Introductory Variational Calculus on Manifolds

Ivo Terek*

Fix a (connected) differentiable manifold Q. Here we use the letter Q instead of the usual M because we regard our manifold as the configuration space of some mechanical system – setup where most of the applications of what we'll see here will occur. If $(q^1, \ldots, q^n)$ denote local coordinates on Q, the induced local coordinates on the tangent bundle TQ will be denoted by $(q^1, \ldots, q^n, v^1, \ldots, v^n)$, while the local coordinates on the cotangent bundle T^*Q will be denoted by $(q^1, \ldots, q^n, p_1, \ldots, p_n)$. Elements of TQ will be denoted by $(x, \boldsymbol{v})$, where $x \in Q$ and $\boldsymbol{v} \in T_xQ$, while elements of T^*Q will be denoted by (x, p), where $x \in Q$ and $\mathsf{p} \in T_x^*Q$. I will try to avoid it, but abuses of notation (although harmless) and mild simplifications might occur here and there[1]. I have deliberately tried to avoid symplectic geometry language (which would appear naturally here), to minimize pre-requisites for reading these notes. A short list of references that I used to understand this material is provided in the end, and you're welcome to contact me regarding corrections or mistakes here.

I'd also like to thank Caleb Dilsavor for the several discussions we had and for the helpful comments while writing this.

Contents

*terekcouto.1@osu.edu

[1]Mostly treating $(x, \boldsymbol{v})$ and $\boldsymbol{v}$ as the same thing, and similarly for (x, p) and p, whenever x is clear from context. I'll also almost always assume that our curves are "small", in the sense that their image is contained in a single coordinate chart (this is always true when Q is an open subset of $\mathbb{R}^n$, which is already a pretty useful case by itself). Or, in some examples, the functions considered might not be smooth along the whole manifold, but only on a certain open subset.

1 Basic definitions and examples

Definition 1.

- A *time-dependent Lagrangian* on Q is a smooth function $L\colon \mathbb{R}\times TQ\to\mathbb{R}$.
- A *time-dependent Hamiltonian* on Q is a smooth function $H\colon \mathbb{R}\times T^*Q\to\mathbb{R}$.

If there is no dependence on the time parameter $t\in\mathbb{R}$ (or, that is to say, if the domains considered are just TQ and T^*Q), we'll say that L or H are *autonomous*.

Example 2.

(1) If (Q,g) is a pseudo-Riemannian manifold and $V\colon \mathbb{R}\times Q\to\mathbb{R}$ is smooth, the function $L\colon \mathbb{R}\times TQ\to\mathbb{R}$ given by

$$L(t,x,\boldsymbol{v}) = \frac{1}{2}g_x(\boldsymbol{v},\boldsymbol{v}) - V(t,x)$$

is a time-dependent Lagrangian on Q. Lagrangians of this form in pseudo-Riemannian manifolds are called *natural*.

(2) If (Q,g) is a pseudo-Riemannian manifold and $V\colon \mathbb{R}\times Q\to\mathbb{R}$ is smooth, the function $H\colon T^*Q\to\mathbb{R}$ given by

$$H(t,x,\mathsf{p}) = \frac{1}{2}g^*_x(\mathsf{p},\mathsf{p}) + V(t,x)$$

is a time-dependent Hamiltonian on Q. Here, g^* is a fiber metric on T^*Q induced by g. More precisely, if $g^\sharp\colon T^*Q\to TQ$ is the musical isomorphism induced by g, we set $g^*_x(\mathsf{p},\mathsf{q}) = g_x(g^\sharp(\mathsf{p}),g^\sharp(\mathsf{q}))$, where $x\in Q$ and $\mathsf{p},\mathsf{q}\in T^*_xQ$.

(3) If $\alpha\in\Omega^1(Q)$ is fixed, we that $L_\alpha\colon TQ\to\mathbb{R}$ given by

$$L_\alpha(x,\boldsymbol{v}) = \alpha_x(\boldsymbol{v})$$

is an autonomous Lagrangian on Q.

(4) If $\boldsymbol{X}\in\mathfrak{X}(Q)$ is fixed, we have that $H^{\boldsymbol{X}}\colon T^*Q\to\mathbb{R}$ given by

$$H^{\boldsymbol{X}}(x,\mathsf{p}) = \mathsf{p}(\boldsymbol{X}_x)$$

is an autonomous Hamiltonian on Q.

We will focus on Lagrangians first.

Definition 3. Let $L\colon \mathbb{R}\times TQ\to\mathbb{R}$ be a time-dependent Lagrangian on Q. The *action functional* $\mathcal{A}_L$ associated to L is defined by

$$\mathcal{A}_L[x] = \int_a^b L(t,x(t),\dot{x}(t))\,\mathrm{d}t,$$

for any smooth curve $x\colon [a,b]\to Q$.

Example 4.

(1) Let (Q, g) be a Riemannian manifold. The *arc-length functional* $\mathscr{L}$ is given by

$$\mathscr{L}[x] = \int_a^b \|\dot{x}(t)\|_{x(t)} \, \mathrm{d}t,$$

for any curve $x\colon [a,b] \to Q$. We have that $\mathscr{L} = \mathscr{A}_L$, where $L\colon TQ \to \mathbb{R}$ is given by $L(x, v) = \|v\|_x$. In this case L is not smooth near the zero section of TQ, so this example does not quite belong in our setup here without some modifications. Instead, we consider the energy functional $\mathscr{E}$ given by

$$\mathscr{E}[x] = \frac{1}{2}\int_a^b g_{x(t)}(\dot{x}(t), \dot{x}(t)) \, \mathrm{d}t,$$

which also is significant when (Q, g) is pseudo-Riemannian instead of Riemannian. We have $\mathscr{E} = \mathscr{A}_L$, where this time $L(x, v) = g_x(v, v)/2$. We will see in due time that the geodesics of (Q, g) are precisely the critical points of $\mathscr{E}$. So we will need to define what is a critical point of an action functional.

(2) Let $\alpha \in \Omega^1(Q)$ and consider again $L_\alpha\colon TQ \to \mathbb{R}$ given by $L_\alpha(x, v) = \alpha_x(v)$. The action functional associated to L_α is given by

$$\mathscr{A}_{L_\alpha}[x] = \int_a^b \alpha_{x(t)}(\dot{x}(t)) \, \mathrm{d}t = \int_{x[a,b]} \alpha = \int_{[a,b]} x^*\alpha,$$

for any smooth $x\colon [a,b] \to Q$.

Definition 5. Let $L\colon \mathbb{R} \times TQ \to \mathbb{R}$ be a time-dependent Lagrangian on Q. We will say that a curve $x\colon [a,b] \to Q$ is *L-critical* or *a critical point of the action functional* $\mathscr{A}_L$ if for every variation $\widetilde{x}\colon (-\epsilon, \epsilon) \times [a,b] \to Q$ of x with fixed points, that is, satisfying

- $\widetilde{x}(0, t) = x(t)$ for all $t \in [a,b]$, and;
- $\widetilde{x}(s, a) = x(a)$ and $\widetilde{x}(s, b) = x(b)$ for all $s \in (-\epsilon, \epsilon)$,

we have that $\left.\dfrac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathscr{A}_L(\widetilde{x}(s, \cdot)) = 0$. We will denote this last derivative by $\mathrm{d}(\mathscr{A}_L)_x(\widetilde{x})$.

Remark. The notation $\mathrm{d}(\mathscr{A}_L)_x(\widetilde{x})$ is justified by the following interpretation: the domain of $\mathscr{A}_L$ is seen as a sort of infinite-dimensional manifold, and a vector tangent to it at the curve $x\colon [a,b] \to \mathbb{R}$ is the velocity vector of a curve of curves that starts at x. This curve of curves is a variation $\widetilde{x}$ with fixed endpoints, and the vector field

$$\boldsymbol{V}_{\widetilde{x}}(t) = \frac{\partial \widetilde{x}}{\partial s}(0, t)$$

along x is called the *variational vector field* of $\widetilde{x}$. Every vector field $\boldsymbol{V}$ along x is $\boldsymbol{V}_{\widetilde{x}}$ for some variation $\widetilde{x}$, and the variation $\widetilde{x}$ can be chosen with fixed endpoints, provided $\boldsymbol{V}(a) = \boldsymbol{V}(b) = 0$. The proof is simple: choose any Riemannian metric g on Q and consider the associated exponential map exp. By compactness of $[a,b]$, there is $\epsilon > 0$

for which $\widetilde{x}\colon (-\epsilon,\epsilon)\times[a,b]\to Q$ given by $\widetilde{x}(s,t)=\exp_{x(t)}(s\boldsymbol{V}(t))$ is well-defined and smooth. Indeed we have $\widetilde{x}(0,t)=\exp_{x(t)}(0)=x(t)$ for all $t\in[a,b]$, and also

$$\frac{\partial\widetilde{x}}{\partial s}(0,t)=\mathrm{d}(\exp_{x(t)})_0(\boldsymbol{V}(t))=\boldsymbol{V}(t),$$

since the derivative of the exponential map at the origin is the identity.

Proposition 6. *Let $L\colon \mathbb{R}\times TQ\to\mathbb{R}$ be a time-dependent Lagrangian. Then a smooth curve $x\colon[a,b]\to Q$ is L-critical if and only if for any natural coordinate system on TQ the* Euler-Lagrange equations *hold:*

$$\frac{\partial L}{\partial q^k}(t,x(t),\dot{x}(t))-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\right)=0,\qquad k=1,\ldots,n.$$

Remark. Some texts write these equations in a suggestive way, as $\nabla_q L-\dfrac{\mathrm{d}}{\mathrm{d}t}\nabla_v L=0$.

Proof: Let $\widetilde{x}\colon(-\epsilon,\epsilon)\times[a,b]\to Q$ be a smooth variation with fixed endpoints. If the path x is described in coordinates by

$$(x(t),\dot{x}(t))=(q^1(t),\ldots,q^n(t),v^1(t),\ldots,v^n(t)),$$

let the variation $\widetilde{x}$ be described by

$$\left(\widetilde{x}(s,t),\frac{\partial\widetilde{x}}{\partial t}(s,t)\right)=(\widetilde{q}^1(s,t),\ldots,\widetilde{q}^n(s,t),\widetilde{v}^1(s,t),\ldots,\widetilde{v}^n(s,t)).$$

Also regard L as a function of the $2n+1$ variables $(t,q^1,\ldots,q^n,v^1,\ldots,v^n)$. So we compute

$$\begin{aligned}
\mathrm{d}(\mathscr{A}_L)_x(\widetilde{x})&=\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\int_a^b L\left(t,\widetilde{x}(s,t),\frac{\partial\widetilde{x}}{\partial t}(s,t)\right)\,\mathrm{d}t\\
&=\int_a^b\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}L\left(t,\widetilde{x}(s,t),\frac{\partial\widetilde{x}}{\partial t}(s,t)\right)\,\mathrm{d}t\\
&=\int_a^b\left(\sum_{k=1}^n\frac{\partial L}{\partial q^k}(t,x(t),\dot{x}(t))\frac{\partial\widetilde{q}^k}{\partial s}(0,t)+\sum_{k=1}^n\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\frac{\partial\widetilde{v}^k}{\partial s}(0,t)\right)\,\mathrm{d}t\\
&\overset{(1)}{=}\int_a^b\left(\sum_{k=1}^n\frac{\partial L}{\partial q^k}(t,x(t),\dot{x}(t))\frac{\partial\widetilde{q}^k}{\partial s}(0,t)+\sum_{k=1}^n\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\frac{\partial^2\widetilde{q}^k}{\partial t\partial s}(0,t)\right)\,\mathrm{d}t\\
&\overset{(2)}{=}\int_a^b\left(\sum_{k=1}^n\frac{\partial L}{\partial q^k}(t,x(t),\dot{x}(t))\frac{\partial\widetilde{q}^k}{\partial s}(0,t)-\sum_{k=1}^n\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\right)\frac{\partial\widetilde{q}^k}{\partial s}(0,t)\right)\,\mathrm{d}t\\
&\qquad+\sum_{k=1}^n\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\frac{\partial\widetilde{q}^k}{\partial s}(0,t)\bigg|_a^b\\
&\overset{(3)}{=}\int_a^b\sum_{k=1}^n\left(\frac{\partial L}{\partial q^k}(t,x(t),\dot{x}(t))-\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\right)\right)\frac{\partial\widetilde{q}^k}{\partial s}(0,t)\,\mathrm{d}t,
\end{aligned}$$

where we use in order:

(1) $\widetilde{v}^k(s,t) = (\partial\widetilde{q}^k/\partial t)(s,t)$ for all s and t and that $\partial/\partial t$ and $\partial/\partial s$ commute.

(2) Integration by parts in the second group of terms.

(3) $(\partial\widetilde{q}^k/\partial s)(0,a) = (\partial\widetilde{q}^k/\partial s)(0,b) = 0$ for $k = 1,\ldots,n$, in view of the endpoint conditions $\widetilde{x}(s,a) = x(a)$ and $\widetilde{x}(s,b) = x(b)$ mentioned before.

However, variations can be chosen suitably to make each one of the summands which $(\partial\widetilde{q}^k/\partial s)(0,t)$ is multiplied against vanish. Thus

$$\frac{\partial L}{\partial q^k}(t,x(t),\dot{x}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\right) = 0, \qquad k = 1,\ldots,n.$$

The converse follows from (roughly) the same computation done above: for simplicity, assume that the images of $[a,b]$ under $\widetilde{x}_s = \widetilde{x}(s,\cdot)$ are covered by two coordinate charts. Take a point $c \in [a,b]$ such that all the $\widetilde{x}_s[a,c]$ are in the first chart and all the $\widetilde{x}_s[c,b]$ in the second (this can be arranged for reducing neighborhoods as needed). Write $\mathscr{A}[\widetilde{x}(s,\cdot)]$ as a sum of two integrals, and compute the s-derivative of each one. Each one vanishes because the Euler-Lagrange equations are satisfied in each chart. The same idea applies in the general case, since we can always cover the image $\widetilde{x}[(-\epsilon,\epsilon)\times[a,b]]$ by a finite number of charts (perhaps reducing ϵ is necessary), due to the compactness of $[a,b]$. □

Remark. The Euler-Lagrange equations are geometric, in the sense that their vanishing does not depend on the system of coordinates. Namely, if $(\hat{q}^1,\ldots,\hat{q}^n,\hat{v}^1,\ldots,\hat{v}^n)$ is another set of tangent coordinates for TQ, we have that

$$\frac{\partial L}{\partial \hat{q}^k}(x(t),\dot{x}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \hat{v}^k}(x(t),\dot{x}(t))\right) = \sum_{\ell=1}^n \left(\frac{\partial L}{\partial q^\ell}(x(t),\dot{x}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^\ell}(x(t),\dot{x}(t))\right)\right)\frac{\partial q^\ell}{\partial \hat{q}^k}(x(t)),$$

for all $k = 1,\ldots,n$.

