^a \to \bar{x}^a = R^a_{\ b} x^b - w^a t + c^a \ , \tag{6.1}$$

with $R^T R = \mathbb{I}$. For Noether's theorem we need the corresponding infinitesimal transformations,

$$\Delta t = \epsilon \quad , \quad \Delta x^a = \omega^a_{\ b} x^b - \nu^a t + \epsilon^a \ , \tag{6.2}$$

with ω anti-symmetric ($R = \mathbb{I} + \omega, R^T R = \mathbb{I} \Rightarrow \omega^T = -\omega$, note that only in 3 spatial dimensions one can trade $\omega^a_{\ b}$ for a vector ω^a with $\Delta\vec{x} = \vec{\omega} \times \vec{x}$). Using the standard rule (5.9), we convert this into a variation of the path $x^a(t)$, namely

$$\delta x^a = -\epsilon \dot{x}^a + \omega^a_{\ b} x^b - \nu^a t + \epsilon^a \ , \tag{6.3}$$

with ω anti-symmetric ($R = \mathbb{I} + \omega, R^T R = \mathbb{I} \Rightarrow \omega^T = -\omega$,

The Lagrangian is obviously strictly invariant under space-translations and rotations, and is also quasi-invariant under time-translations (cf. (5.13)) and Galilean boosts,

$$\delta_\nu x^a = -\nu^a t \quad \Rightarrow \quad \delta_\nu \mathcal{L} = -m\dot{\vec{x}}.\vec{\nu} = \tfrac{d}{dt}(-m\vec{x}.\vec{\nu}) \ , \tag{6.4}$$

and one obtains the corresponding conserved quantities

$$
\begin{array}{lll}
\delta x^a = -\epsilon \dot{x}^a & \Rightarrow & \text{Energy} \quad E = \mathcal{H} = p_a \dot{x}^a - \mathcal{L} = T \\
\delta x^a = \epsilon^a & \Rightarrow & \text{Momentum} \quad p^a \\
\delta x^a = \omega^a_{\ b} x^b & \Rightarrow & \text{Angular Momentum} \quad \ell^{ab} = x^a p^b - x^b p^a \\
 & & (\text{in 3 spatial dimensions :} \quad \vec{\ell} = \vec{x} \times \vec{p}) \\
\delta x^a = -\nu^a t & \Rightarrow & (???) \quad G^a = -p^a t + m x^a
\end{array}
\tag{6.5}
$$

The only quantity that perhaps deserves some comment / explanation is G^a (which, as far as I know, does not have a standard name - perhaps because it turns out to be a not particularly interesting quantity): first of all, calculating the time-derivative of G^a and using $p^a = m\dot{x}^a$, one sees that $\dot{G}^a = -\dot{p}^a t$ so that conservation of G is implied by momentum conservation; secondly, rewriting the definition of G^a as

$$
x^a(t) = (p^a/m)t + (G^a/m) \tag{6.6}
$$

shows that $G^a = mx^a(0)$ is esentially just the rather tautological conserved "initial position" of the particle.

All of this generalises in a straightforward way to a closed N-particle system defined to be a system of N-particles, with coordinates $\vec{x}_A$, not subject to any external forces and interacting via a potential of the form $U = U(|\vec{x}_A - \vec{x}_B|)$,

$$
\mathcal{L} = T - U = \sum_{A=1}^{N} \tfrac{1}{2} m_A \dot{\vec{x}}_A^2 - U(|\vec{x}_A - \vec{x}_B|) \quad . \tag{6.7}
$$

Since the potential is strictly invariant under all Galilei transformations, and the kinetic term is like that of a single particle (quasi-)invariant as well, one again finds conserved quantities, the total energy, total momentum, total angular momentum and "total G^a",

$$
E = T + U \quad , \quad P^a = \sum_A p^a_A \quad , \quad L^{ab} = \sum_A \ell^{ab}_A \quad , \quad G^a = \sum_A G^a_A \quad , \tag{6.8}
$$

and G^a is related to the initial position of the center of mass coordinate $\vec{X} = \sum_A m_A \vec{x}_A / \sum_A m_A$.

7 Equivalent Lagrangians / Lagrangian Gauge Transformations

Two Lagrangians $\mathcal{L}_{1,2}$ that differ by a total time-derivative lead to the same Euler-Lagrange equations,

$$
\mathcal{L}_2(q^a, \dot{q}^a, t) = \mathcal{L}_1(q^a, \dot{q}^a, t) + \frac{d}{dt} G(q^a, t) \quad \Rightarrow \quad (\mathcal{E}_a)_2 = (\mathcal{E}_a)_1 \quad . \tag{7.1}
$$

This is evident from Hamilton's principle, since the corresponding actions only differ by a boundary term, but it is also straightforward and instructive to check this explicitly. It is sufficient to verify that the Euler-Lagrange equations for a Lagrangian that is itself a total derivative are identically satisfied,

$$
\mathcal{L}(q^a(t), \dot{q}^a(t), t) = \frac{d}{dt} G(q^a(t), t) \quad \Rightarrow \quad \mathcal{E}_a = 0 \quad \text{identically} \quad . \tag{7.2}
$$

The proof of this is straightforward. First of all, such a Lagrangian depends at most linearly on the velocities $\dot{q}^a(t)$,

$$\mathcal{L} = \frac{d}{dt}G = \frac{\partial G}{\partial t} + \frac{\partial G}{\partial q^b}\dot{q}^b \quad . \tag{7.3}$$

so that

$$\frac{\partial \mathcal{L}}{\partial \dot{q}^a} = \frac{\partial G}{\partial q^a} \quad . \tag{7.4}$$

Moreover, one has the general result that the operators d/dt and $\partial/\partial q^a$ commute with each other when acting on functions $G(q^a, \dot{q}^a, \ldots, t)$, because d/dt acting on such functions can be written as a sum of operators all of which manifestly commute with $\partial/\partial q^a$,

$$\frac{d}{dt}G = \frac{\partial}{\partial t}G + \dot{q}^a\frac{\partial}{\partial q^a}G + \ldots \quad \Rightarrow \quad \left[\frac{d}{dt}, \frac{\partial}{\partial q^a}\right]G = 0 \quad . \tag{7.5}$$

These two identities now imply that

$$\left(\frac{\partial}{\partial q^a} - \frac{d}{dt}\frac{\partial}{\partial \dot{q}^a}\right)\frac{d}{dt}G = \left(\frac{\partial}{\partial q^a}\frac{d}{dt} - \frac{d}{dt}\frac{\partial}{\partial q^a}\right)G = 0 \quad , \tag{7.6}$$

as was to be shown.

One can also show the converse: if a Lagrangian L gives rise to Euler-Lagrange equations that are identically satisfied then (locally) the Lagrangian is a total derivative. The proof is simple. Assume that $L(q, \dot{q}; t)$ satisfies

$$\frac{\partial L}{\partial q^a} \equiv \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} \tag{7.7}$$

identically. The left-hand side does evidently not depend on the acceleration $\ddot{q}^b$. The right-hand side, on the other hand, will in general depend on $\ddot{q}^b$ - unless L is at most linear in $\dot{q}^b$. Thus a necessary condition for L to give rise to identically satisfied Euler-Lagrange equations is that it is of the form

$$L(q, \dot{q}; t) = L^0(q; t) + L^1_b(q; t)\dot{q}^b \quad . \tag{7.8}$$

Therefore

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^a} = \frac{d}{dt}L^1_a = \frac{\partial L^1_a}{\partial t} + \frac{\partial L^1_a}{\partial q^b}\dot{q}^b \tag{7.9}$$

and

$$\frac{\partial L}{\partial q^a} = \frac{\partial L^0}{\partial q^a} + \frac{\partial L^1_b}{\partial q^a}\dot{q}^b \quad . \tag{7.10}$$

Thus the requirement that the Euler-Lagrange equations are identities lead to the conditions

$$\frac{\partial L^1_a}{\partial t} = \frac{\partial L^0}{\partial q^a} \quad , \quad \frac{\partial L^1_a}{\partial q^b} = \frac{\partial L^1_b}{\partial q^a} \quad . \tag{7.11}$$

These are precisely the integrability conditions which guarantee the local existence of a function $G(q, t)$ such that

$$L^0 = \frac{\partial G}{\partial t} \quad , \quad L^1_b = \frac{\partial G}{\partial q^b} \quad , \tag{7.12}$$

and therefore

$$L = L^0 + L^1_b\dot{q}^b = \frac{d}{dt}G \quad , \tag{7.13}$$

as was to be shown.

This freedom in the choice of $\mathcal{L}$,

$$\mathcal{L}_1 \to \mathcal{L}_2 = \mathcal{L}_1 + \frac{d}{dt}G \tag{7.14}$$

is sometimes referred to as a Lagrangian (or mechanical) gauge transformation, not only because (like a gauge transformation) it leaves the equations of motion invariant, but also since its effect on the canonical momenta

$$p_a = \frac{\partial \mathcal{L}}{\partial \dot{q}^a} \tag{7.15}$$

is

$$(p_a)_1 \to (p_a)_2 = (p_a)_1 + \frac{\partial G}{\partial q^a} \equiv (p_a)_1 + \partial_a G \ \ , \tag{7.16}$$

which is reminiscent of that of standard Maxwell gauge transformations.

8 Lagrangian Gauge Transformations and the Noether Theorem

If the Lagrangian $\mathcal{L}_1$ is quasi-invariant,

$$\delta_s \mathcal{L}_1 = \frac{d}{dt}F_1 \ \ , \tag{8.1}$$

then also $\mathcal{L}_2 = \mathcal{L}_1 + (d/dt)G$ is quasi-invariant because (by definition / construction) the variation commutes with the total time-derivative,

$$\delta_s \mathcal{L}_2 = \frac{d}{dt}F_1 + \delta_s \frac{d}{dt}G = \frac{d}{dt}(F_1 + \delta_s G) \equiv \frac{d}{dt}F_2 \ \ , \tag{8.2}$$

with

$$F_2 = F_1 + \delta_s G = F_1 + (\partial_a G)\delta_s q^a \ \ . \tag{8.3}$$

According to (5.6) the conserved Noether charge associated to the quasi-invariance of $\mathcal{L}_1$ is

$$P_1 = (p_a)_1 \delta_s q^a - F_1 \ \ , \tag{8.4}$$

and that associated to the quasi-invariance of $\mathcal{L}_2$ is

$$P_2 = (p_a)_2 \delta_s q^a - F_2 = (p_a)_1 \delta_s q^a + (\partial_a G)\delta_s q^a - (F_1 + \delta_s G) = P_1 \ \ . \tag{8.5}$$

Thus (in mechanics!) the conserved Noether charge is invariant under Lagrangian gauge transformations.