Let's see some instances of how the Euler-Lagrange equations can appear in Mathematics and Physics:

Example 7 (Down to earth applications)**.** If $Q = \mathbb{R}$, then $T\mathbb{R} = \mathbb{R}^2$, and the Euler-Lagrange equations may be used to study problems such as: given $[a,b] \subseteq \mathbb{R}$ and numbers $y_a, y_b \in \mathbb{R}$, find a (smooth) function $y\colon [a,b] \to \mathbb{R}$ such that $y(a) = y_a$ and $y(b) = y_b$, minimizing (or maximizing) the integral

$$\mathscr{F}[y] = \int_a^b L(x,y(x),\dot{y}(x))\,\mathrm{d}x.$$

As a concrete example, let's study the variational problem

$$\begin{cases} \mathscr{F}[y] = \displaystyle\int_1^2 \frac{\dot{y}(x)^2}{x^3}\,\mathrm{d}x \\ y(1) = 0, \quad y(2) = 15. \end{cases}$$

The "time-dependent" Lagrangian here is given by $L(x, y, \dot{y}) = \dot{y}^2/x^3$, and even though it is not defined on the whole $\mathbb{R} \times T\mathbb{R} \cong \mathbb{R}^3$, we may still apply what we have seen above, since the integral is taken over an interval far from zero. We have

$$0 = \frac{\partial L}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial L}{\partial \dot{y}} = 0 - \frac{\mathrm{d}}{\mathrm{d}x}\frac{2\dot{y}(x)}{x^3} = -\frac{2\ddot{y}(x)x^3 - 6x^2\dot{y}(x)}{x^6} \implies \ddot{y}(x) - \frac{3}{x}\dot{y}(x) = 0.$$

This equation can be solved by reducing the order and using an integrating factor. We directly obtain that $y(x) = c_1x^4 + c_2$, where $c_1, c_2 \in \mathbb{R}$. The boundary conditions give $0 = c_1 + c_2$ and $15 = 16c_1 + c_2$, whence $c_1 = 1$ and $c_2 = -1$. Thus, the candidate to extremize $\mathscr{F}$ is $y(x) = x^4 - 1$. Deciding whether this is a (local?) minimum or maximum for $\mathscr{F}$ in general is difficult, but in this case we can do a somewhat direct analysis. Let $\eta_1, \eta_2 \colon [a, b] \to \mathbb{R}$ be smooth and satisfying all the endpoint conditions $\eta_1(a) = \eta_1(b) = \eta_2(a) = \eta_2(b) = 0$. Those conditions ensure that $y + s_1\eta_1 + s_2\eta_2$ is also a candidate to a solution to the variational problem considered, no matter the values of s_1 and s_2 chosen. So we compute

$$\begin{aligned}
\left.\frac{\partial^2}{\partial s_1 \partial s_2}\right|_{s_1=s_2=0} \mathscr{F}[y + s_1\eta_2 + s_2\eta_2] &= \left.\frac{\partial^2}{\partial s_1 \partial s_2}\right|_{s_1=s_2=0} \int_1^2 \frac{(y'(x) + s_1\eta_1'(x) + s_2\eta_2'(x))^2}{x^3}\,\mathrm{d}x \\
&= \left.\frac{\partial}{\partial s_1}\right|_{s_1=0} \int_1^2 \frac{2\eta_2'(x)(y'(x) + s_1\eta_1'(x))}{x^3}\,\mathrm{d}x \\
&= \int_1^2 \frac{2\eta_1'(x)\eta_2'(x)}{x^3}\,\mathrm{d}x.
\end{aligned}$$

Now, the quadratic form over the space of smooth $\eta \colon [a, b] \to \mathbb{R}$ with $\eta(a) = \eta(b) = 0$ given by

$$\eta \mapsto \int_1^2 \frac{2\eta'(x)^2}{x^3}\,\mathrm{d}x$$

is clearly positive-definite, so $y(x) = x^4 - 1$ maximizes $\mathscr{F}$. The fact that this "Hessian" was "equal" to $\mathscr{F}$ itself in this case was just a happy coincidence. Choosing other Lagrangians, we can look for surfaces of revolutions generated by graphs of functions $y = y(x)$ minimizing the area functional

$$\mathscr{A}[y] = \int_a^b 2\pi y(x)\sqrt{1 + \dot{y}(x)^2}\,\mathrm{d}x,$$

or understand the famous *brachistochrone* problem (studied by the Bernoullis, Newton, L'Hospital, among others): finding a curve in the plane joining two points which minimizes the time a particle takes to move from one point to another, only under influence of gravity – the functional that maps a curve to such time is

$$\mathscr{B}[y] = \int_a^b \sqrt{\frac{1 + \dot{y}(x)^2}{2gy(x)}}\,\mathrm{d}x,$$

where g is the standard acceleration caused by gravity.

Example 8 (Geodesics)**.** Let (Q, g) be a pseudo-Riemannian manifold. *Geodesics* of (Q, g) are the critical points of the energy functional

$$\mathscr{A}_L[x] = \frac{1}{2}\int_a^b g_{x(t)}(\dot{x}(t), \dot{x}(t))\,\mathrm{d}t,$$

where $L\colon TQ \to \mathbb{R}$ is given by $L(x, \boldsymbol{v}) = g_x(\boldsymbol{v}, \boldsymbol{v})/2$. Let's use this to show that geodesics described in coordinates by

$$(x(t), \dot{x}(t)) = (q^1(t), \ldots, q^n(t), v^1(t), \ldots, v^n(t))$$

are characterized by the differential equations

$$\frac{\mathrm{d}v^k}{\mathrm{d}t}(t) + \sum_{i,j=1}^n \Gamma_{ij}^k(x(t))v^i(t)v^j(t) = 0, \qquad k = 1, \ldots, n,$$

where the *Christoffel symbols* are given by

$$\Gamma_{ij}^k = \sum_{r=1}^n \frac{g^{kr}}{2}\left(\frac{\partial g_{ir}}{\partial q^j} + \frac{\partial g_{jr}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^r}\right)$$

and $(g^{ij})_{i,j=1}^n$ is the inverse matrix of $(g_{ij})_{i,j=1}^n$. Along the curve x, we have

$$L(x(t), \dot{x}(t)) = \frac{1}{2}\sum_{i,j=1}^n g_{ij}(x(t))v^i(t)v^j(t).$$

With this, we compute:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(x(t), \dot{x}(t))\right) - \frac{\partial L}{\partial q^k}(x(t), \dot{x}(t)) = \\
&\quad= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\sum_{i,j=1}^n g_{ij}(x(t))\delta_k^i v^j(t) + \frac{1}{2}\sum_{i,j=1}^n g_{ij}(x(t))v^i(t)\delta_k^j\right) - \frac{1}{2}\sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial q^k}(x(t))v^i(t)v^j(t) \\
&\quad= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{i=1}^n g_{ik}(x(t))v^i(t)\right) - \frac{1}{2}\sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial q^k}(x(t))v^i(t)v^j(t) \\
&\quad= \sum_{i,j=1}^n \frac{\partial g_{ik}}{\partial q^j}(x(t))\frac{\mathrm{d}q^j}{\mathrm{d}t}(t)v^i(t) + \sum_{i=1}^n g_{ik}(x(t))\frac{\mathrm{d}v^i}{\mathrm{d}t}(t) - \frac{1}{2}\sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial q^k}(x(t))v^i(t)v^j(t) \\
&\quad= \sum_{i,j=1}^n \frac{\partial g_{ik}}{\partial q^j}(x(t))v^j(t)v^i(t) + \sum_{i=1}^n g_{ik}(x(t))\frac{\mathrm{d}v^i}{\mathrm{d}t}(t) - \frac{1}{2}\sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial q^k}(x(t))v^i(t)v^j(t) \\
&\quad= \sum_{i=1}^n g_{ik}(x(t))\frac{\mathrm{d}v^i}{\mathrm{d}t}(t) + \frac{1}{2}\sum_{i,j=1}^n \left(\frac{\partial g_{ik}}{\partial q^j}(x(t)) + \frac{\partial g_{jk}}{\partial q^i}(x(t)) - \frac{\partial g_{ij}}{\partial q^k}(x(t))\right)v^i(t)v^j(t).
\end{aligned}$$

Thus, along geodesics we have that

$$\sum_{i=1}^n g_{ik}(x(t))\frac{\mathrm{d}v^i}{\mathrm{d}t}(t) + \frac{1}{2}\sum_{i,j=1}^n \left(\frac{\partial g_{ik}}{\partial q^j}(x(t)) + \frac{\partial g_{jk}}{\partial q^i}(x(t)) - \frac{\partial g_{ij}}{\partial q^k}(x(t))\right)v^i(t)v^j(t) = 0,$$

for all $k = 1, \ldots, n$. Raising the index k (or, in more details: multiplying everything by g^{kr}, summing over k, and then renaming $r \to k$), we obtain

$$\frac{\mathrm{d}v^k}{\mathrm{d}t}(t) + \sum_{i,j=1}^{k} \Gamma_{ij}^k(x(t))v^i(t)v^j(t) = 0, \qquad k = 1, \ldots, n,$$

as wanted. More commonly, we express this in terms of second derivatives of the $q^k(t)$ as

$$\frac{\mathrm{d}^2q^k}{\mathrm{d}t^2}(t) + \sum_{i,j=1}^{k} \Gamma_{ij}^k(x(t))\frac{\mathrm{d}q^i}{\mathrm{d}t}(t)\frac{\mathrm{d}q^j}{\mathrm{d}t}(t) = 0, \qquad k = 1, \ldots, n.$$

Note here that the Euler-Lagrange equations are not the classic geodesic differential equations: this process of raising the index is necessary (but in general this step is trivial, for example if one deals with a coordinate system for which the coordinate vector fields are orthogonal). The proof of Proposition 6 (p. 4) in fact gives us that

$$\mathrm{d}\mathscr{E}_x(\widetilde{x}) = \int_a^b g_{x(t)}\left(\frac{D\dot{x}}{\mathrm{d}t}(t), \frac{\partial \widetilde{x}}{\partial s}(0,t)\right) \mathrm{d}t,$$

where $D/\mathrm{d}t$ is the covariant derivative operator induced along x by the Levi-Civita connection of (Q, g).

Sub-examples:

- Let $\mathbb{H}^2 = \mathbb{R} \times \mathbb{R}_{>0}$ be the *hyperbolic plane* equipped with the Riemannian metric

$$g = \frac{\mathrm{d}x^2 + \mathrm{d}y^2}{y^2}.$$

We consider the Lagrangian

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}\frac{\dot{x}^2 + \dot{y}^2}{y^2},$$

where we (abusively) use $\dot{x}$ and $\dot{y}$ as coordinates in $T\mathbb{H}^2$, as $x = q^1$, $y = q^2$, $\dot{x} = v^1$ and $\dot{y} = v^2$. One equation is

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{x}}{y^2}\right) - 0 = \frac{\ddot{x}y^2 - 2y\dot{x}\dot{y}}{y^4},$$

and the other is

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\dot{y}}{y^2}\right) + \frac{\dot{x}^2 + \dot{y}^2}{y^3} = \frac{\ddot{y}y^2 - 2y\dot{y}^2}{y^4} + \frac{\dot{x}^2 + \dot{y}^2}{y^3}.$$

Simplified, they are nothing more than

$$\ddot{x} - \frac{2}{y}\dot{x}\dot{y} = 0 \quad \text{and} \quad \ddot{y} + \frac{1}{y}\dot{x}^2 - \frac{1}{y}\dot{y}^2 = 0.$$

So we conclude that

$$\Gamma^x_{xy} = -\frac{1}{y}, \quad \Gamma^y_{xx} = \frac{1}{y} \quad \text{and} \quad \Gamma^y_{yy} = -\frac{1}{y},$$

while all the remaining symbols vanish (pay attention to the factor 2 gone from the mixed symbol). Solving these equations, we have that the geodesics in $\mathbb{H}^2$ are vertical lines and half-circles centered in the x-axis (and hence meeting it orthogonally).

- In the half-plane $\mathbb{R}_{>0} \times \mathbb{R}$, consider now the *Grushin metric*

$$g = \mathrm{d}x^2 + \frac{\mathrm{d}y^2}{x^2}.$$

Using the same notations as in the $\mathbb{H}^2$ example, we consider the Lagrangian

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}\left(\dot{x}^2 + \frac{\dot{y}^2}{x^2}\right).$$

So we have

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{\mathrm{d}}{\mathrm{d}t}(\dot{x}) + \frac{\dot{y}^2}{x^3} = \ddot{x} + \frac{1}{x^3}\dot{y}^2, \\ 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\dot{y}}{x^2} - 0 = \frac{\ddot{y}x^2 - 2x\dot{x}\dot{y}}{x^4}, \end{aligned}$$

thus from

$$\ddot{x} + \frac{1}{x^3}\dot{y}^2 = 0 \quad \text{and} \quad \ddot{y} - \frac{2}{x}\dot{x}\dot{y} = 0$$

we can read $\Gamma^x_{yy} = 1/x^3$ and $\Gamma^y_{xy} = -1/x$, while all the remaining symbols vanish.

■

Example 9. Let's generalize the previous example. Let (Q, g) be a pseudo-Riemannian manifold, $V\colon \mathbb{R} \times Q \to \mathbb{R}$ a smooth function, and consider again the time-dependent Lagrangian with potential energy $L\colon \mathbb{R} \times TQ \to \mathbb{R}$ given by

$$L(t, x, \boldsymbol{v}) = \frac{1}{2}g_x(\boldsymbol{v}, \boldsymbol{v}) - V(t, x).$$

Describe a curve $x\colon [a, b] \to Q$ in coordinates by

$$(x(t), \dot{x}(t)) = (q^1(t), \dots, q^n(t), v^1(t), \dots, v^n(t)),$$

so that

$$L(t, x(t), \dot{x}(t)) = \frac{1}{2}\sum_{i,j=1}^n g_{ij}(x(t))v^i(t)v^j(t) - V(t, x(t)).$$

So we compute the Euler-Lagrange equations as:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(t,x(t),\dot{x}(t))\right)-\frac{\partial L}{\partial q^k}(t,x(t),\dot{x}(t))= \\
&=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial v^k}\frac{1}{2}\sum_{i,j=1}^n g_{ij}(x(t))v^i(t)v^j(t)\right)+\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial v^k}V(t,x(t))-\frac{\partial}{\partial q^k}\left(\frac{1}{2}\sum_{i,j=1}^n g_{ij}(x(t))v^i(t)v^j(t)-V(t,x(t))\right) \\
&=\sum_{i=1}^n g_{ik}(x(t))\frac{\mathrm{d}v^i}{\mathrm{d}t}(t)+\frac{1}{2}\sum_{i,j=1}^n\left(\frac{\partial g_{ik}}{\partial q^j}(x(t))+\frac{\partial g_{jk}}{\partial q^i}(x(t))-\frac{\partial g_{ij}}{\partial q^k}(x(t))\right)v^i(t)v^j(t)+\frac{\partial V}{\partial q^k}(t,x(t)).
\end{aligned}$$

So, along critical points of $\mathscr{A}_L$ we have

$$\sum_{i=1}^n g_{ik}(x(t))\frac{\mathrm{d}v^i}{\mathrm{d}t}(t)+\frac{1}{2}\sum_{i,j=1}^n\left(\frac{\partial g_{ik}}{\partial q^j}(x(t))+\frac{\partial g_{jk}}{\partial q^i}(x(t))-\frac{\partial g_{ij}}{\partial q^k}(x(t))\right)v^i(t)v^j(t)=-\frac{\partial V}{\partial q^k}(t,x(t)),$$

for all $k=1,\ldots,n$. Raising the index k, we obtain

$$\frac{\mathrm{d}^2q^k}{\mathrm{d}t^2}(t)+\sum_{i,j=1}^k\Gamma_{ij}^k(x(t))\frac{\mathrm{d}q^i}{\mathrm{d}t}(t)\frac{\mathrm{d}q^j}{\mathrm{d}t}(t)=-((\mathrm{grad}_g V(t,\cdot))_{x(t)})^k,$$

thus leading us to *Newton's equation*

$$\frac{D\dot{x}}{\mathrm{d}t}(t)=-(\mathrm{grad}_g V(t,\cdot))_{x(t)},$$

for all $t\in[a,b]$. Here, the gradient is computed only with respect to Q. In this setting, we define the *force* $\boldsymbol{F}\colon\mathbb{R}\times Q\to TQ$ by $\boldsymbol{F}(t,x)=-(\mathrm{grad}_g V(t,\cdot))_x$, so that

$$\boldsymbol{F}(t,x(t))=\frac{D\dot{x}}{\mathrm{d}t}(t)$$

resembles directly the classical $F=\mathrm{m}a$. The reason for the absence of the mass m of the particle whose path is modeled by the curve x in the formula above is that the usually such mass is already taken into account in the definition of the metric g, which is positive-definite when modelling a mechanical system. The mass is then an operator $\mathrm{m}: TQ\to T^*Q$, whose action is summarized in the "definition" of linear momentum $\mathrm{p}=\mathrm{m}\boldsymbol{v}$ (of course, this is nothing more than the identification between TQ and T^*Q given by g, tentatively phrased in physics language). So m is included in the definition of $\boldsymbol{F}$, formed as a gradient (i.e., if F is a gradient taken with respect to an inner product which does not take into account m, we have $\boldsymbol{F}=F/\mathrm{m}$). In more general settings, when dealing with non-conservative systems, we would consider time-dependent non-positional forces $\boldsymbol{F}\colon\mathbb{R}\times TQ\to TQ$, $\boldsymbol{F}=\boldsymbol{F}(t,x,\boldsymbol{v})$.