9 Some Exotic Examples (with Scaling Symmetries)

1. Conformal Particle

   In its simplest incarnation this is described by an action for a a particle in one dimension moving in an inverse-square potential. Thus the action is, calling the single coordinate now $x = x(t)$,

   $$S = \int dt\, \mathcal{L} = \int dt \left(\frac{1}{2}m\dot{x}^2 - \frac{\alpha}{x^2}\right) \ \ . \tag{9.1}$$

   By inspection this action is invariant under

- time translations

$$\bar{t} = t + \lambda \quad , \quad \bar{x}(\bar{t}) = x(t) \tag{9.2}$$

- and scalings

$$\bar{t} = \lambda t \quad , \quad \bar{x}(\bar{t}) = \sqrt{\lambda} x(t) \ . \tag{9.3}$$

The former leads to energy conservation, with conserved charge

$$\mathcal{H} = E = \frac{1}{2} m \dot{x}^2 + \frac{\alpha}{x^2} \ . \tag{9.4}$$

Infinitesimally, the scaling symmetry reads

$$\lambda = 1 + \epsilon \quad \Rightarrow \quad \bar{t} = t + \epsilon t \quad , \quad \bar{x}(\bar{t}) = (1 + \epsilon/2) x(t) \ . \tag{9.5}$$

Thus the corresponding symmetry variation is

$$X(t) = \epsilon t \quad , \quad Y(t) = \frac{1}{2} \epsilon x(t) \quad \Rightarrow \quad \delta x(t) = \frac{1}{2} \epsilon x(t) - \epsilon t \dot{x}(t) \ . \tag{9.6}$$

Under this variation, the Lagrangian changes as

$$\delta L = \frac{d}{dt}(-X\mathcal{L}) = \frac{d}{dt}(-\epsilon t \mathcal{L}) \tag{9.7}$$

in agreement with the general result (B.9),

$$\bar{\delta}\mathcal{L} = 0 \quad \Rightarrow \quad \delta\mathcal{L} = -\frac{d}{dt}(X\mathcal{L}) \ . \tag{9.8}$$

Therefore the corresponding conserved quantity is

$$P = p_a \delta q^a + X\mathcal{L} = p_a Y^a - X\mathcal{H} = \epsilon(\tfrac{1}{2} m x \dot{x} - tE) \ . \tag{9.9}$$

It is straightforward to verify that indeed this quantity is conserved for a solution to the equations of motion,

$$m\ddot{x} = \frac{2\alpha}{x^3} \quad \Rightarrow \quad \frac{d}{dt} P = 0 \ . \tag{9.10}$$

Using E and P, one can now algebraically solve for x as a function of t, with E and P playing the role of constants of integration. One finds

$$x(t) = \pm (mE)^{-1/2} \left(2(P + tE)^2 + \alpha m\right)^{1/2} \quad , \tag{9.11}$$

and one can (and should) verify explicitly that this solves the equations of motion.

2. Scale-invariant Time-Dependent (Inverted) Oscillator

   The action for a general time-dependent oscillator (with time-dependent frequency $\omega(t)$ but time-independent mass m) is

$$S = \int dt \left(\frac{m}{2} \dot{x}^2 - \frac{m\omega(t)^2}{2} x^2 \right) = \frac{m}{2} \int dt \left(\dot{x}^2 - \omega(t)^2 x^2 \right) \ . \tag{9.12}$$

   When ω is time-independent, one has time translation invariance and the corresponding conserved energy. Generically, when $\omega = \omega(t)$ is time-dependent, there is no conserved quantity. However, for the special choice

$$\omega(t)^2 = \omega_0^2 t^{-2} \tag{9.13}$$

(one can also consider $\omega_0^2 < 0$) this action again has a scaling symmetry (actually the same scaling symmetry as the one in the previous example), namely

$$\bar{t} = \lambda t \quad , \quad \bar{x}(\bar{t}) = \sqrt{\lambda} x(t) \quad \Rightarrow \quad \delta x(t) = \frac{1}{2}\epsilon x(t) - \epsilon t \dot{x}(t) \ . \tag{9.14}$$

Thus the conserved charge is (again as in the previous example)

$$P = \epsilon(\tfrac{1}{2} m x \dot{x} - tE) \ , \tag{9.15}$$

the difference being that in this case the energy

$$E = \tfrac{1}{2} m(\dot{x}^2 + \omega_0^2 x^2/t^2) \tag{9.16}$$

is *not* conserved (while P nevertheless *is* conserved, as one can easily verify).

3. "Damped" Harmonic Oscillator

The action is

$$S = \tfrac{1}{2}\int d\tau \left(m(y')^2 - ky^2\right) \mathrm{e}^{\tau} \equiv \int d\tau \mathcal{L}_0 \mathrm{e}^{\tau} \ , \tag{9.17}$$

with constant m, k, and with $y' = dy/d\tau$, and with $\mathcal{L}_0$ the Lagrangian of a completely standard time-independent harmonic oscillator. This turns out to be the previous example in disguise, and we will come back to this below. Here we will look at this model in its own right.