Example 10 (Fiberwise linear Lagrangians)**.** Let $\alpha\in\Omega^1(Q)$ and consider the Lagrangian $L\colon TQ\to\mathbb{R}$ given by $L_\alpha(x,\boldsymbol{v})=\alpha_x(\boldsymbol{v})$. Assume given a curve $x\colon[a,b]\to Q$. Take local coordinates $(q^1,\ldots,q^n,v^1,\ldots,v^n)$ and describe the curve by

$$(x(t),\dot{x}(t))=(q^1(t),\ldots,q^n(t),v^1(t),\ldots,v^n(t)).$$

Also write $\alpha = \sum_{j=1}^{n} \alpha_j \, \mathrm{d}q^j$ and $v = \sum_{j=1}^{n} v^j \frac{\partial}{\partial q^j}\Big|_x$, so that $L_\alpha(x, v) = \sum_{j=1}^{n} \alpha_j(x) v^j$. Thus we compute

$$
\begin{aligned}
\frac{\partial L_\alpha}{\partial q^k}(x(t), \dot{x}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_\alpha}{\partial v^k}(x(t), \dot{x}(t)) &= \sum_{j=1}^{n} \frac{\partial \alpha_j}{\partial q^k}(x(t)) v^j(t) - \frac{\mathrm{d}}{\mathrm{d}t}(\alpha_k(x(t))) \\
&= \sum_{j=1}^{n} \frac{\partial \alpha_j}{\partial q^k}(x(t)) v^j(t) - \sum_{j=1}^{n} \frac{\partial \alpha_k}{\partial q^j}(x(t)) \frac{\mathrm{d}q^j}{\mathrm{d}t}(t) \\
&= \sum_{j=1}^{n} \left(\frac{\partial \alpha_j}{\partial q^k}(x(t)) - \frac{\partial \alpha_k}{\partial q^j}(x(t)) \right) v^j(t),
\end{aligned}
$$

using that $\mathrm{d}q^j/\mathrm{d}t = v^j$ for all $j = 1, \ldots, n$. If the curve x is L-critical, then the last above expression vanishes for all $k = 1, \ldots, n$. Then sum over k and conclude that $\mathrm{d}\alpha_{x(t)}(\dot{x}(t), \cdot) = 0$. Conversely, if $\mathrm{d}\alpha_{x(t)}(\dot{x}(t), \cdot) = 0$, we may evaluate it at convenient tangent vectors and conclude that all the n Euler-Lagrange equations hold. So we conclude that $x : [a, b] \to Q$ is L-critical if and only if $\mathrm{d}\alpha_{x(t)}(\dot{x}(t), \cdot) = 0$ for all $t \in [a, b]$. Now, from the proof of Proposition 6 (p. 4), we'll also see that

$$
\mathrm{d}(\mathscr{A}_{L_\alpha})_x(\widetilde{x}) = \int_a^b \mathrm{d}\alpha_{x(t)} \left(\dot{x}(t), \frac{\partial \widetilde{x}}{\partial s}(0, t) \right) \mathrm{d}t.
$$

2 Fiber derivatives and Legendre transformations

Consider an autonomous Lagrangian $L\colon TQ \to \mathbb{R}$. For a point $x \in Q$, let also $L_x \doteq L\big|_{T_xQ}\colon T_xQ \to \mathbb{R}$ be the restriction of L to the fiber T_xQ. For $\boldsymbol{v} \in T_xQ$, we may compute the derivative $\mathrm{d}(L_x)_{\boldsymbol{v}}\colon T_{\boldsymbol{v}}T_xQ \to \mathbb{R}$, but $T_{\boldsymbol{v}}T_xQ \cong T_xQ$, and so we may regard such map as an element of T_x^*Q. Taking derivatives of L only along fiber directions leads us to the:

Definition 11. Let $L\colon TQ \to \mathbb{R}$ be an autonomous Lagrangian. The *fiber derivative* of L is the map $\mathbb{F}L\colon TQ \to T^*Q$ given by

$$\mathbb{F}L(x,\boldsymbol{v})\boldsymbol{w} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} L(x,\boldsymbol{v}+t\boldsymbol{w}).$$

We say that L is *regular* if $\mathbb{F}L$ is a local diffeomorphism, and that it is *hyperregular* if $\mathbb{F}L$ is a diffeomorphism.

Example 12.

(1) If (Q,g) is a pseudo-Riemannian manifold, $V\colon Q \to \mathbb{R}$ is smooth and the Lagrangian $L\colon TQ \to \mathbb{R}$ is given by

$$L(x,\boldsymbol{v}) = \frac{1}{2}g_x(\boldsymbol{v},\boldsymbol{v}) - V(x),$$

then

$$\begin{aligned}\mathbb{F}L(x,\boldsymbol{v})\boldsymbol{w} &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} L(x,\boldsymbol{v}+t\boldsymbol{w}) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(\frac{1}{2}g_x(\boldsymbol{v}+t\boldsymbol{w},\boldsymbol{v}+t\boldsymbol{w}) - V(x)\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(\frac{1}{2}\big(g_x(\boldsymbol{v},\boldsymbol{v}) + 2tg_x(\boldsymbol{v},\boldsymbol{w}) + t^2 g_x(\boldsymbol{w},\boldsymbol{w})\big) - V(x))\right) \\ &= g_x(\boldsymbol{v},\boldsymbol{w}).\end{aligned}$$

So $\mathbb{F}L(x,\boldsymbol{v}) = (x,\boldsymbol{v}_\flat)$, and $\mathbb{F}L = \flat$ is the usual musical isomorphism $TQ \to T^*Q$ induced by g. Thus L is hyperregular.

(2) Let $\alpha \in \Omega^1(Q)$ and consider again $L_\alpha\colon TQ \to \mathbb{R}$ given by $L_\alpha(x,\boldsymbol{v}) = \alpha_x(\boldsymbol{v})$. Then

$$\begin{aligned}\mathbb{F}L_\alpha(x,\boldsymbol{v})\boldsymbol{w} &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} L_\alpha(x,\boldsymbol{v}+t\boldsymbol{w}) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \alpha_x(\boldsymbol{v}+t\boldsymbol{w}) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (\alpha_x(\boldsymbol{v}) + t\alpha_x(\boldsymbol{w})) = \alpha_x(\boldsymbol{w}).\end{aligned}$$

So, the expression $\mathbb{F}L_\alpha(x,\boldsymbol{v}) = (x,\alpha_x)$ says that L_α is not regular, as it does not depend on $\boldsymbol{v}$ whatsoever (in particular, it is not injective when restricted any open subset of TQ).

We can do the same thing for T^*Q and autonomous Hamiltonians. Namely, if we're given $H\colon T^*Q \to \mathbb{R}$ and $x \in Q$, we may again restrict $H_x = H\big|_{T_x^*Q}\colon T_x^*Q \to \mathbb{R}$, and for $\mathsf{p} \in T_x^*Q$, compute the derivative $\mathrm{d}(H_x)_{\mathsf{p}}\colon T_{\mathsf{p}}T_x^*Q \to \mathbb{R}$. Like before, $T_{\mathsf{p}}T_x^*Q \cong T_x^*Q$ allows us to see this derivative as an element of $(T_x^*Q)^* \cong T_xQ$. So we have the:

Definition 13. Let $H\colon T^*Q \to \mathbb{R}$ be an autonomous Hamiltonian. The fiber derivative of H is the map $\mathbb{F}H\colon T^*Q \to TQ$ given by

$$\mathbb{F}H(x,\mathsf{p})\mathsf{q} = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} H(x,\mathsf{p}+t\mathsf{q}).$$

We'll say that H is regular or hyperregular according to whether $\mathbb{F}H$ is a local or global diffeomorphism.

Remark. By now you should have realized that these definitions may be written in a much more general setting: let $\pi\colon E \to Q$ be a fiber bundle, with typical fiber F. Put $E_x = \pi^{-1}(x)$. So each E_x is diffeomorphic to F and $E = \bigsqcup_{x\in Q} E_x$. Consider a smooth map $L\colon E \to \mathbb{R}$. For $x \in Q$, restrict $L\big|_{E_x}\colon E_x \to \mathbb{R}$, and for $\xi \in E_x$, compute the differential $\mathrm{d}(L\big|_{E_x})_\xi\colon T_\xi E_x \to \mathbb{R}$. This defines a map by $\mathbb{F}L(x,\xi) = \mathrm{d}(L\big|_{E_x})_\xi$. In the particular case where we have a vector bundle, then we may identify $T_\xi E_x \cong E_x$ and see $\mathbb{F}L(x,\xi)$ as an element of the dual space E_x^*. Thus we have a bundle morphism $\mathbb{F}L\colon E \to E^*$. If $E = TQ$, we have our first definition of fiber derivative. The above definition is just the case $E = T^*Q$ with the natural identification between a finite-dimensional vector space and its bidual, and we call our maps H instead of L.

Example 14 (Analogues)**.**

(1) If (Q,g) be a pseudo-Riemannian manifold, $V\colon Q \to \mathbb{R}$ is smooth and the Hamiltonian $H\colon T^*Q \to \mathbb{R}$ is given by

$$H(x,\mathsf{p}) = \frac{1}{2}g_x^*(\mathsf{p},\mathsf{p}) + V(x),$$

then

$$\begin{aligned}
\mathbb{F}H(x,\mathsf{p})\mathsf{q} &= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} H(x,\mathsf{p}+t\mathsf{q}) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(\frac{1}{2}g_x^*(\mathsf{p}+t\mathsf{q},\mathsf{p}+t\mathsf{q}) + V(x)\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(\frac{1}{2}\big(g_x^*(\mathsf{p},\mathsf{p}) + 2t g_x^*(\mathsf{p},\mathsf{q}) + t^2 g_x^*(\mathsf{q},\mathsf{q})\big) + V(x)\right) \\
&= g_x^*(\mathsf{p},\mathsf{q}) \\
&= g_x(\mathsf{p}^\sharp,\mathsf{q}^\sharp) \\
&= \mathsf{q}(\mathsf{p}^\sharp),
\end{aligned}$$

so that $\mathbb{F}H(x,\mathsf{p}) = (x,\mathsf{p}^\sharp)$ and $\mathbb{F}H = \sharp$ is the usual (inverse) musical isomorphism $T^*Q \to TQ$ induced by g.

(2) Let $\boldsymbol{X} \in \mathfrak{X}(Q)$ and consider $H^{\boldsymbol{X}}\colon T^*Q \to \mathbb{R}$ given by $H^{\boldsymbol{X}}(x, \mathsf{p}) = \mathsf{p}(\boldsymbol{X}_x)$. We have that

$$\begin{aligned} \mathbb{F}H^{\boldsymbol{X}}(x,\mathsf{p})\mathsf{q} &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} H^{\boldsymbol{X}}(x, \mathsf{p}+t\mathsf{q}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\mathsf{p}+t\mathsf{q})(\boldsymbol{X}_x) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (\mathsf{p}(\boldsymbol{X}_x) + t\mathsf{q}(\boldsymbol{X}_x)) = \mathsf{q}(\boldsymbol{X}_x). \end{aligned}$$

So $\mathbb{F}H^{\boldsymbol{X}}(x,\mathsf{p}) = (x, \boldsymbol{X}_x)$, and $H^{\boldsymbol{X}}$ is not regular.

Back to Lagrangians. Essentially, we want to see $\mathbb{F}L(x,\boldsymbol{v}) = (x,\mathsf{p})$ as a change of coordinates. The first step is to understand what is p in terms of x and $\boldsymbol{v}$, at least locally. So let $(q^1, \ldots, q^n, v^1, \ldots, v^n)$ be local coordinates in TQ, and compute

$$\begin{aligned} p_k(x,\boldsymbol{v}) &= \mathbb{F}L(x,\boldsymbol{v}) \left.\frac{\partial}{\partial q^k}\right|_x = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} L\left(x, \boldsymbol{v} + t\left.\frac{\partial}{\partial q^k}\right|_x\right) \\ &= \sum_{i=1}^n \frac{\partial L}{\partial v^i}(x,\boldsymbol{v}) \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (v^i + t\delta^i_k) \\ &= \sum_{i=1}^n \frac{\partial L}{\partial v^i}(x,\boldsymbol{v})\delta^i_k \\ &= \frac{\partial L}{\partial v^k}(x,\boldsymbol{v}). \end{aligned}$$

Usually, people write this suggestively as $\mathsf{p} = \partial L/\partial \boldsymbol{v}$. The quantity p_k is then called the *momentum* of L in the direction q^k, and the formula

$$\mathbb{F}L(x,\boldsymbol{v}) = \left(x, \sum_{k=1}^n \frac{\partial L}{\partial v^k}(x,\boldsymbol{v})\, \mathrm{d}q^k|_x\right)$$

holds. A quick criterion to see whether the above actually define local coordinates on T^*Q is the following:

Proposition 15. *Let $L\colon TQ \to \mathbb{R}$ be a Lagrangian. Then L is regular if and only if given any set of local coordinates $(q^1, \ldots, q^n, v^1, \ldots, v^n)$, we have the* Legendre condition

$$\det\left(\left(\frac{\partial^2 L}{\partial v^j \partial v^k}(x,\boldsymbol{v})\right)_{j,k=1}^n\right) \neq 0.$$

Remark. This does not require any proof: it is just the condition needed to apply the Inverse Function Theorem to the set of equations $\mathsf{p} = \partial L/\partial \boldsymbol{v}$. Such condition does not imply hyperregularity. For example, for $Q = \mathbb{R}$, define $L\colon T\mathbb{R} \cong \mathbb{R}^2 \to \mathbb{R}$ by $L(x,v) = \mathrm{e}^v$. Then $(\partial^2 L/\partial v^2)(x,v) = \mathrm{e}^v \neq 0$ (so L is regular), but

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \mathrm{e}^{v+tw} = \mathrm{e}^v w$$

says that the fiber derivative $\mathbb{F}L(x,v) = (x, \mathrm{e}^v)$ is not surjective (and hence L is not hyperregular). Here, of course, we're identifying $T^*\mathbb{R} \cong T\mathbb{R} \cong \mathbb{R}^2$ using the multiplication of $\mathbb{R}$, so e^v actually denotes the map $T_x\mathbb{R} \cong \mathbb{R} \ni w \mapsto \mathrm{e}^v w \in \mathbb{R}$.

Similar computations can be done for a Hamiltonian $H\colon T^*Q \to \mathbb{R}$: if we take local coordinates $(q^1,\ldots,q^n,p_1,\ldots,p_n)$ in T^*Q, we may compute

$$\begin{aligned} v^k(x,\mathsf{p}) &= \mathbb{F}H(x,\mathsf{p})\,\mathrm{d}q^k\big|_x = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} H(x,\mathsf{p}+t\,\mathrm{d}q^k\big|_x) \\ &= \sum_{i=1}^n \frac{\partial H}{\partial p_i}(x,p)\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(p_i+t\delta_i^k) \\ &= \sum_{i=1}^n \frac{\partial H}{\partial p_i}(x,\mathsf{p})\delta_i^k \\ &= \frac{\partial H}{\partial p_k}(x,\mathsf{p}), \end{aligned}$$

which gives

$$\mathbb{F}H(x,\mathsf{p}) = \left(x, \sum_{k=1}^n \frac{\partial H}{\partial p_k}(x,\mathsf{p})\frac{\partial}{\partial q^k}\bigg|_x\right).$$

Expressing v^k's in terms of p and p_k's in terms of v might suggest a deeper relation between those coordinates. For suitable choices of L and H, we indeed obtain something.

Definition 16. Let $L\colon TQ \to \mathbb{R}$ be an autonomous Lagrangian. Define the *energy of L* as the map $E_L\colon TQ \to \mathbb{R}$ given by

$$E_L(x,v) = \mathbb{F}L(x,v)v - L(x,v).$$

Similarly, one defines E_H for an autonomous Hamiltonian $H\colon T^*Q \to \mathbb{R}$.

Fiber derivatives and energy maps will provide the relation between Lagrangians and Hamiltonians:

Theorem 17. *Let $L\colon TQ \to \mathbb{R}$ be a hyperregular Lagrangian and define a Hamiltonian $H\colon T^*Q \to \mathbb{R}$ by $H = E_L \circ (\mathbb{F}L)^{-1}$. We say that H is the Legendre transformation of L. Then:*

(i) H is hyperregular and $\mathbb{F}H = (\mathbb{F}L)^{-1}$.

(ii) the associated Lagrangian $E_H \circ (\mathbb{F}H)^{-1}$ to H is the original L.

In other words, (i) tells us how to compute the inverse to $\mathbb{F}L$, and (ii) tells us that the Legendre transformation is an involution.

Proof: We'll do coordinate computations.

(i) First write $\mathbb{F}L(x,v) = (x,\mathsf{p})$, as before. We have that

$$H(x,\mathsf{p}) = \sum_{i=1}^n p_i(x,v)v^i - L(x,v),$$

and so

$$\frac{\partial H}{\partial p_k}(x,\mathsf{p}) = \frac{\partial}{\partial p_k}\bigg|_x \sum_{i=1}^n \Big(p_i(x,\boldsymbol{v})v^i - L(x,\boldsymbol{v})\Big) = v^k,$$

which says that

$$\mathbb{F}H \circ \mathbb{F}L(x,\boldsymbol{v}) = \left(x, \sum_{k=1}^n \frac{\partial H}{\partial p_k}(x,\mathsf{p})\frac{\partial}{\partial q^k}\bigg|_x\right) = \left(x, \sum_{k=1}^n v^k\frac{\partial}{\partial q^k}\bigg|_x\right) = (x,\boldsymbol{v}).$$

For the reverse composition, write $\mathbb{F}H(x,\mathsf{p}) = (x,\boldsymbol{v})$, so that

$$L(x,\boldsymbol{v}) = \sum_{i=1}^n p_i v^i(x,\mathsf{p}) - H(x,\mathsf{p}),$$

and thus

$$\frac{\partial L}{\partial v^k}(x,\boldsymbol{v}) = \frac{\partial}{\partial v^k}\bigg|_x \left(\sum_{i=1}^n p_i v^i(x,\mathsf{p}) - H(x,\mathsf{p})\right) = p_k,$$

leading to

$$\mathbb{F}L \circ \mathbb{F}H(x,\mathsf{p}) = \left(x, \sum_{k=1}^n \frac{\partial L}{\partial v^k}(x,\boldsymbol{v})\,\mathrm{d}q^k\big|_x\right) = \left(x, \sum_{k=1}^n p_k\,\mathrm{d}q^k\big|_x\right) = (x,\mathsf{p}),$$

as wanted.

(ii) With the relation between $\boldsymbol{v}$ and p in mind, we have that the associated Lagrangian to H is

$$\begin{aligned}(x,\boldsymbol{v}) \mapsto E_H(x,\mathsf{p}) &= \mathbb{F}H(x,\mathsf{p})\mathsf{p} - H(x,\mathsf{p})\\ &= \mathbb{F}H(x,\mathsf{p})\mathsf{p} - E_L(x,\boldsymbol{v})\\ &= \mathbb{F}H(x,\mathsf{p})\mathsf{p} - \mathbb{F}L(x,\boldsymbol{v})\boldsymbol{v} + L(x,\boldsymbol{v})\\ &= L(x,\boldsymbol{v}),\end{aligned}$$

since both fiber derivatives equal the sum $\sum_{k=1}^n p_k v^k$ and thus cancel each other.

$\square$

Now, let's see concrete examples:

Example 18.

(1) Let (Q,g) be a pseudo-Riemannian manifold, $V\colon Q \to \mathbb{R}$ be smooth, and our usual Lagrangian $L\colon TQ \to \mathbb{R}$ given by

$$L(x,\boldsymbol{v}) = \frac{1}{2}g_x(\boldsymbol{v},\boldsymbol{v}) - V(x).$$

We have seen in Example 12 (p. 12) that L is hyperregular and $\mathsf{p} \doteq \mathbb{F}L(x, \boldsymbol{v}) = \boldsymbol{v}_\flat$, and thus $\boldsymbol{v} = \mathsf{p}^\sharp$. So, we compute the associated Hamiltonian as:

$$\begin{aligned} H(x, \mathsf{p}) &= E_L(x, \mathsf{p}^\sharp) \\ &= \mathbb{F}L(x, \mathsf{p}^\sharp)\mathsf{p}^\sharp - L(x, \mathsf{p}^\sharp) \\ &= \mathsf{p}(\mathsf{p}^\sharp) - \left(\frac{1}{2} g_x(\mathsf{p}^\sharp, \mathsf{p}^\sharp) - V(x)\right) \\ &= g_x(\mathsf{p}^\sharp, \mathsf{p}^\sharp) - \frac{1}{2} g_x(\mathsf{p}^\sharp, \mathsf{p}^\sharp) + V(x) \\ &= \frac{1}{2} g_x^*(\mathsf{p}, \mathsf{p}) + V(x). \end{aligned}$$

This in particular justifies our choice of sign for V: minus in the Lagrangian and plus in the Hamiltonian. It is perhaps more natural to think of the total energy of a system as the sum of the kinetic energy and the potential energy. Of course, we know from classical mechanics that this total energy is constant along curves obeying Newton's equation $(D\dot{x}/\mathrm{d}t)(t) = \boldsymbol{F}(t, x(t))$, but this will be an easy consequence of what's to come. If one wants to carry out the inverse Legendre transform to get a further feeling about what is going on here, the procedure is the same: start with the Hamiltonian $H(x, \mathsf{p})$ above. In Example 14 (p. 13) we have seen that H is hyperregular and $\boldsymbol{v} = \mathbb{F}H(x, \mathsf{p}) = \mathsf{p}^\sharp$, so that $\mathsf{p} = \boldsymbol{v}_\flat$ and thus the associated Lagrangian is

$$\begin{aligned} L(x, \boldsymbol{v}) &= E_H(x, \boldsymbol{v}_\flat) \\ &= \mathbb{F}H(x, \boldsymbol{v}_\flat)\boldsymbol{v}_\flat - H(x, \boldsymbol{v}_\flat) \\ &= \boldsymbol{v}_\flat(\boldsymbol{v}) - \left(\frac{1}{2} g_x^*(\boldsymbol{v}_\flat, \boldsymbol{v}_\flat) + V(x)\right) \\ &= g_x(\boldsymbol{v}, \boldsymbol{v}) - \frac{1}{2} g_x(\boldsymbol{v}, \boldsymbol{v}) - V(x) \\ &= \frac{1}{2} g_x(\boldsymbol{v}, \boldsymbol{v}) - V(x), \end{aligned}$$

as predicted by Theorem 17.

(2) Let $\alpha \in \Omega^1(Q)$ and $\boldsymbol{X} \in \mathfrak{X}(Q)$, and consider L_α and $H^{\boldsymbol{X}}$, as before. We have seen that they're not even regular, but we could still try to compute their energies. We have

$$\begin{aligned} E_{L_\alpha}(x, \boldsymbol{v}) &= \mathbb{F}L_\alpha(x, \boldsymbol{v})\boldsymbol{v} - L_\alpha(x, \boldsymbol{v}) = \alpha_x(\boldsymbol{v}) - \alpha_x(\boldsymbol{v}) = 0, \\ E_{H^{\boldsymbol{X}}}(x, \mathsf{p}) &= \mathbb{F}H^{\boldsymbol{X}}(x, \mathsf{p})\mathsf{p} - H^{\boldsymbol{X}}(x, \mathsf{p}) = \mathsf{p}(\boldsymbol{X}_x) - \mathsf{p}(\boldsymbol{X}_x) = 0. \end{aligned}$$

Maybe that was to be expected.

Proceeding, it is natural to wonder what happens with the Euler-Lagrange equations after undergoing the Legendre transformation. The answer, which consists of the equivalence between the Lagrangian and Hamiltonian formalisms of classical mechanics, is in the:

Theorem 19. *Let $L\colon TQ \to \mathbb{R}$ and $H\colon T^*Q \to \mathbb{R}$ be associated via Legendre transform. Given a curve $x\colon [a,b] \to \mathbb{R}$, set $(x(t), \mathsf{p}(t)) = \mathbb{F}L(x(t), \dot{x}(t))$ and describe x in coordinates by*

$$(x(t), \dot{x}(t), \mathsf{p}(t)) = (q^1(t), \dots, q^n(t), v^1(t), \dots, v^n(t), p_1(t), \dots, p_n(t)).$$

Then the following are equivalent:

(i) The Euler-Lagrange equations

$$\frac{\partial L}{\partial q^k}(x(t), \dot{x}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(x(t), \dot{x}(t))\right) = 0$$

hold for $k = 1, \dots, n$.

(ii) The Hamilton equations

$$\frac{\mathrm{d}q^k}{\mathrm{d}t}(t) = \frac{\partial H}{\partial p_k}(x(t), \mathsf{p}(t)) \qquad \text{and} \qquad \frac{\mathrm{d}p_k}{\mathrm{d}t}(t) = -\frac{\partial H}{\partial q^k}(x(t), \mathsf{p}(t))$$

hold for all $k = 1, \dots, n$.

Proof: Assume that x satisfies the Euler-Lagrange equations. Then

$$H(x(t), \mathsf{p}(t)) = \sum_{i=1}^n p_i(t) v^i(t) - L(x(t), \dot{x}(t))$$

gives us that

$$\begin{aligned}\frac{\partial H}{\partial p_k}(x(t), \mathsf{p}(t)) &= \frac{\partial}{\partial p_k}\bigg|_{(x(t),\mathsf{p}(t))}\left(\sum_{i=1}^n p_i(t) v^i(t) - L(x(t), \dot{x}(t))\right) \\ &= \sum_{i=1}^n \delta_i^k v^i(t) - 0 = v^k(t) \\ &= \frac{\mathrm{d}q^k}{\mathrm{d}t}(t),\end{aligned}$$

and also

$$\begin{aligned}\frac{\partial H}{\partial q^k}(x(t), \mathsf{p}(t)) &= \frac{\partial}{\partial q^k}\bigg|_{(x(t),\mathsf{p}(t))}\left(\sum_{i=1}^n p_i(t) v^i(t) - L(x(t), \dot{x}(t))\right) \\ &= 0 - \frac{\partial L}{\partial q^k}(x(t), \dot{x}(t)) = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(x(t), \dot{x}(t))\right) \\ &= -\frac{\mathrm{d}p_k}{\mathrm{d}t}(t),\end{aligned}$$

where we first use the Euler-Lagrange-equations, and then the relation $p_k = \partial L/\partial v^k$ given by the Legendre transformation. Conversely, assume that x satisfies the Hamilton equations. Since

$$L(x(t), \dot{x}(t)) = \sum_{i=1}^n p_i(t) v^i(t) - H(x(t), \mathsf{p}(t)),$$

we have that

$$\frac{\partial L}{\partial q^k}(x(t), \dot{x}(t)) = \frac{\partial}{\partial q^k}\bigg|_{(x(t),\dot{x}(t))} \left(\sum_{i=1}^{n} p_i(t) v^i(t) - H(x(t), \mathsf{p}(t)) \right)$$
$$= 0 - \frac{\partial H}{\partial q^k}(x(t), \mathsf{p}(t)) = \frac{\mathrm{d}p_k}{\mathrm{d}t}(t),$$

and so

$$\frac{\partial L}{\partial q^k}(x(t), \dot{x}(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(x(t), \dot{x}(t)) \right) = \frac{\mathrm{d}p_k}{\mathrm{d}t}(t) - \frac{\mathrm{d}p_k}{\mathrm{d}t}(t) = 0,$$

as wanted. □

Perhaps the only thing left to observe here is that while the Euler-Lagrange equations arose directly from a variational principle (i.e., characterizing critical points of the action functional $\mathscr{A}_L$), it is not clear whether there is a variational principle generating Hamilton's equations directly, avoiding a transference process like Legendre transformations. It turns out that the answer to this is affirmative. Here's the:

Theorem 20. *For a given Hamiltonian $H\colon T^*Q \to \mathbb{R}$ and a curve $(x, \mathsf{p})\colon [a,b] \to T^*Q$, define the* action integral of H *as*

$$\mathscr{A}^H[x, \mathsf{p}] = \int_a^b \mathsf{p}(t)\big(\dot{x}(t)\big) - H(x(t), \mathsf{p}(t))\,\mathrm{d}t.$$

If we describe a critical point of $\mathscr{A}^H$ as

$$(x(t), \mathsf{p}(t)) = (q^1(t), \ldots, q^n(t), p_1(t), \ldots, p_n(t)),$$

then (x, p) satisfies Hamilton's equations

$$\frac{\mathrm{d}q^k}{\mathrm{d}t}(t) = \frac{\partial H}{\partial p_k}(x(t), \mathsf{p}(t)) \quad \textit{and} \quad \frac{\mathrm{d}p_k}{\mathrm{d}t}(t) = -\frac{\partial H}{\partial q^k}(x(t), \mathsf{p}(t)),$$

for $k = 1, \ldots, n$.

Proof: Let $(\widetilde{x}, \widetilde{\mathsf{p}})\colon (-\epsilon, \epsilon) \times [a,b] \to T^*Q$ a variation of (x, p) with fixed endpoints, that is, satisfying the conditions

- $(\widetilde{x}(0,t), \widetilde{\mathsf{p}}(0,t)) = (x(t), \mathsf{p}(t))$ for all $t \in [a,b]$, and;
- $\widetilde{x}(s,a) = x(a)$ and $\widetilde{x}(s,b) = x(b)$ for all $s \in (-\epsilon, \epsilon)$.