Note first of all that the equations of motion are of the standard harmonic oscillator type, with a damping term $\sim y' = dy/d\tau$,

$$my'' + my' + ky = 0 \ . \tag{9.18}$$

By inspection again, this action has a symmetry involving a translation of τ,

$$\bar{\tau} = \tau + \lambda \quad , \quad \bar{y}(\bar{\tau}) = \mathrm{e}^{-\lambda/2} y(\tau) \ . \tag{9.19}$$

By the (by now hopefully standard) rules, this leads to

(a) the symmetry variation (setting $\epsilon = 1$)

$$\delta y = -y' - y/2 \ . \tag{9.20}$$

(b) the variation of the Lagrangian ($X = 1$)

$$\delta \mathcal{L} = \frac{d}{dt}(-\mathcal{L}) \tag{9.21}$$

(c) the conserved charge

$$P = p_y \delta y + \mathcal{L} = -\tfrac{1}{2} m y y' \mathrm{e}^{\tau} - E \ . \tag{9.22}$$

with (non-conserved) energy

$$E = \tfrac{1}{2}\left(m(y')^2 + ky^2\right) \mathrm{e}^{\tau} \ . \tag{9.23}$$

4. "Damped" Harmonic Oscillator Revisited

   Taking our cue from the overall $\exp\tau$-factor of the Lagrangian, we eliminate it by introducing the new time coordinate t via

$$t = \mathrm{e}^{\tau} \quad \Rightarrow \quad d\tau = t^{-1}dt \quad \Rightarrow \quad \frac{dy}{d\tau} = t\frac{dy}{dt} \quad . \tag{9.24}$$

   We also introduce

$$x(t) = ty(t) \quad \Rightarrow \quad \dot{x} = t\dot{y} + y \quad \Rightarrow \quad \dot{x}^2 = t^2\dot{y}^2 + \frac{d}{dt}(ty^2) \quad . \tag{9.25}$$

   Thus, up to a total derivative the action is

$$S \to \frac{1}{2}\int dt\left(m\dot{x}^2 - kx^2/t^2\right) \quad , \tag{9.26}$$

   which is precisely of the scale-invariant oscillator form discussed above.

   In particular, in terms of y and τ, the free particle equation of motion can be made to look like that of a particle with friction,

$$\ddot{x} = 0 \quad \Leftrightarrow \quad y'' + y' = 0 \quad . \tag{9.27}$$

A Noether Theorem: Traditional Proof

The usual proof of the Noether theorem for mechanics that one can find in the literature (if done (a) correctly and (b) in this generality - neither (a) nor (b) is guaranteed) consists of the following steps:

1. One considers transformations the time coordinate t (the independent variable) and of the paths $q^a(t)$ (the dependent variables)

$$t \to \bar{t}(t, q^a(t), \dot{q}^a(t), \ldots) \quad , \quad q^a(t) \to \bar{q}^a(t, q^a(t), \dot{q}^a(t), \ldots) \equiv \bar{q}^a(\bar{t}) \tag{A.1}$$

   (thinking of $\bar{t}$ as the new independent variable) that depend on a continuous parameter (more generally, one considers transformations that form a finite-dimensional group) and that leave the action invariant, in the sense that

$$\int_{\bar{t}_1}^{\bar{t}_2} d\bar{t}\;\mathcal{L}(\bar{q}^a, \frac{d\bar{q}^a}{d\bar{t}}, \bar{t}) = \int_{t_1}^{t_2} dt\;\mathcal{L}(q^a, \frac{dq^a}{dt}, t) \tag{A.2}$$

   (perhaps up to boundary / total derivative terms).

2. Infinitesimally, one can write such transformations as

$$\begin{aligned} t \to \bar{t} &= t + \epsilon X(q^b(t), \dot{q}^b(t), \ldots, t) \equiv t + \bar{\delta}t \\ q^a(t) \to \bar{q}^a(\bar{t}) &= q^a(t) + \epsilon Y^a(q^b(t), \dot{q}^b(t), \ldots, t) \equiv q^a(t) + \bar{\delta}q^a(t) \quad . \end{aligned} \tag{A.3}$$

   Note that, for $X \neq 0$, $\bar{\delta}t$ and $\bar{\delta}q^a(t)$ are *not* variations in the sense defined in section 3: for a variation, t is not transformed, and $\delta q^a(t)$ is the difference of two paths at the same point t. Here t is transformed, and

$$\bar{\delta}q^a(t) = \bar{q}^a(\bar{t}) - q^a(t) \tag{A.4}$$

is the difference of two paths at two different points t and $\bar{t}$, and therefore in particular does not satisfy the characteristic identity (2.5). As a consequence, also the VME (3.2) does not hold for these infinitesimal transformations.

3. The condition for invariance (up to a total derivative) can now be written as

$$\bar{\delta}\mathcal{L} \equiv \frac{d}{d\epsilon}\left[\frac{d\bar{t}}{dt}\,\mathcal{L}(\bar{q}^a, \frac{d\bar{q}^a}{d\bar{t}}, \bar{t})\right]_{\epsilon=0} \stackrel{!}{=} \frac{d}{dt}f(q^a, \ldots, t) \quad . \tag{A.5}$$

   Note that, for the same reasons as outlined above, $\bar{\delta}\mathcal{L}$ is not, in general, a variation of $\mathcal{L}$ in the sense defined here. So for now, we just treat is as an abbreviation for the 1st-order (linear in ϵ, or linear in X and Y^a) change of $\mathcal{L}dt$ under the above transformation.