Let's use $\langle \cdot, \cdot \rangle$ for the standard inner product in $\mathbb{R}^{2n}$ and for the natural pairing between vectors and covectors. Identifying $(\mathbb{R}^n)^* \cong \mathbb{R}^n$ and write all vectors as

$$\frac{\partial H}{\partial q} = \left(\frac{\partial H}{\partial q^1}, \ldots, \frac{\partial H}{\partial q^n} \right), \qquad \frac{\partial H}{\partial p}, \quad \frac{\partial \widetilde{\mathsf{p}}}{\partial s},$$

and so on. So

$$\mathscr{A}^H[x, \mathsf{p}] = \int_a^b \langle \mathsf{p}(t), \dot{x}(t) \rangle - H(x(t), \mathsf{p}(t))\,\mathrm{d}t.$$

With this, let's compute:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathscr{A}^H[\widetilde{x}(s,\cdot),\widetilde{\mathsf{p}}(s,\cdot)] &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \int_a^b \left\langle \widetilde{\mathsf{p}}(s,t), \frac{\partial \widetilde{x}}{\partial t}(s,t)\right\rangle - H(\widetilde{x}(s,t),\widetilde{\mathsf{p}}(s,t))\,\mathrm{d}t \\
&= \int_a^b \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \left(\left\langle \widetilde{\mathsf{p}}(s,t), \frac{\partial \widetilde{x}}{\partial t}(s,t)\right\rangle - H(\widetilde{x}(s,t),\widetilde{\mathsf{p}}(s,t))\right)\mathrm{d}t \\
&= \int_a^b \left\langle \frac{\partial \widetilde{\mathsf{p}}}{\partial s}(0,t), \frac{\mathrm{d}x}{\mathrm{d}t}(t)\right\rangle + \left\langle \mathsf{p}(t), \frac{\partial^2 \widetilde{x}}{\partial s \partial t}(0,t)\right\rangle \\
&\quad - \left\langle \frac{\partial H}{\partial q}(x(t),\mathsf{p}(t)), \frac{\partial \widetilde{x}}{\partial s}(0,t)\right\rangle - \left\langle \frac{\partial H}{\partial p}(x(t),\mathsf{p}(t)), \frac{\partial \widetilde{\mathsf{p}}}{\partial s}(0,t)\right\rangle \mathrm{d}t \\
&= \int_a^b \left\langle \frac{\partial \widetilde{\mathsf{p}}}{\partial s}(s,0), \frac{\mathrm{d}x}{\mathrm{d}t}(t) - \frac{\partial H}{\partial p}(x(t),\mathsf{p}(t))\right\rangle - \left\langle \frac{\partial \widetilde{x}}{\partial s}(0,t), \frac{\mathrm{d}\mathsf{p}}{\mathrm{d}t}(t) + \frac{\partial H}{\partial q}(x(t),\mathsf{p}(t))\right\rangle \mathrm{d}t \\
&\quad + \left\langle \mathsf{p}(t), \frac{\partial \widetilde{x}}{\partial s}(0,t)\right\rangle\bigg|_a^b
\end{aligned}$$

If (x,p) is critical for $\mathscr{A}^H$, since the variation has fixed points we obtain that

$$\int_a^b \left\langle \frac{\partial \widetilde{\mathsf{p}}}{\partial s}(0,t), \frac{\mathrm{d}x}{\mathrm{d}t}(t) - \frac{\partial H}{\partial p}(x(t),\mathsf{p}(t))\right\rangle - \left\langle \frac{\partial \widetilde{x}}{\partial s}(0,t), \frac{\mathrm{d}\mathsf{p}}{\mathrm{d}t}(t) + \frac{\partial H}{\partial q}(x(t),\mathsf{p}(t))\right\rangle \mathrm{d}t = 0.$$

But variations can be chosen suitably, with $\partial\widetilde{\mathsf{p}}/\partial s = 0$ and $\partial\widetilde{x}/\partial s$ still arbitrary, so that

$$\int_a^b \left\langle \frac{\partial \widetilde{x}}{\partial s}(0,t), \frac{\mathrm{d}\mathsf{p}}{\mathrm{d}t}(t) + \frac{\partial H}{\partial q}(x(t),\mathsf{p}(t))\right\rangle \mathrm{d}t = 0$$

for arbitrary variational fields $\partial\widetilde{x}/\partial s$ implies

$$\frac{\mathrm{d}p_k}{\mathrm{d}t}(t) = -\frac{\partial H}{\partial q^k}(x(t),\mathsf{p}(t)), \qquad k = 1,\ldots,n.$$

Back to the previous integral, now we can say that

$$\int_a^b \left\langle \frac{\partial \widetilde{\mathsf{p}}}{\partial s}(0,t), \frac{\mathrm{d}x}{\mathrm{d}t}(t) - \frac{\partial H}{\partial p}(x(t),\mathsf{p}(t))\right\rangle \mathrm{d}t = 0$$

for arbitrary $\partial\widetilde{\mathsf{p}}/\partial s$ implies that

$$\frac{\mathrm{d}q^k}{\mathrm{d}t}(t) = \frac{\partial H}{\partial p_k}(x(t),\mathsf{p}(t)), \qquad k = 1,\ldots,n,$$

as wanted. □

Remark. There's something deeper going on with this action integral, regarding the particular structure of T^*Q. Let $\pi\colon T^*Q \to Q$ be the usual projection, $\pi(x,\mathsf{p}) = x$. There is a natural way to define a 1-form in T^*Q. Namely, given $(x,\mathsf{p}) \in T^*Q$ and $\boldsymbol{X}_{(x,\mathsf{p})} \in T_{(x,\mathsf{p})}T^*Q$, we need to produce a real number. Projecting $\boldsymbol{X}_{(x,\mathsf{p})}$ we obtain a vector $\mathrm{d}\pi_{(x,\mathsf{p})}(\boldsymbol{X}_{(x,\mathsf{p})}) \in T_xQ$, on which p can act on. So we define the so called *tautological* 1*-form* $\lambda \in \Omega^1(T^*Q)$ by the formula

$$\lambda_{(x,\mathsf{p})}(\boldsymbol{X}_{(x,\mathsf{p})}) = \mathsf{p}\left(\mathrm{d}\pi_{(x,\mathsf{p})}(\boldsymbol{X}_{(x,\mathsf{p})})\right).$$

A simple justification for the name "tautological" is the fact that any 1-form $\alpha \in \Omega^1(Q)$ is actually a map $\alpha\colon Q \to T^*Q$, and the pull-back of λ is just $\alpha^*\lambda = \alpha$. In fact, this property *characterizes* λ, which is given in coordinates by $\lambda = \sum_{k=1}^n p_k\,\mathrm{d}q^k$, since

$$\lambda_{(x,\mathsf{p})}\left(\left.\frac{\partial}{\partial q^k}\right|_{(x,\mathsf{p})}\right) = \mathsf{p}\left(\mathrm{d}\pi_{(x,\mathsf{p})}\left(\left.\frac{\partial}{\partial q^k}\right|_{(x,\mathsf{p})}\right)\right) = \mathsf{p}\left(\left.\frac{\partial}{\partial q^k}\right|_{x}\right) = p_k(x,\mathsf{p})$$

and $\partial/\partial p_k$ is always vertical. All that said, note that

$$\lambda_{(x(t),\mathsf{p}(t))}\left(\frac{\mathrm{d}}{\mathrm{d}t}(x(t),\mathsf{p}(t))\right) = \mathsf{p}(t)\big(\dot{x}(t)\big).$$

The integrand of the action integral $\mathscr{A}^H$ is the so-called action form $\lambda_H = \lambda - H\,\mathrm{d}t$. One can do the calculations from the previous proof using λ directly and Cartan's magic formula.

Let's conclude this section noting one last consequence of the Euler-Lagrange equations, regarding the total energy E_L:

Proposition 21. *If a curve $x\colon [a,b] \to Q$ is L-critical, then E_L is constant along x.*

Proof: Compute in coordinates:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}E_L(x(t),\dot{x}(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbb{F}L(x(t),\dot{x}(t))\dot{x}(t) - L(x(t),\dot{x}(t))\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=1}^n \frac{\partial L}{\partial v^k}(x(t),\dot{x}(t))v^k(t)\right) - \sum_{k=1}^n \frac{\partial L}{\partial q^k}(x(t),\dot{x}(t))v^k(t) - \sum_{k=1}^n \frac{\partial L}{\partial v^k}(x(t),\dot{x}(t))\frac{\mathrm{d}v^k}{\mathrm{d}t}(t) \\
&= \sum_{k=1}^n \left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(x(t),\dot{x}(t))\right) - \frac{\partial L}{\partial q^k}(x(t),\dot{x}(t))\right)v^k(t) \\
&= 0,
\end{aligned}$$

since each one of the n Euler-Lagrange equations kills a summand. $\square$

Remark.

- This has a somewhat obvious adaptation for time-dependent Lagrangians, with the necessary adjustments done in the definition of fiber derivative.
- This result also goes by the slogan "the Hamiltonian is constant along trajectories of the action functional associated to the Lagrangian". And this is now perfectly justified, as the Hamiltonian is E_L up to a change of coordinates.

Example 22. Item (1) of Example 18 above says that in a conservative mechanical system, the sum of the kinetic energy with the potential energy is constant along paths satisfying Newton's equation

$$\frac{D\dot{x}}{\mathrm{d}t}(t) = \boldsymbol{F}(t, x(t)).$$

In particular, when there is no potential, we conclude that in any pseudo-Riemannian manifold, geodesics have constant speed (and so their causal character is determined by a single tangent vector to it).

Looking for *conservation laws* of mechanical systems, such as the one given Proposition 21 is our goal in the next and last session. In particular, we will see Noether's theorem, which is a rich source of conservation laws, exploiting symmetries of L.

3 Invariance and Noether's Theorem

Let's start by formalizing what we mean by a symmetry of a Lagrangian:

Definition 23. Let $L\colon TQ \to \mathbb{R}$ be a Lagrangian.

(i) A diffeomorphism $\varphi\colon Q \to Q$ is a *symmetry* of L (or, L is *φ-invariant*) if

$$L(\varphi(x), \mathrm{d}\varphi_x(\boldsymbol{v})) = L(x, \boldsymbol{v}),$$

for all $(x, \boldsymbol{v}) \in TQ$.

(ii) A vector field $\boldsymbol{X} \in \mathfrak{X}(Q)$ is an *infinitesimal symmetry* of L if for every time t for which the flow $\Phi_{t,\boldsymbol{X}}$ is defined, L is $\Phi_{t,\boldsymbol{X}}$-invariant (in an adequate domain).

Let's see a few examples:

Example 24.

(1) Let (Q, g) be a pseudo-Riemannian manifold, $V\colon Q \to \mathbb{R}$ be smooth, and consider the natural Lagrangian $L\colon TQ \to \mathbb{R}$ given by

$$L(x, \boldsymbol{v}) = \frac{1}{2} g_x(\boldsymbol{v}, \boldsymbol{v}) - V(x).$$

Then any $\varphi \in \mathrm{Iso}(Q, g)$ satisfying the additional property $V(\varphi(x)) = V(x)$ for all $x \in Q$ leaves L invariant. This happens, for example, when $Q = \mathbb{R}^n$ and $V(x) = \widetilde{V}(\|x\|^2)$ depends only on the distance between the point x and the origin.

Sub-examples:

- Consider a mass $\mathrm{m} > 0$, g the acceleration of gravity, $Q = \mathbb{R}^3$ with coordinates (x, y, z), equipped with the Riemannian metric $\mathrm{m}\langle\cdot,\cdot\rangle$ (here $\langle\cdot,\cdot\rangle$ is the standard metric in $\mathbb{R}^3$), and let $V\colon \mathbb{R}^3 \to \mathbb{R}$ be the gravitational potential $V(x, y, z) = \mathrm{m}gz$. Thus

$$L(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{\mathrm{m}}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^3) - \mathrm{m}gz.$$

Note that with respect to the metric $m\langle\cdot,\cdot\rangle$, the conservative force is given by $\boldsymbol{F}(x, y, z) = -g\partial_z$. So for a motion $\gamma\colon I \to \mathbb{R}^3$ given by $\gamma(t) = (x(t), y(t), z(t))$, Newton's equation $\ddot{\gamma}(t) = \boldsymbol{F}(\gamma(t))$ (the Levi-Civita connection of $m\langle\cdot,\cdot\rangle$ is the same as $\langle\cdot,\cdot\rangle$'s) give us that

$$\begin{cases} \ddot{x}(t) = 0 \\ \ddot{y}(t) = 0 \\ \ddot{z}(t) = -g \end{cases} \implies \begin{cases} x(t) = a_x + v_x t \\ y(t) = a_y + v_y t \\ z(t) = a_z + v_z t - \dfrac{gt^2}{2}, \end{cases}$$

where $\boldsymbol{a}_\gamma = (a_x, a_y, a_z)$ is the starting point and $\boldsymbol{v}_\gamma = (v_x, v_y, v_z)$ is the initial velocity of the motion. Now, for every $\theta \in \mathbb{R}$, we have that

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

leaves L invariant. Since this is true for all θ, one might expect that taking the θ-derivative of something would give us something of physical significance. Indeed, we know that the angular momentum $-\mathrm{m}\dot{x}y + \mathrm{m}x\dot{y}$ is constant along the motion γ, as

$$\frac{\mathrm{d}}{\mathrm{d}t}(-\mathrm{m}v_x(a_y + v_y t) + \mathrm{m}(a_x + v_x t)a_y) = -\mathrm{m}v_x v_y + \mathrm{m}v_x v_y = 0,$$

but this can be also expressed in terms of R_θ, as follows: since

$$\left.\frac{\partial}{\partial\theta}\right|_{\theta=0} R_\theta(x,y,z) = -y\partial_x\big|_{(x,y,z)} + x\partial_y\big|_{(x,y,z)},$$

we have

$$\mathbb{F}L(x,y,z,\dot{x},\dot{y},\dot{z})\left(\left.\frac{\mathrm{d}}{\mathrm{d}\theta}\right|_{\theta=0} R_\theta(x,y,z)\right) = -\mathrm{m}\dot{x}y + \mathrm{m}x\dot{y},$$

by item (1) of Example 12. The above quantity, in general, is called a *Noether charge*. So we see that the Noether charge is constant along motions of the mechanical system $(\mathbb{R}^3, \mathrm{m}\langle\cdot,\cdot\rangle, -\mathrm{d}V)$. Observe also that $-y\partial_x + x\partial_y$ is a infinitesimal symmetry of L (and a Killing vector field for $(\mathbb{R}^3, \mathrm{m}\langle\cdot,\cdot\rangle)$. This is a particular instance of something more general that we'll see in details soon. Such idea can also be used to actually get information about the motions of the system themselves. For example, for each $s \in \mathbb{R}$, consider the translation $\tau_s\colon \mathbb{R}^3 \to \mathbb{R}^3$ given by $\tau_s(x,y,z) = (x+s,y,z)$. Since the total derivative of τ_s at any point is the identity, we have that L is τ_s-invariant for all $s \in \mathbb{R}$. Setting up a second Noether charge

$$\mathbb{F}L(x,y,z,\dot{x},\dot{y},\dot{z})\left(\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \tau_s(x,y,z)\right) = \mathrm{m}\dot{x},$$

we reconfirm that $\dot{x}(t)$ is constant for a motion, and so $x(t) = a_x + v_x t$. Same thing for $y(t)$.

- Let $V = 0$ and $\boldsymbol{\xi} \in \mathfrak{X}(Q)$ be a Killing vector field along Q. Then, where defined, the flow $\Phi_{s,\boldsymbol{\xi}}$ consists of isometries. Let $x\colon [a,b] \to Q$ is a geodesic of Q whose image is contained in the domain of the flow of $\boldsymbol{\xi}$, and compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} g_{x(t)}(\dot{x}(t), \boldsymbol{\xi}_{x(t)}) &= \dot{x}(t)\big(g(\dot{x},\boldsymbol{\xi})\big) \\ &= g_{x(t)}\left(\frac{D\dot{x}}{\mathrm{d}t}(t), \boldsymbol{\xi}_{x(t)}\right) + g_{x(t)}(\dot{x}(t), \nabla_{\dot{x}(t)}\boldsymbol{\xi}) = 0,\end{aligned}$$

since $D\dot{x}/\mathrm{d}t = \mathbf{0}$ and $\nabla\boldsymbol{\xi}$ is skew-adjoint. So, we conclude that the function $[a,b] \ni t \mapsto g_{x(t)}(\dot{x}(t), \boldsymbol{\xi}_{x(t)}) \in \mathbb{R}$ is constant. But this can be expressed in terms of a Noether charge, since

$$\mathbb{F}L(x,v)\left(\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi_{t,\boldsymbol{\xi}}(x)\right) = \mathbb{F}L(x,v)\boldsymbol{\xi}_x = g_x(v, \boldsymbol{\xi}_x).$$

In other words, the Noether charge is constant along geodesics (L-critical curves, when $V = 0$).

- As a consequence of the previous point: let $Q \subseteq \mathbb{R}^3$ be a surface covered by a *Clairaut chart*, i.e., coordinates (u, v) for which the metric of $\mathbb{R}^3$ induced on Q takes the form $\mathrm{d}s^2 = E(u)\,\mathrm{d}u^2 + G(u)\,\mathrm{d}v^2$, where E and G are certain smooth functions. We easily see that the coordinate vector field ∂_v is a Killing vector field along Q. So, if x is a geodesic in Q described in coordinates by $x(t) = (u(t), v(t))$, we have that

$$\left\langle \dot{x}(t), \partial_v\big|_{x(t)} \right\rangle_{x(t)} = G(u(t))\dot{v}(t) = c,$$

for some real constant c, for all time t. In particular, if x has unit speed and $\theta(t)$ is the angle between $\dot{x}(t)$ and $\partial_u\big|_{x(t)}$, we use that

$$\left\langle \dot{x}(t), \partial_v\big|_{x(t)} \right\rangle = \|\dot{x}(t)\| \|\partial_v\big|_{x(t)}\| \cos\left(\frac{\pi}{2} - \theta(t)\right) = \sqrt{G(u(t))}\sin\theta(t)$$

and obtain the famous *Clairaut relation* (renaming c):

$$\sqrt{G(u(t))}\sin\theta(t) = c.$$

As a consequence, we see that for Clairaut surfaces Q, unit speed geodesics are forced to remain in the region of Q where $G \geq c^2$.

(2) Let $\alpha \in \Omega^1(Q)$ and consider the fiberwise linear Lagrangian $L_\alpha\colon TQ \to \mathbb{R}$, and consider $\boldsymbol{X} \in \mathfrak{X}(Q)$ be an infinitesimal symmetry of L_α. This means that $\Phi_{t,\boldsymbol{X}}^*\alpha = \alpha$, for all times t for which this makes sense. Then we claim that along L_α-critical curves x, the function $t \mapsto \alpha_{x(t)}(\boldsymbol{X}_{x(t)})$ is constant. Indeed, we have by Cartan's magic formula that $\mathscr{L}_{\boldsymbol{X}}\alpha = \mathrm{d}(\iota_{\boldsymbol{X}}\alpha) + \iota_{\boldsymbol{X}}\mathrm{d}\alpha$, but $\iota_{\boldsymbol{X}}\mathrm{d}\alpha = -\mathrm{d}\alpha(\cdot, \boldsymbol{X})$, $\iota_{\boldsymbol{X}}\alpha = \alpha(\boldsymbol{X})$ and $\mathscr{L}_{\boldsymbol{X}}\alpha = 0$ by assumption, so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha_{x(t)}(\boldsymbol{X}_{x(t)}) = \mathrm{d}\alpha_{x(t)}(\dot{x}(t), \boldsymbol{X}_{x(t)}) = 0$$

for L_α-critical x, by the chain rule and Example 10. But we can again express this in terms of a Noether charge:

$$\mathbb{F}L_\alpha(x, \boldsymbol{v})\left(\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\Phi_{s,\boldsymbol{X}}(x)\right) = \mathbb{F}L_\alpha(x, \boldsymbol{v})\boldsymbol{X}_x = \alpha_x(\boldsymbol{X}_x).$$

So again, the Noether charge is constant along L_α-critical curves.

(3) Assume that we $Q = \mathbb{R}^n$ and we have global coordinates $\boldsymbol{q} = (q^1, \ldots, q^n)$. Consider a smooth Lagrangian $L\colon T\mathbb{R}^n \cong \mathbb{R}^{2n} \to \mathbb{R}$. Assume that for a fixed index k, we have that $\partial L/\partial q^k = 0$. So, the Euler-Lagrange equations tell us that along L-critical curves x, the momentum

$$\frac{\partial L}{\partial v^k}(x(t), \dot{x}(t))$$

is a constant function of t. Albeit silly, we can express this in terms of a Noether charge: we have that for each $s \in \mathbb{R}$, the translation map $\varphi_s \colon \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\varphi_s(\boldsymbol{q}) = (q^1, \dots, q^{k-1}, q^k + s, q^{k+1}, \dots, q^k)$$

leaves L invariant. And we have that

$$\mathbb{F}L(\boldsymbol{q}, \boldsymbol{v}) \left(\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \varphi_s(\boldsymbol{q}) \right) = \mathbb{F}L(\boldsymbol{q}, \boldsymbol{v}) \left(\frac{\partial}{\partial q^k}\bigg|_{\boldsymbol{q}} \right) = \frac{\partial L}{\partial v^k}(\boldsymbol{q}, \boldsymbol{v}).$$

Such a coordinate q^k is called *cyclic*. Just in the same way that $p_k = \partial L/\partial v^k$ is called a momentum, $\partial L/\partial q^k$ is called the *force* of L in the direction of q^k. With this terminology and bearing in mind the general idea that the force is the derivative of the momentum (this is the content of the Euler-Lagrange equations), what we have done above just says that if there is no force in a given direction, the corresponding momentum is constant. The conservation of energy given in Proposition 21 can also be expressed as a Noether charge.

Now, we have expressed conservations laws that we already had proofs for, in terms of Noether charges. But what if we wanted to discover new conservation laws? So, to begin with, what do all the R_θ, flows of vector fields, ad families of translations have in common? All of them are actually 1-*parameter groups* of diffeomorphisms, that is, satisfying the properties $\varphi_0 = \mathrm{Id}_Q$ and $\varphi_{s+s'} = \varphi_s \circ \varphi_{s'}$. This is the key for motivating the correct statement of the:

Theorem 25 (Noether). *Let $L \colon TQ \to \mathbb{R}$ be a Lagrangian, and $(\varphi_s)_{s\in\mathbb{R}}$ a 1-parameter group of diffeomorphisms of Q leaving L invariant. Then the* Noether charge *associated to the $(\varphi_s)_{s\in\mathbb{R}}$, defined by*

$$\mathcal{J}(x, \boldsymbol{v}) \doteq \mathbb{F}L(x, \boldsymbol{v}) \left(\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \varphi_s(x) \right),$$

is constant along L-critical curves in Q.

Proof: Let x be a L-critical curve and, for each s, put $x_s = \varphi_s \circ x$. In particular, the chain rule gives $\dot{x}_s(t) = \mathrm{d}(\varphi_s)_{x(t)}(\dot{x}(t))$. We'll do a coordinate computation, writing

$$(x_s(t), \dot{x}_s(t)) = (q^1(s,t), \dots, q^n(s,t), v^1(s,t), \dots, v^n(s,t)).$$

Since $L(x_s(t), \dot{x}_s(t)) = L(x(t), \dot{x}(t))$ for every $s \in \mathbb{R}$, we may apply $\partial/\partial s$ to get

$$\begin{aligned} 0 &= \sum_{k=1}^n \frac{\partial L}{\partial q^k}(x_s(t), \dot{x}_s(t)) \frac{\partial q^k}{\partial s}(s,t) + \sum_{k=1}^n \frac{\partial L}{\partial v^k}(x_s(t), \dot{x}_s(t)) \frac{\partial v^k}{\partial s}(s,t) \\ &\overset{(1)}{=} \sum_{k=1}^n \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(x_s(t), \dot{x}_s(t)) \right) \frac{\partial q^k}{\partial s}(s,t) + \sum_{k=1}^n \frac{\partial L}{\partial v^k}(x_s(t), \dot{x}_s(t)) \frac{\partial^2 q^k}{\partial s \partial t}(s,t) \\ &\overset{(2)}{=} \sum_{k=1}^n \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^k}(x_s(t), \dot{x}_s(t)) \frac{\partial q^k}{\partial s}(s,t) \right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=1}^n \frac{\partial L}{\partial v^k}(x_s(t), \dot{x}_s(t)) \frac{\partial q^k}{\partial s}(s,t) \right), \end{aligned}$$

where in (1) we use the Euler-Lagrange equations for x_s and in (2) we use that $\partial/\partial s$ and $\partial/\partial t$ commute, and the product rule. Now set $s = 0$ (i.e., $x_0 = x$) and recognize the equality

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbb{F}L(x(t),\dot{x}(t))\left(\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\varphi_s(x(t))\right)\right) = \frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{J}(x(t),\dot{x}(t))),$$

as wanted. □

There is, however, a more powerful generalization of Noether's theorem. Thinking of "symmetries" should be the same as thinking of groups (this is why people should care about groups, to begin with). We can see the action of $(\varphi_s)_{s\in\mathbb{R}}$ in Q as a group action $\mathbb{R} \circlearrowright Q$. So, with this idea in mind, fix a Lie group G with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$, and let $G \circlearrowright Q$ be a smooth action in the configuration space. That is, we have a map $G \times Q \to Q$ satisfying the axioms

(a) $e_G \cdot x = x$ for all $x \in Q$, where e_G is the identity[2] of G,

(b) $g \cdot (h \cdot x) = (gh) \cdot x$ for all $x \in Q$ and $g, h \in G$.

Fixed $g \in G$, the map $A_g \colon Q \to Q$ given by $A_g(x) = g \cdot x$ is called the *action map* of g, and it is usually identified with g itself. And fixed $x \in Q$, the map $\mathcal{O}_x \colon G \to Q$ given by $\mathcal{O}_x(g) = g \cdot x$ is called the *orbit map* of x. Note the relation $A_g(x) = \mathcal{O}_x(g)$.

Now, we also obtain an action $G \circlearrowright TQ$ given via derivatives. Namely, if we regard an element $g \in G$ as a diffeomorphism $g \colon Q \to Q$, then we set

$$g \cdot (x, \boldsymbol{v}) = (gx, \mathrm{d}g_x(\boldsymbol{v})).$$

This is indeed an action, since the derivative of $e_G = \mathrm{Id}_Q$ is again the identity and we apply the chain rule with the action property $g(hx) = (gh)x$ to get

$$\begin{aligned} g \cdot (h \cdot (x, \boldsymbol{v})) &= g \cdot (hx, \mathrm{d}h_x(\boldsymbol{v})) \\ &= (g(h(x)), \mathrm{d}g_{hx}\mathrm{d}h_x(\boldsymbol{v})) \\ &= ((gh)x, \mathrm{d}(gh)_x(\boldsymbol{v})) \\ &= (gh) \cdot (x, \boldsymbol{v}), \end{aligned}$$

as wanted. With this, the following definition makes sense:

Definition 26. A Lagrangian $L \colon TQ \to \mathbb{R}$ is G-invariant if $L(g \cdot (x, \boldsymbol{v})) = L(x, \boldsymbol{v})$, for every $g \in G$.

Remark. In the sense of Definition 23, note that this is the same as saying that L is g-invariant for every $g \in G$, in view of the definition $G \circlearrowright TQ$ and identifying g with its action map $TQ \to TQ$.

Before proceeding, let's make one more comparison with the case $G = \mathbb{R}$: the derivative $x \mapsto \partial_s\big|_{s=0}\varphi_s(x)$ actually defined a tangent vector in T_xQ, which we can see as coming from the element 1 of the Lie algebra $\mathrm{Lie}(\mathbb{R})$. So, elements of $\mathfrak{g}$ will generate vector fields in Q via the action. More precisely, we have the:

[2] The letter e comes from German, *einselement* (unit element).

Definition 27. The *infinitesimal action* (or *action field*) $X^\# \in \mathfrak{X}(Q)$ of an element $X \in \mathfrak{g}$ is defined by

$$X^\#_x \doteq \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \mathrm{e}^{tX}x = \mathrm{d}(\mathcal{O}_x)_{e_G}(X).$$

Obviously $(X+Y)^\# = X^\# + Y^\#$ and $(\lambda X)^\# = \lambda X^\#$, whenever $X, Y \in \mathfrak{g}$ and $\lambda \in \mathbb{R}$. So we can put the pieces together and state the:

Theorem 28 (Noether)**.** *Let $L\colon TQ \to \mathbb{R}$ be a G-invariant Lagrangian. Then, for every element $X \in \mathfrak{g}$, the map $\mathcal{J}^X\colon TQ \to \mathbb{R}$ (called the* Noether charge generated by X*) defined by $\mathcal{J}^X(x, v) = \mathbb{F}L(x,v)X^\#_x$, is constant along L-critical curves in Q.*

Remark. Note indeed that for $G = \mathbb{R}$ with $s \cdot x = \varphi_s(x)$ and $1 \in \mathrm{Lie}(\mathbb{R})$, we have that $1^\#_x = \partial_s\big|_{s=0}\varphi_s(x)$, so Theorem 25 is a particular case of Theorem 28 above, with the Noether charge there being the Noether charge here generated by 1.

Proof: Fix $X \in \mathfrak{g}$ and define an action $\mathbb{R} \circlearrowright Q$ by $s \cdot x \doteq \mathrm{e}^{sX}x$. If $\gamma_X\colon \mathbb{R} \to G$ is given by $\gamma_X(s) = \mathrm{e}^{sX}$, observe that given $x \in Q$, we have $\mathcal{O}^{\mathbb{R}}_x(s) = (\mathcal{O}^G_x \circ \gamma_X)(s)$. Thus

$$1^{\#_{\mathbb{R}}}_x = \mathrm{d}(\mathcal{O}^{\mathbb{R}}_x)_0(X) = \mathrm{d}(\mathcal{O}^G_x \circ \gamma_X)_0(X) = \mathrm{d}(\mathcal{O}^G_x)_{\gamma_X(0)}(\gamma_X'(0)) = \mathrm{d}(\mathcal{O}^G_x)_{e_G}(X) = X^{\#_G}_x,$$

since $\gamma_X(0) = e_G$ and $\gamma_X'(0) = X$. So the Noether charge $\mathcal{J}^X$ generated by X via the action of G equals the Noether charge $\mathcal{J}$ generated by 1 via the action of $\mathbb{R}$. Clearly L is $\mathbb{R}$-invariant (since it is G-invariant), and thus such charge is constant along L-critical curves, by Theorem 25. $\square$

Remark.

- We have observed that map $\mathfrak{g} \in X \mapsto X^\# \in \mathfrak{X}(Q)$ is linear, and since $\mathbb{F}L$ takes values in T^*Q, we may see the charge $\mathcal{J}$ as a linear map $\mathfrak{g} \ni X \mapsto \mathcal{J}^X \in \mathscr{C}^\infty(TQ)$. So we get $m = \dim \mathfrak{g}$ independent conservation laws.
- From a logical standpoint, it follows from the above proof the remarkable fact that Noether's theorem for $\mathbb{R}$-actions is equivalent to its version for more general G-actions.

We'll conclude the discussion with two last examples:

Example 29.

(1) Let $Q = G$ acting on itself, $\alpha \in \Omega^1(G)$ and consider the fiberwise linear Lagrangian $L_\alpha\colon TG \to \mathbb{R}$ given by $L_\alpha(x, v) = \alpha_x(v)$.

- Assume the action is given by left translations (i.e., $g \cdot x \doteq gx$) and that α is a left-invariant 1-form, that is, satisfying $(\mathrm{L}_g)^*\alpha = \alpha$ for all $g \in G$. Then L_α is G-invariant, as

  $$L(\mathrm{L}_g(x), \mathrm{d}(\mathrm{L}_g)_x(v)) = \alpha_{\mathrm{L}_g(x)}\big(\mathrm{d}(\mathrm{L}_g)_x(v)\big) = ((\mathrm{L}_g)^*\alpha)_x(v) = \alpha_x(v) = L(x, v),$$

  for all $g \in G$. So, given $X \in \mathfrak{g}$, we compute the infinitesimal action of X as

  $$X^\#_x = \mathrm{d}(\mathcal{O}_x)_{e_G}(X) = \mathrm{d}(\mathrm{R}_x)_{e_G}(X) = (X^R)_x,$$

where $X^R \in \mathfrak{X}^R(G)$ is the right-invariant vector field defined by X. Then, the Noether charge generated by X is $\mathcal{J}^X(x,v) = \mathbb{F}L_\alpha(x,v)X^\#_x = \alpha_x((X^R)_x)$. By Noether's theorem, we conclude that if $x\colon [a,b] \to G$ is any L_α-critical curve, i.e., satisfying the condition $\mathrm{d}\alpha_{x(t)}(\dot{x}(t),\cdot) = 0$ (seen in Example 10), then the function $[a,b] \ni t \mapsto \alpha_{x(t)}((X^R)_{x(t)}) \in \mathbb{R}$ is constant.