4. Using

$$d\bar{t} = dt + \epsilon\frac{dX}{dt}dt \equiv (1 + \epsilon\dot{X})dt \tag{A.6}$$

   and (to first order in ϵ)

$$\frac{d\bar{q}^a(\bar{t})}{d\bar{t}} = \frac{1}{1+\epsilon\dot{X}}\frac{d}{dt}(q^a(t) + \epsilon Y^a) = \frac{dq^a}{dt} + \epsilon(\dot{Y}^a - \dot{X}^a\dot{q}^a) \tag{A.7}$$

   one finds

$$\bar{\delta}\mathcal{L} = \dot{X}\mathcal{L} + \frac{\partial\mathcal{L}}{\partial q^a}Y^a + \frac{\partial\mathcal{L}}{\partial\dot{q}^a}(\dot{Y}^a - \dot{X}\dot{q}^a) + X\frac{\partial\mathcal{L}}{\partial t} \quad . \tag{A.8}$$

5. Using

$$\frac{\partial\mathcal{L}}{\partial t} = \frac{d}{dt}\left(\mathcal{L} - \frac{\partial\mathcal{L}}{\partial\dot{q}^a}\dot{q}^a\right) - \left(\frac{\partial\mathcal{L}}{\partial q^a} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}^a}\right)\dot{q}^a \equiv -\frac{d}{dt}\mathcal{H} - \mathcal{E}_a\dot{q}^a \tag{A.9}$$

   (this is the identity underlying the usual proof of energy conservation for a time-independent system), this can be written as

$$\bar{\delta}\mathcal{L} = \frac{d}{dt}(p_aY^a - X\mathcal{H}) + \mathcal{E}_a(Y^a - X\dot{q}^a) \quad . \tag{A.10}$$

6. Thus, if $\bar{\delta}\mathcal{L}$ is a total derivative, then there is a conserved quantity

$$\bar{Q} = p_aY^a - X\mathcal{H} - f: \quad \mathcal{E}_a = 0 \quad \Rightarrow \quad \frac{d}{dt}\bar{Q} = 0 \quad . \tag{A.11}$$

While the above argument is correct, it gives undue prominence to transformations of t (and the integration measure dt). These should not play an essential role because integration is defined in such a way that it is invariant under transformations of the integration variable. And the above procedure leads to totallly unnecessary complications and confusions, in particular when the Noether theorem is applied to action principles for fields (where the integrals are not over the worldline of a particle but over a region of space-time).

B Noether Theorem: Equivalence with the Variational Version

Therefore, let us look at the above procedure in a different way. First of all, we make the crucial observation that (for $X \neq 0$) the infinitesimal change ϵY^a of q^a as defined in (A.3),

$$\bar{t} = t + \epsilon X \quad , \quad \bar{q}^a(\bar{t}) = q^a(t) + \epsilon Y^a \quad , \tag{B.1}$$

is *not* a variation, because it is a difference between two paths at two different points, not at the same point,

$$\epsilon Y^a = \bar{q}^a(\bar{t}) - q^a(t) \neq \delta q^a(t) \quad . \tag{B.2}$$

As a consequence for the thus defined quantity

$$\bar{\delta} q^a(t) = \bar{q}^a(\bar{t}) - q^a(t) \tag{B.3}$$

one has

$$\frac{d}{dt}\bar{\delta} q^a(t) \neq \bar{\delta}\frac{d}{dt} q^a(t) \quad . \tag{B.4}$$

To rectify this, we substitute $\bar{t}$ and Taylor expand to find (to order ϵ),

$$\bar{q}^a(\bar{t}) = \bar{q}^a(t) + \epsilon X \frac{d\bar{q}^a}{d\bar{t}} + \ldots = \bar{q}^a(t) + \epsilon X \dot{q}^a + \ldots \quad . \tag{B.5}$$

Thus we have the original paths $q^a(t)$, and we have also obtained new paths $\bar{q}^a(t)$, and their difference (at the same point)

$$\bar{q}^a(t) - q^a(t) = \bar{q}^a(\bar{t}) - q^a(t) - \epsilon X \dot{q}^a(t) = \epsilon(Y^a - X\dot{q}^a(t)) \tag{B.6}$$

is a variation. Dropping the ϵ (i.e. we understand that when we write δ, we mean something infinitesimal - alternatively, we absorb ϵ into the definition of X and Y^a), the transformation (A.3) is thus equivalent to the *variation*

$$\delta q^a(t) = Y^a - X\dot{q}^a(t) \tag{B.7}$$

(with t not transformed).

Now let us go back to identity (A.10). First of all, in the last term we already recognise the variation that we have just defined, and we can also rewrite the other terms to arrive at

$$\bar{\delta}\mathcal{L} = \frac{d}{dt}(X\mathcal{L}) + \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^a}(Y^a - X\dot{q}^a)\right) + \mathcal{E}_a(Y^a - X\dot{q}^a) \quad . \tag{B.8}$$

We see that the 2nd and 3rd terms are precisely the terms that appear in the VME (3.2), and therefore

$$\delta\mathcal{L} = \bar{\delta}\mathcal{L} - \frac{d}{dt}(X\mathcal{L}) \quad . \tag{B.9}$$

We see that these two procedures to obtain the infinitesimal change in the action only differ by a total time derivative. In particular, the (true) variation is a total derivative if and only if $\bar{\delta}\mathcal{L}$ is a total derivative,

$$\bar{\delta}\mathcal{L} = \frac{d}{dt}f \quad \Leftrightarrow \quad \delta\mathcal{L} = \frac{d}{dt}F = \frac{d}{dt}(f - X\mathcal{L}) \quad , \tag{B.10}$$

and the conserved quantity (5.6) predicted by the variational version of Noether's theorem is

$$P_\delta = p_a \delta q^a - \mathcal{F}_\delta = p_a(Y^a - X\dot{q}^a) - (f - X\mathcal{L}) = p_a Y^a - X\mathcal{H} - f = \bar{Q} \tag{B.11}$$

agrees precisely with that previously derived in (A.11).

Thus the criterion for the transformation (A.3) to be a symmetry of the action is equivalent to the requirement that the variation (B.7) is a (quasi-)symmetry of $\mathcal{L}$.

In particular, when applied to time translations,

$$\bar{t} = t + \epsilon \quad , \quad \bar{q}^a(\bar{t}) = q^a(t) \ , \tag{B.12}$$

this leads to the variation (again dropping the ϵ)

$$\delta q^a(t) = -\epsilon \dot{q}^a(t) \ , \tag{B.13}$$

as anticipated in the discussion around (5.13).

C Aspects of Variational Calculus

In section 4 we have seen that the Euler-Lagrange equations are the necessary conditions for functions $q^a(t)$ to extremise a functional of the form $S[q] = \int dt \mathcal{L}(q, \dot{q}, t)$ (necessary but not necessarily sufficient because, as mentioned, a stationary configuration could also be a "saddle-point"). Clearly this is a very general result that transcends its origin in the mechanics of point particles, and other situations or problems where one is required to find functions that extremise a given functional abound in mathematics and physics: minimising the length / area / volume of a configuration, minimising the time it takes to travel between 2 points, finding a static configuration of minimal (potential) energy, etc. etc.

To mentally accomplish this shift away from purely mechanical problems, we now denote the *independent variable(s)* by x or x^i (instead of t), and the *dependent variable(s)* by $y = f(x)$ or $y_A = f_A(x)$ etc. (instead of $q^a(t)$). Let us start by considering the (already very rich and interesting) case of one independent variable x and one dependent variable $y = f(x)$, which we can visualise as describing a curve in the (x, y)-plane. The analog of the action and Lagrangian is then a functional of the form

$$J[y] = \int_{x_1}^{x_2} dx \ F[y(x), y'(x), x] \ , \tag{C.1}$$

and the task of *variational calculus* is to find the curve (or condition for the curve) that extremises the functional $J[y]$, typically with fixed boundary conditions $y_1 = f(x_1), y_2 = f(x_2)$.

Here are some examples:

1. The simplest non-trivial example of this kind is the functional which assigns to a curve $y(x)$ the *length* of the curve between the points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$,

   $$L[y] = \int_{P_1}^{P_2} ds \equiv \int_{P_1}^{P_2} \sqrt{dx^2 + dy^2} = \int_{x_1}^{x_2} dx \sqrt{1 + y'(x)^2} \ , \tag{C.2}$$

   and it allows one to pose (and answer!) the question which curve between two points in the (x, y)-plane mnimises the distance along the curve (as we will see below, none too surprisingly but reassuringly the answer turns out to be "a straight line").

2. As a variation (pun not intended) of this theme, consider the question which curve minimises the area of the surface that one obtains by rotating the curve around the x-axis or

the y-axis. In the former case, the surface is formed by circles of radius $y(x)$, so that the area is

$$A_1[y] = 2\pi \int_{x_1}^{x_2} dx\; y(x)\sqrt{1+y'(x)^2} \quad , \tag{C.3}$$

while in the latter case it is formed by circles of radius x so that the area is

$$A_2[y] = 2\pi \int_{x_1}^{x_2} dx\; x\sqrt{1+y'(x)^2} \quad . \tag{C.4}$$

3. Another simple example arises from the question what would have to be the shape of a shute or slide between 2 points P_i that minimises the time it takes for a particle in a constant gravitational field to slide from P_1 to P_2 (with zero initial velocity and zero friction, say).[1] To find the functional for this problem, let us (wlog) assume that $y_1 = 0$. With the initial condition $v_1 = v(x_1) = 0$, energy conservation tells us that

$$mv^2/2 - mgy = mv_1^2/2 - mgy_1 = 0 \tag{C.5}$$

all along the path, with $v = ds/dt$. Thus the time can be calculated from

$$T[y] = \int_{t_1}^{t_2} dt = \int_{P_1}^{P_2} ds/v = \int_{x_1}^{x_2} dx\sqrt{1+y'(x)^2}/\sqrt{2gy(x)} \quad . \tag{C.6}$$

4. As a final example, consider the question what is the shape of a chain (or high-voltage power line) of fixed length L_0 (and constant mass per unit length m) suspended between 2 points (pylons) with coordinates (x_i, y_i) in a constant gravitational field. In that case, one needs to minimise the potential energy

$$E[y] = \int_{x_1}^{x_2} dx\; mgy(x)\sqrt{1+y'(x)^2} \tag{C.7}$$

subject to the condition that $L[y] = L_0$ is fixed.

These examples should suffice to illustrate the ubiquity and importance of variational problems in mathematics and physics.

To determine the equation that $y(x)$ has to satisfy to extremise the functional $J[y]$, we could now proceed as in section 4 and simply consider variations $y(x) \to y(x) + \delta y(x)$ with $\delta y(x_1) = \delta y(x_2) = 0$ and demand stationarity of $J[y]$, i.e. $\delta J[y] = 0$ for all $\delta y(x)$. For present purposes it is perhaps more instructive to rephrase this as follows: clearly a necessary condition for $y(x)$ to be a (local) extremum of $J[y]$ is that the (ordinary) function of ϵ one obtains by replacing $y(x)$ by $y(x) + \epsilon\eta(x)$ (for some arbitrary function $\eta(x)$ with $\eta(x_1) = \eta(x_2) = 0$) has a critical point at $\epsilon = 0$,

$$J_\eta(\epsilon) \equiv J[y(x) + \epsilon\eta(x)] \quad , \quad \frac{d}{d\epsilon}J_\eta(\epsilon)|_{\epsilon=0} \stackrel{!}{=} 0 \quad . \tag{C.8}$$

[1] This *brachistochrone* (shortest time) problem is also of historical importance: posed by Jakob Bernoulli in 1696, it stimulated the development of variational calculus, Euler subsequently showing how such problems could be reduced to differential equations, the equations that we have so far (mixing variational calculus and analytical mechanics language) referred to as the *Euler-Lagrange equations*.