- Assume the action is given by right translations (i.e., $g \cdot x \doteq xg^{-1}$; the inverse is needed here to make this a left action, or else we'd have a right action) and that α is a right-invariant 1-form, that is, satisfying $(\mathrm{R}_g)^*\alpha = \alpha$ for all $g \in G$. Like in the previous case, L_α is G-invariant. So, given $X \in \mathfrak{g}$, we compute the infinitesimal action of X as

$$\begin{aligned} X^\#_x &= \mathrm{d}(\mathcal{O}_x)_{e_G}(X) = \mathrm{d}(\mathrm{L}_x \circ \mathrm{inv})_{e_G}(X) \\ &= \mathrm{d}(\mathrm{L}_x)_{e_G} \circ \mathrm{d}(\mathrm{inv})_{e_G}(X) = \mathrm{d}(\mathrm{L}_x)_{e_G}(-X) \\ &= -(X^L)_x, \end{aligned}$$

where $\mathrm{inv}\colon G \to G$ is the inversion mapping[3], $X^L \in \mathfrak{X}^L(G)$ is the left-invariant vector field defined by X. This time, the Noether charge generated by X is $\mathcal{J}^X(x,v) = \mathbb{F}L_\alpha(x,v)X^\#_x = -\alpha_x((X^L)_x)$. By Noether's theorem, we conclude that $[a,b] \ni t \mapsto \alpha_{x(t)}((X^L)_{x(t)}) \in \mathbb{R}$ is constant along L_α-critical curves x.

One could also consider the case where the G acts on itself by conjugation (i.e., $g \cdot x = gxg^{-1}$) and that α is a bi-invariant 1-form, that is, both left and right-invariant. But this does not give us anything new, since in this case both $\alpha_x((X^L)_x)$ and $\alpha_x((X^R)_x)$ are already constant functions of x, for any $X \in \mathfrak{g}$. Mimicking the argument given in the previous two cases (which indeed say something, as first we have that left-invariant α acts on right-invariant X^R, and then right-invariant α acts on left-invariant X^L), we would have that L_α is G-invariant, and for $X \in \mathfrak{g}$ the action field is $X^\#_x = (X^R)_x - (X^L)_x$ (which makes sense, as we're combining the previous two actions, hence combining the infinitesimal actions as well). The Noether charge generated by X is

$$\mathcal{J}^X(x,v) = \alpha_x((X^R)_x) - \alpha_x((X^L)_x).$$

Since $(X^R)_{e_G} - (X^L)_{e_G} = X - X = 0$, Noether's theorem gives that along L_α-critical curves x, we have that $\alpha_{x(t)}((X^R)_{x(t)}) = \alpha_{x(t)}((X^L)_{x(t)})$ for all t.

(2) Let $Q = \mathbb{R}^3$ with $G = \mathrm{SO}(3)$ acting via evaluation, and consider a mass $\mathrm{m} > 0$. Consider the natural Lagrangian $L\colon T\mathbb{R}^3 \cong \mathbb{R}^6 \to \mathbb{R}$ given by

$$L(x,y,z,\dot{x},\dot{y},\dot{z}) = \frac{\mathrm{m}}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2) - f(x^2+y^2+z^2),$$

where this time $f\colon \mathbb{R} \to \mathbb{R}$ is smooth. In other words, the potential energy depends only on the distance between the (base) point (x,y,z) and the origin of $\mathbb{R}^3$, and thus

[3]Satisfying $\mathrm{d}(\mathrm{inv})_{e_G} = -\mathrm{Id}_{\mathfrak{g}}$.

L is SO(3)-invariant. Gradients of such potential functions are called *central force fields*. In this case, we have that

$$\boldsymbol{F}(x,y,z) = -\frac{2}{\mathrm{m}} f'(x^2+y^2+z^2)(x\partial_x\big|_{(x,y,z)} + y\partial_y\big|_{(x,y,z)} + z\partial_z\big|_{(x,y,z)}),$$

and thus for any L-critical γ we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{m}\gamma(t)\times\dot{\gamma}(t) &= \mathrm{m}\dot{\gamma}(t)\times\dot{\gamma}(t) + \mathrm{m}\gamma(t)\times\ddot{\gamma}(t) \\ &= \mathrm{m}\gamma(t)\times\boldsymbol{F}(\gamma(t)) \\ &= \mathrm{m}\gamma(t)\times\left(-\frac{2}{\mathrm{m}}f'(\|\gamma(t)\|^2)\gamma(t)\right) \\ &= \mathbf{0}.\end{aligned}$$

This means that the angular momentum $\mathrm{m}\gamma(t)\times\dot{\gamma}(t)$ is constant. Now, recall that given $\boldsymbol{v} = (a,b,c) \in \mathbb{R}^3$, the matrix representing the linear map $\boldsymbol{v}\times \_ \colon \mathbb{R}^3 \to \mathbb{R}^3$ with respect to the standard basis is

$$\mathsf{A}_{\boldsymbol{v}} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix},$$

and the map $\mathbb{R}^3 \ni \boldsymbol{v} \mapsto \mathsf{A}_{\boldsymbol{v}} \in \mathfrak{so}(3)$ is a Lie algebra isomorphism, as can be checked by a direct computation with matrices (linearity is obvious), or by using the double cross product formula (here $\langle\cdot,\cdot\rangle$ denotes the standard inner product in $\mathbb{R}^3$):

$$\begin{aligned}[\mathsf{A}_{\boldsymbol{v}},\mathsf{A}_{\boldsymbol{w}}]\boldsymbol{p} &= \boldsymbol{v}\times(\boldsymbol{w}\times\boldsymbol{p}) - \boldsymbol{w}\times(\boldsymbol{v}\times\boldsymbol{p}) \\ &= \langle\boldsymbol{v},\boldsymbol{p}\rangle\boldsymbol{w} - \langle\boldsymbol{v},\boldsymbol{w}\rangle\boldsymbol{p} - \langle\boldsymbol{w},\boldsymbol{p}\rangle\boldsymbol{v} + \langle\boldsymbol{w},\boldsymbol{v}\rangle\boldsymbol{p} \\ &= \langle\boldsymbol{v},\boldsymbol{p}\rangle\boldsymbol{w} - \langle\boldsymbol{w},\boldsymbol{p}\rangle\boldsymbol{v} \\ &= \boldsymbol{p}\times(\boldsymbol{w}\times\boldsymbol{v}) \\ &= (\boldsymbol{v}\times\boldsymbol{w})\times\boldsymbol{p} \\ &= \mathsf{A}_{\boldsymbol{v}\times\boldsymbol{w}}\boldsymbol{p}.\end{aligned}$$

Anyway, this means that via the above isomorphism, we can generate Noether charges associated to vectors in $\mathbb{R}^3$. Since the action is given by evaluation (and the total derivative of a linear map at any point is itself), we clearly have that the infinitesimal action is given by $\boldsymbol{w}^\#_{\boldsymbol{p}} = \boldsymbol{w}\times\boldsymbol{p} \in T_{\boldsymbol{p}}\mathbb{R}^3$. With this, we compute

$$\mathcal{J}^{\boldsymbol{w}}(\boldsymbol{p},\boldsymbol{v}) = \mathbb{F}L(\boldsymbol{p},\boldsymbol{v})\boldsymbol{w}^\#_{\boldsymbol{p}} = \mathrm{m}\langle\boldsymbol{v},\boldsymbol{w}^\#_{\boldsymbol{p}}\rangle = \mathrm{m}\langle\boldsymbol{v},\boldsymbol{w}\times\boldsymbol{p}\rangle = \mathrm{m}\det(\boldsymbol{p},\boldsymbol{v},\boldsymbol{w}).$$

Noether's theorem gives that along a motion γ of the mechanical system considered, for every $\boldsymbol{w} \in \mathbb{R}^3$, the map $t \mapsto \mathrm{m}\det(\gamma(t),\dot{\gamma}(t),\boldsymbol{w})$ is a constant (which depends on $\boldsymbol{w}$). Taking $\boldsymbol{w}$ to be each one of the vectors in the standard basis of $\mathbb{R}^3$, we conclude again that the angular momentum $\mathrm{m}\gamma(t)\times\dot{\gamma}(t)$ is constant. To wit, writing $\gamma(t) = (x(t),y(t),z(t))$, we get constants $c_1,c_2,c_3 \in \mathbb{R}$ such that

$$\begin{cases} c_1 = \mathrm{m}(y(t)\dot{z}(t) - \dot{y}(t)z(t)) \\ c_2 = \mathrm{m}(\dot{x}(t)z(t) - x(t)\dot{z}(t)) \\ c_3 = \mathrm{m}(x(t)\dot{y}(t) - \dot{x}(t)y(t)), \end{cases}$$

and the expressions on the right side are the components of $\mathrm{m}\gamma(t) \times \dot{\gamma}(t)$.

4 Further generalizations on Lagrangians

What if instead of considering curves in Q, we considered m-surfaces? Fix $m \geq 1$ and let $\Omega \subseteq \mathbb{R}^m$ be a compact subset with non-empty interior and regular boundary $\partial\Omega$. Let's denote points in Ω by $u = (u^1, \ldots, u^m)$ and say that a map $x\colon \Omega \to Q$ is a m-surface. A variation of x is a map $\widetilde{x}\colon (-\epsilon, \epsilon) \times \Omega \to Q$ satisfying $\widetilde{x}(0, u) = x(u)$, and we'll say that $\widetilde{x}$ is a variation with fixed boundary if it also satisfies the condition $\widetilde{x}(s, \cdot)|_{\partial\Omega} = x\big|_{\partial\Omega}$ for all $s \in (-\epsilon, \epsilon)$. Now, our Lagrangians should be able to accept as inputs all the partial derivatives of x together.

So, before proceeding, we should understand how to describe the Whitney sum

$$TQ^{\oplus m} = \bigsqcup_{x \in Q} \underbrace{(T_xQ \oplus \cdots \oplus T_xQ)}_{m \text{ copies}}.$$

Its elements have the form $(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$, where $x \in Q$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m \in T_xQ$. Coordinates $(q^1, \ldots, q^m)$ will induce coordinates in each copy of TQ inside $TQ^{\oplus m}$, which we will treat as independent, and label with additional indices:

$$\left(q^1, \ldots, q^n, v^1_{(1)}, \ldots, v^n_{(1)}, \ldots, v^1_{(m)}, \ldots, v^n_{(m)}\right).$$

In the (frequent) case when $n = 1$, we omit the upper indices and the parentheses in the lower index, so that $v^1_{(\ell)} = v_\ell$. Observe also that $\dim TQ^{\oplus m} = n(m+1)$. This leads us to the:

Definition 30.

(i) A *Lagrangian on m parameters* is a map $L\colon \mathbb{R}^m \times TQ^{\oplus m} \to \mathbb{R}$, where $TQ^{\oplus m}$ denotes the Whitney sum of m copies of TQ. We will say that L is *autonomous* if it does not depend explicitly on the u-variable (or, in other words, if its domain is only $TQ^{\oplus m}$).

(ii) The *action functional of L* is defined by

$$\mathcal{A}_L[x] = \int_\Omega L\left(u, x(u), \frac{\partial x}{\partial u^1}(u), \ldots, \frac{\partial x}{\partial u^m}(u)\right) \mathrm{d}u^1 \cdots \mathrm{d}u^m,$$

where $x\colon \Omega \to Q$ is a m-surface.

(iii) We will say that x is L-critical if

$$\mathrm{d}(\mathcal{A}_L)_x(\widetilde{x}) \doteq \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathcal{A}_L[\widetilde{x}(s, \cdot)] = 0$$

for every variation $\widetilde{x}$ of x.

Remark. For convenience of notation, we set $\mathrm{d}u = \mathrm{d}u^1 \cdots \mathrm{d}u^m$ and $\nabla x(u)$ for the vector of partial derivatives of x, to simply write

$$\mathcal{A}_L[x] = \int_\Omega L(u, x(u), \nabla x(u))\, \mathrm{d}u.$$

Let's see how our old examples can be generalized:

Example 31.

(1) Let (Q,g) be a pseudo-Riemannian manifold, and $V\colon \mathbb{R}^m \times Q \to \mathbb{R}$ a smooth map. A natural Lagrangian on m parameters $L\colon \mathbb{R}^m \times TQ^{\oplus m} \to \mathbb{R}$ can be defined by

$$L(u,x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m) = \frac{1}{2}\sum_{r=1}^m g_x(\boldsymbol{v}_r,\boldsymbol{v}_r) - V(u,x).$$

(2) Let $\Phi \in \Gamma((T^*Q)^{\otimes m})$ be a field of covariant m-tensors on Q. Define the Lagrangian associated to Φ by $L_\Phi\colon TQ^{\oplus m} \to \mathbb{R}$ by $L_\Phi(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m) = \Phi_x(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m)$. Of particular interest will be the case when Φ is a m-form, in which case we have that

$$\mathcal{A}_L[x] = \int_\Omega \Phi_{x(u)}\left(\frac{\partial x}{\partial u^1}(u),\ldots,\frac{\partial x}{\partial u^m}(u)\right)\,\mathrm{d}u = \int_\Omega x^*\Phi.$$

The Euler-Lagrange equations have the following generalization:

Proposition 32. *Let $L\colon \mathbb{R}^m \times TQ^{\oplus m} \to \mathbb{R}$ be a Lagrangian on m parameters. Then a m-surface $x\colon \Omega \to Q$ is L-critical if and only if for any natural coordinate system on $TQ^{\oplus m}$ the* Euler-Lagrange equations *hold:*

$$\frac{\partial L}{\partial q^k}(u,x(u),\nabla x(u)) - \sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\frac{\partial L}{\partial v^k_{(\ell)}}(u,x(u),\nabla x(u))\right) = 0, \qquad k = 1,\ldots,n.$$

Proof: The argument is the same given for $m=1$ previously, but let's just deduce the equations for completeness: describe everything in coordinates as

$$(x(u),\nabla x(u)) = \left(\widetilde{q}^1(s,u),\ldots,\widetilde{q}^n(s,u),\widetilde{v}^1_{(1)}(s,u),\ldots,\widetilde{v}^n_{(1)}(s,u),\ldots,\widetilde{v}^1_{(m)}(s,u),\ldots,\widetilde{v}^n_{(m)}(s,u)\right),$$

which is to say that all the relations

$$\widetilde{v}^k_{(\ell)}(s,u) = \frac{\partial \widetilde{q}^K}{\partial u^\ell}(s,u), \qquad k = 1,\ldots,n, \quad \ell = 1,\ldots,m$$

hold. So we repeat the strategy from before:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0}\int_\Omega L(u,\widetilde{x}(s,u),\nabla_u\widetilde{x}(s,u))\,\mathrm{d}u = \int_\Omega \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} L(u,\widetilde{x}(s,u),\nabla_u\widetilde{x}(s,u))\,\mathrm{d}u \\
&= \int_\Omega \sum_{k=1}^n \frac{\partial L}{\partial q^k}(u,x(u),\nabla x(u))\frac{\partial \widetilde{q}^k}{\partial s}(0,u) + \sum_{k=1}^n\sum_{\ell=1}^m \frac{\partial L}{\partial v^k_{(\ell)}}(u,x(u),\nabla x(u))\frac{\partial \widetilde{v}^k_{(\ell)}}{\partial s}(0,u)\,\mathrm{d}u \\
&= \int_\Omega \sum_{k=1}^n \frac{\partial L}{\partial q^k}(u,x(u),\nabla x(u))\frac{\partial \widetilde{q}^k}{\partial s}(0,u) + \sum_{k=1}^n\sum_{\ell=1}^m \frac{\partial L}{\partial v^k_{(\ell)}}(u,x(u),\nabla x(u))\frac{\partial^2 \widetilde{q}^k}{\partial u^\ell \partial s}(0,u)\,\mathrm{d}u \\
&\stackrel{(*)}{=} \int_\Omega \sum_{k=1}^n \frac{\partial L}{\partial q^k}(u,x(u),\nabla x(u))\frac{\partial \widetilde{q}^k}{\partial s}(0,u) - \sum_{k=1}^n\sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\frac{\partial L}{\partial v^k_{(\ell)}}(u,x(u),\nabla x(u))\right)\frac{\partial \widetilde{q}^k}{\partial s}(0,u)\,\mathrm{d}u \\
&= \int_\Omega \sum_{k=1}^n \left(\frac{\partial L}{\partial q^k}(u,x(u),\nabla x(u)) - \sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\frac{\partial L}{\partial v^k_{(\ell)}}(u,x(u),\nabla x(u))\right)\right)\frac{\partial \widetilde{q}^k}{\partial s}(0,u)\,\mathrm{d}u,
\end{aligned}$$

where we use in $(*)$ the multivariable integration by parts formula together with the condition $\widetilde{x}(s,\cdot)\big|_{\partial\Omega} = x\big|_{\partial\Omega}$, which ensures that the boundary terms vanish. $\square$