Performing this differentiation, one finds

$$\int_{x_1}^{x_2} dx \left(\frac{\partial F}{\partial y(x)} \eta(x) + \frac{\partial F}{\partial y'(x)} \eta'(x) \right) = 0 \quad \forall\, \eta(x) \ . \tag{C.9}$$

Integrating by parts the 2nd term (i.e. performing again the step that led to the VME (3.2), now inside the integral) and using the boundary conditions on $\eta(x)$, one finds that this can be rewritten as

$$\int_{x_1}^{x_2} dx \left(\frac{\partial F}{\partial y(x)} - \frac{d}{dx} \frac{\partial F}{\partial y'(x)} \right) \eta(x) = 0 \quad \forall\, \eta(x) \quad \Leftrightarrow \quad \frac{\partial F}{\partial y(x)} - \frac{d}{dx} \frac{\partial F}{\partial y'(x)} = 0 \ . \tag{C.10}$$

None too surprisingly, these are precisely the Euler-Lagrange equations for a 1-dimensional system (with $t \to x, q(t) \to y(x)$).

Applied to Example 1 (the shortest path bewteen 2 points in the place), this gives us the condition

$$\frac{d}{dx} \frac{y'(x)}{\sqrt{1 + y'(x)^2}} = 0 \quad \Rightarrow \quad y'(x) = \text{const.} \ , \tag{C.11}$$

so reassuringly this is the equation of a straight line, $y(x) = ax + b$, with a, b determined by $y(x_1) = y_1$ and $y(x_2) = y_2$.

Before blindly applying these equations to some other examples, let us oberve that, as in the case of mechanics, we have conserved quantities (1st integrals of the Euler-Lagrange equations) if e.g. the integrand $F[y, y', x]$ does not depend explicitly either on y (cyclic coordinate, "momentum conservation") or on x ("energy conservation"),

$$\begin{aligned} F = F(y', x) \quad &\Rightarrow \quad \frac{\partial F}{\partial y'} = \text{const.} \\ F = F(y, y') \quad &\Rightarrow \quad y' \frac{\partial F}{\partial y'} - F = \text{const.} \end{aligned} \tag{C.12}$$

In particular, applied to Example 2, for $A_1[y]$ we have "energy conservation", leading to

$$y\sqrt{1 + y'^2} - yy'^2 / \sqrt{1 + y'^2} = c_1 \quad \Rightarrow \quad y' = \sqrt{1 - (y/c_1)^2} \ , \tag{C.13}$$

with solution

$$y(x) = c_1 \cosh((x - x_0)/c_1) \tag{C.14}$$

(with x_0 and c_1 to be determined by the boundary conditions). Here x_0 is the value of x at which $y(x)$ takes on its (absolute) minimum, but depending on the boundary conditions x_0 may or may not lie in the interval (x_1, x_2).

For $A_2[y]$, on the other hand, we have momentum conservation,

$$xy' / \sqrt{1 + y'^2} = c_2 \quad \Rightarrow \quad y' = c_2 / \sqrt{x^2 - c_2^2} \ , \tag{C.15}$$

with solution

$$y(x) = c_2 \cosh^{-1}(x/c_2) + y_0 \quad \Leftrightarrow \quad x(y) = c_2 \cosh((y - y_0)/c_2) \ . \tag{C.16}$$

Reassuringly, the functions that minimise the areas one obtains upon rotation around the x-axis or y-axis are related by $x \leftrightarrow y$ ...

More fun with soap bubbles, catenoids, brachistochrones etc. can be had with these and other examples, but also more pain with solving the differential (ot some transcendental algebraic) equations, so this shall suffice.

I will close this section with some remarks on generalisations:

1. If one has several independent variables $x^i, i = 1, \dots, D$ and dependent variables $y_A(x^i)$, with
$$F = F(y_A(x^i), \partial y_A(x^i)/\partial x^j, x^k) \tag{C.17}$$
and a functional of the form
$$J[y_A] = \int d^D x\, F(y_A(x^i), \partial y_A(x^i)/\partial x^j, x^k) \ , \tag{C.18}$$
then by the same procedure one finds the Euler-Lagrange equations
$$\frac{\partial F}{\partial y_A} - \sum_i \frac{d}{dx^i} \frac{\partial F}{\partial(\partial y_A/\partial x^i)} = 0 \ , \tag{C.19}$$
where the symbol d/dx^i means that the operator acts on both the implicit and the explicit x^i-dependence of F, but still acts as a partial derivative in the sense that the other independent variables are to be held fixed.

2. If one has problems with supplementary conditions / constraints of the kind that appears in Example 4 (an integrated constraint, *isoperimetric boundary conditions*), then the way to take them into acccount is via the method of *Lagrange multipliers*. In the case at hand, this amounts to looking at the functional
$$J[y] = E[y] + \lambda L[y] \ , \tag{C.20}$$
with a real parameter λ. The solution now depends on 3 parameters, λ and the constants of integration, and these are then to be chosen in such a way that the boundary conditions and the constraint are satisfied. Rescaling $E[y] \to E[y]/mg$ (which has no influence on the curves that extremise it), explicitly $J[y]$ is
$$J[y] = \int_{x_1}^{x_2} dx\ (y(x) + \lambda)\sqrt{1 + y'(x)^2} \ . \tag{C.21}$$
Note that this differs from the problem for $A_1[y]$ solved above only by the constant shift $y(x) \to y(x) + \lambda$, so from (C.14) we can immediately write down the general solution
$$y(x) = c_1 \cosh((x - x_0)/c_1) - \lambda \ . \tag{C.22}$$
One can now calculate
$$\sqrt{1 + y'(x)^2} = \cosh((x - x_0)/c_1) \ , \tag{C.23}$$
and thus the length $L[y]$ for this class of solutions is
$$L[y] = \int_{x_1}^{x_2} dx \sqrt{1 + y'(x)^2} = c_1 \sinh((x - x_0)/c_1)|_{x_1}^{x_2} \ . \tag{C.24}$$
The 3 equations $L[y] = L_0$ and $y(x_i) = y_i$ now determine the 3 constants λ, c_1, x_0 in terms of L_0, y_i (via some in general transcendental equations).

3. Similarly, if one has holonomic constraints of the type $G_\alpha(y,x) = 0$, then these can be taken into account by introducing Lagrange multipliers $\lambda_\alpha(x)$ (these are now functions of the coordinate x as they need to be able to impose the constraints "pointwise") and modifying the integrand F of $J[y]$ to

$$F[y, y', x] \to \tilde{F}[y, \lambda, y', x] = F[y, y', x] + \sum \lambda_\alpha(x) G_\alpha(y,x) \ . \tag{C.25}$$

The Euler-Lagrange equations for $y(x)$ now take the form

$$\frac{\partial F}{\partial y(x)} - \frac{d}{dx}\frac{\partial F}{\partial y'(x)} = -\sum_\alpha \lambda_\alpha \frac{\partial G_\alpha}{\partial y(x)} \ , \tag{C.26}$$

(which should be familiar as the Lagrange equations of the 1st kind), while the Euler-Lagrange equations for the $\lambda_\alpha(x)$ simply impose the constraints $G_\alpha = 0$.

D Hamilton-Jacobi Identities from the VME

As a final variation of this theme of extracting information from the VME (3.2), we now consider the case where the 1st term on the rhs is zero, i.e. where the Euler-Lagrange equations are satisfied. Thus we restrict the $q^a(t)$ and their variations to the space of classical solutions. Since a classical solution is uniquely specified by the initial and final conditions $q(t_1) = q_1, q(t_2) = q_2$, in order to be able to perform variations in the space of solutions, we must be able to vary the q_i and / or the t_i. Let us choose to fix q_1 and t_1, and consider q_2 and t_2 as variables. Then the action, evaluated on a classical solution connecting the point q_1 at time t_1 to the point q_2 at time t_2 can be regarded as an ordinary function of the variables (q_2^a, t_2),

$$S[q] \to S(q_2, t_2) = \int_{t_1}^{t_2} dt\ \mathcal{L}(q, \dot{q}, t) \ . \tag{D.1}$$

Since $q(t)$ satisfies the Euler-Lagrange equations, and we have fixed $\delta q^a(t_1) = 0$, the integrated version (4.1) of the VME reduces to

$$\delta S(q_2, t_2) = p_a(t_2)\delta q^a(t_2) \quad \Rightarrow \quad p_a = \frac{\partial S(q,t)}{\partial q^a} \ , \tag{D.2}$$

i.e. the momentum can be calculated by differentiating the classical action with respect to the coordinates.

Furthermore, since one obviously has

$$\frac{d}{dt_2}S(q_2, t_2) = \frac{d}{dt_2}\int^{t_2} dt\ \mathcal{L}(q(t), \dot{q}(t), t) = \mathcal{L}(q(t_2), \dot{q}(t_2), t_2) \tag{D.3}$$

one deduces

$$\mathcal{L} = \frac{d}{dt}S(q,t) = \frac{\partial S}{\partial q^a}\dot{q}^a + \frac{\partial S}{\partial t} = p_a \dot{q}^a + \frac{\partial S}{\partial t} \ , \tag{D.4}$$

or

$$\frac{\partial S}{\partial t} = \mathcal{L} - p_a \dot{q}^a = -\mathcal{H} \ , \tag{D.5}$$

i.e. the time-derivative of the classical action is minus the energy.

<u>Remarks:</u>

1. Obviously, if one knows all the classical solutions, one can calculate the classical action $S(q,t)$. The content of the *Hamilton-Jacobi (HJ) theory* is the (non-trivial) converse to this, that in principle the solution to the non-linear partial differential HJ equation

$$\frac{\partial S}{\partial t} + \mathcal{H}(q^a, p_a = \frac{\partial S}{\partial q^a}) = 0 \tag{D.6}$$

   for the single function S provides one with the complete solution to the Euler-Lagrange equations. It is remarkable that there are situations in which indeed this is the method of choice to solve the equations, but by-and-large HJ theory is a god-awful mess.

2. Evidently the relations $p_a \to (\partial/\partial q^a)$ and $E \to -(\partial/\partial t)$ are reminiscent of quantum mechanics, and indeed HJ theory is the closest classical counterpart of Schrödinger wave mechanics, and the HJ equation can in some sense be regarded as the classical limit of the Schrödinger equation.