Example 33 (Minimal surfaces)**.** Let $f\colon \Omega \to \mathbb{R}$ be smooth and consider the graph $\mathrm{gr}(f)$ of f, a hypersurface in $\mathbb{R}^{m+1} = \mathbb{R}^m \times \mathbb{R}$. It has the standard Monge parametrization $\varphi\colon \Omega \to \mathrm{gr}(f)$ given by $\varphi(u) = (u, f(u))$. Fix the curve $\varphi[\partial\Omega]$ in $\mathbb{R}^{m+1}$. Among all the graphs which share $\varphi[\partial\Omega]$ as boundary, which one has least "area" (i.e., m-volume)? In other words, when does f minimize the area integral

$$\mathrm{Area}(\mathrm{gr}(f)) = \int_\Omega \sqrt{1 + \|\nabla f(u)\|^2}\,\mathrm{d}u?$$

We can model this situation with what we have done so far by considering $Q = \mathbb{R}$ and the Lagrangian $L\colon T\mathbb{R}^{\oplus m} \to \mathbb{R}$ given by $L(x, v_1, \ldots, v_m) = \sqrt{1 + v_1^2 + \cdots + v_m^2}$. Thus

$$\mathrm{Area}(\mathrm{gr}(f)) = \mathscr{A}_L[f] = \int_\Omega L(f(u), \nabla f(u))\,\mathrm{d}u.$$

Let's compute the Euler-Lagrange equations for this L. We clearly have

$$\frac{\partial L}{\partial x}(x, v_1, \ldots, v_m) = 0 \quad \text{and} \quad \frac{\partial L}{\partial v_\ell}(x, v_1, \ldots, v_m) = \frac{v_\ell}{\sqrt{1 + v_1^2 + \cdots + v_m^2}},$$

and so

$$\begin{aligned} 0 &= \sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\frac{\partial L}{\partial v_\ell}(f(u), \nabla f(u))\right) - \frac{\partial L}{\partial x}(f(u), \nabla f(u)) \\ &= \sum_{\ell=1}^n \frac{\partial}{\partial u^\ell}\left(\frac{(\partial f/\partial u^\ell)(u)}{\sqrt{1 + \|\nabla f(u)\|^2}}\right) \\ &= \mathrm{div}\left(\frac{\nabla f}{\sqrt{1 + \|\nabla f\|^2}}\right)(u). \end{aligned}$$

We call the relation

$$\mathrm{div}\left(\frac{\nabla f}{\sqrt{1 + \|\nabla f\|^2}}\right) = 0$$

the *minimal surface equation in divergence form*. This is also related to the so-called *mean curvature* of $\mathrm{gr}(f)$, which is defined (up to a $\pm 1/m$ factor) as the trace of the second fundamental form of $\mathrm{gr}(f)$ in $\mathbb{R}^{m+1}$. It can be computed as $H = \sum_{i,j=1}^m g^{ij} h_{ij}$, where

$$g_{ij} = \left\langle \frac{\partial \varphi}{\partial u^i}, \frac{\partial \varphi}{\partial u^j} \right\rangle, \qquad h_{ij} = \left\langle \boldsymbol{N}, \frac{\partial^2 \varphi}{\partial u^i \partial u^j} \right\rangle \quad \text{and} \quad (g^{ij})_{i,j=1}^m = \left((g_{ij})_{i,j=1}^m\right)^{-1},$$

and $\boldsymbol{N}$ is a unit normal vector field along $\mathrm{gr}(f)$ in $\mathbb{R}^{m+1}$. If $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$ is the standard basis of $\mathbb{R}^m$, we have that

$$\frac{\partial \varphi}{\partial u^i}(u) = \left(\boldsymbol{e}_i, \frac{\partial f}{\partial u^i}(u)\right) \quad \text{and} \quad \frac{\partial^2 \varphi}{\partial u^i \partial u^j}(u) = \left(\boldsymbol{0}, \frac{\partial^2 f}{\partial u^i \partial u^j}(u)\right),$$

and so

$$g_{ij} = \delta_{ij} + \frac{\partial f}{\partial u^i}\frac{\partial f}{\partial u^j} \implies g^{ij} = \delta^{ij} - \frac{1}{1 + \|\nabla f\|^2}\frac{\partial f}{\partial u^i}\frac{\partial f}{\partial u^j},$$

for all $i, j = 1, \ldots, m$ (for example, by the Sherman-Morrison formula[4]). Now, a unit normal field along $\mathrm{gr}(f)$ in $\mathbb{R}^{m+1}$ is[5]

$$\boldsymbol{N} = \frac{(-\nabla f, 1)}{\sqrt{1 + \|\nabla f\|^2}},$$

which immediately yields

$$h_{ij} = \frac{1}{\sqrt{1 + \|\nabla f\|^2}} \frac{\partial^2 f}{\partial u^i \partial u^j}.$$

Thus, if $\nabla^2 f$ denotes the Hessian of f, we can put all of this together and compute

$$\begin{aligned} H &= \sum_{i,j=1}^{m} g^{ij} h_{ij} \\ &= \sum_{i,j=1}^{n} \left(\delta^{ij} - \frac{1}{1 + \|\nabla f\|^2} \frac{\partial f}{\partial u^i} \frac{\partial f}{\partial u^j} \right) \left(\frac{1}{\sqrt{1 + \|\nabla f\|^2}} \frac{\partial^2 f}{\partial u^i \partial u^j} \right) \\ &= \frac{\triangle f}{\sqrt{1 + \|\nabla f\|^2}} - \frac{(\nabla^2 f)(\nabla f, \nabla f)}{(1 + \|\nabla f\|^2)^{3/2}} \\ &= \operatorname{div} \left(\frac{\nabla f}{\sqrt{1 + \|\nabla f\|^2}} \right), \end{aligned}$$

in view of the product rule[6] for div. Thus, we conclude that a f is a critical point of the area functional if and only if $\mathrm{gr}(f)$ has zero mean curvature. In the case when $m = 2$, eliminating the common denominator and setting $(u^1, u^2) = (u, v)$, we have the classical minimal surface equation

$$\left(1 + \left(\frac{\partial f}{\partial v}\right)^2\right) \frac{\partial^2 f}{\partial u^2} - 2 \frac{\partial f}{\partial u} \frac{\partial f}{\partial v} \frac{\partial^2 f}{\partial u \partial v} + \left(1 + \left(\frac{\partial f}{\partial u}\right)^2\right) \frac{\partial^2 f}{\partial v^2} = 0$$

Example 34 (Harmonic functions)**.** For $Q = \mathbb{R}$, look at $L \colon T\mathbb{R}^{\oplus m} \to \mathbb{R}$ given by

$$L(x, v_1, \ldots, v_m) = \frac{1}{2}(v_1^2 + \cdots + v_m^2)$$

[4]If A is a non-singular square matrix and u and v are column vectors, then $A + uv^\top$ is non-singular if and only if $1 + v^\top A^{-1} u \neq 0$, in which case we have

$$(A + uv^\top)^{-1} = A^{-1} - \frac{A^{-1} u v^\top A^{-1}}{1 + v^\top A^{-1} u}.$$

Our case here is $A = \mathrm{Id}_m$ and $u = v = \nabla f$.

[5]You can do the case $m = 2$ to see what is going on, the general case then becomes clear.

[6]For a vector bundle $E \to (M, g)$ over a pseudo-Riemannian manifold, equipped with a linear connection ∇, we define the divergence of $\sigma \in \Gamma(T^*M \otimes E)$ as $\operatorname{div} \sigma = \operatorname{tr}_g\big((\boldsymbol{X}, \boldsymbol{Y}) \mapsto (\nabla_{\boldsymbol{X}} \sigma)(\boldsymbol{Y})\big)$, and the formula $\operatorname{div}(f\sigma) = f \operatorname{div} \sigma + \sigma(\operatorname{grad}_g f)$ holds for $f \in \mathscr{C}^\infty(M)$.

For a smooth function $\psi\colon \Omega \to \mathbb{R}$, the action functional of L is

$$\mathscr{A}_L[\psi] = \frac{1}{2}\int_\Omega \left(\frac{\partial\psi}{\partial u^1}(u)\right)^2 + \cdots + \left(\frac{\partial\psi}{\partial u^m}(u)\right)^2 \,\mathrm{d}u = \frac{1}{2}\int_\Omega \|\nabla\psi(u)\|^2\,\mathrm{d}u.$$

This is called the *Dirichlet energy* of ψ. As

$$\frac{\partial L}{\partial x}(x, v_1, \ldots, v_m) = 0 \quad \text{and} \quad \frac{\partial L}{\partial v_\ell}(x, v_1, \ldots, v_m) = v_\ell,$$

we have that ψ is L-critical if and only if

$$0 = \sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\frac{\partial L}{\partial v_\ell}(\psi(u), \nabla\psi(u))\right) - \frac{\partial L}{\partial x}(\psi(u), \nabla\psi(u)) = \sum_{\ell=1}^m \frac{\partial^2\psi}{(\partial u^\ell)^2}(u) = \triangle\psi(u).$$

That is, ψ is L-critical if and only if ψ is harmonic ($\triangle\psi = 0$). Since any $A \in \mathrm{O}(m,\mathbb{R})$ is a diffeomorphism of $\mathbb{R}^m$ preserving L, we conclude that if ψ is L-critical, so is $\psi\circ A$, which is to say that if $\triangle\psi = 0$, then $\triangle(\psi\circ A) = 0$ (of course, actually the stronger relation $\triangle(\psi\circ A) = \triangle\psi\circ A$ holds).

Remark. Generalizing even further what we're doing in this section and instead of m-surfaces $\Omega \to Q$, considering maps where the domain is another pseudo-Riemannian manifold, one can prove in the general setting of manifolds that critical points for the Dirichlet energy functional are harmonic functions.

Example 35. Let's go back to Example 31.

(1) The natural Lagrangian $L\colon \mathbb{R}^m \times TQ^{\oplus m} \to \mathbb{R}$ given by

$$L(u, x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \frac{1}{2}\sum_{r=1}^m g_x(\boldsymbol{v}_r, \boldsymbol{v}_r) - V(u,x)$$

on a pseudo-Riemannian manifold (Q, g) is described in coordinates by

$$L(u, x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \frac{1}{2}\sum_{i,j=1}^n\sum_{r=1}^m g_{ij}(x)v^i_{(r)}v^j_{(r)} - V(u,x),$$

whence

$$\frac{\partial L}{\partial q^k}(u, x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \frac{1}{2}\sum_{i,j=1}^n\sum_{r=1}^m \frac{\partial g_{ij}}{\partial q^k}(x)v^i_{(r)}v^j_{(r)} - \frac{\partial V}{\partial q^k}(u,x)$$

and

$$\frac{\partial L}{\partial v^k_{(\ell)}}(u, x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \sum_{i=1}^n g_{ik}(x)v^i_{(\ell)}.$$

Putting all of this together, we compute the Euler-Lagrange equations just like in Example 9 (p. 9) as

$$\sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\frac{\partial L}{\partial v^k_{(\ell)}}(u, x(u), \nabla x(u))\right) - \frac{\partial L}{\partial q^k}(u, x(u), \nabla x(u))$$
$$= \sum_{i=1}^n\sum_{\ell=1}^m g_{ik}(x(u))\frac{\partial v^i_{(\ell)}}{\partial u^\ell}(u) + \sum_{i,j=1}^n\sum_{\ell=1}^m \Gamma_{ijk}(x(u))v^i_{(\ell)}(u)v^j_{(\ell)}(u) + \frac{\partial V}{\partial q^k}(u, x(u)),$$

so that setting this equal to zero and raising k yields

$$\sum_{\ell=1}^{m}\left(\frac{\partial^2 q^k}{(\partial u^\ell)^2}(u)+\sum_{i,j=1}^{n}\Gamma_{ij}^k(x(u))\frac{\partial q^i}{\partial u^\ell}(u)\frac{\partial q^j}{\partial u^\ell}(u)\right)=-\big((\mathrm{grad}_g V(u,\cdot))_{x(u)}\big)^k.$$

This means that if $\boldsymbol{F}\colon \mathbb{R}^m\times Q\to TQ$ is given by $\boldsymbol{F}(u,x)=-(\mathrm{grad}_g V(u,\cdot))_x$, then the condition to be L-critical is a generalized version of Newton's equation:

$$\sum_{\ell=1}^{m}\frac{D}{\partial u^\ell}\left(\frac{\partial x}{\partial u^\ell}\right)=\boldsymbol{F}(u,x(u)).$$

The above is a suggestive notation for what follows: we use the map x to pull back the Levi-Civita connection ∇ in TQ to a connection $x^*\nabla$ in the bundle $x^*(TQ)$ over Ω, and the condition is $\sum_{\ell=1}^m (x^*\nabla)_{\partial/\partial u^\ell}(\partial x/\partial u^\ell)=\boldsymbol{F}(u,x(u))$.

In particular, if $Q=\mathbb{R}^n$ and $g=\mathrm{m}\langle\cdot,\cdot\rangle$ (where $\mathrm{m}>0$ is a fixed mass and $\langle\cdot,\cdot\rangle$ is the standard inner product), the Levi-Civita connection is the standard flat connection, and Newton's equation reads $\triangle x=\boldsymbol{F}$, where $\triangle$ is the Laplace operator applied componentwise.

(2) Let $\Phi\in\Omega^m(Q)$ be a m-form, and consider the Lagrangian $L_\Phi\colon TQ^{\oplus m}\to\mathbb{R}$ defined by Φ. Here, we consider Φ to be totally skew instead of merely a field of covariant m-tensors, to be able to use the exterior derivative $\mathrm{d}\Phi\in\Omega^{m+1}(Q)$. We have a direct generalization of what happened with 1-forms, in Example 10 (p. 10). Namely, a m-surface will be L_Φ-critical if and only if it satisfies the relation

$$\mathrm{d}\Phi_{x(u)}\left(\frac{\partial x}{\partial u^1}(u),\ldots,\frac{\partial x}{\partial u^m}(u),\cdot\right)=0.$$

To illustrate, let's write the Euler-Lagrange equations for the $m=2$. We start with

$$\Phi=\sum_{i,j=1}^{n}\Phi_{ij}\,\mathrm{d}q^i\otimes\mathrm{d}q^j \implies L_\Phi(x,\boldsymbol{v}_1,\boldsymbol{v}_2)=\sum_{i,j=1}^{n}\Phi_{ij}(x)v_{(1)}^i v_{(2)}^j.$$

It follows that $\dfrac{\partial L_\Phi}{\partial q^k}(x,\boldsymbol{v}_1,\boldsymbol{v}_2)=\displaystyle\sum_{i,j=1}^{n}\frac{\partial\Phi_{ij}}{\partial q^k}(x)v_{(1)}^i v_{(2)}^j$ and also that

$$\frac{\partial L_\Phi}{\partial v_{(\ell)}^k}(x,\boldsymbol{v}_1,\boldsymbol{v}_2)=\sum_{i=1}^{n}\Phi_{ik}(x)(v_{(1)}^i\delta_2^\ell-v_{(2)}^i\delta_1^\ell),$$

by using the skew-symmetry of Φ. With this, we have that for a 2-surface described in coordinates as

$$(x(u),\nabla x(u))=\big(q^1(u),\ldots,q^n(u),v_{(1)}^1(u),\ldots,v_{(1)}^n(u),v_{(2)}^1(u),\ldots,v_{(2)}^n(u)\big),$$

the Euler-Lagrange equations read

$$\begin{aligned}
0 &= \frac{\partial L_\Phi}{\partial q^k}(x(u), \nabla x(u)) - \frac{\partial}{\partial u^1}\left(\frac{\partial L_\Phi}{\partial v^k_{(1)}}(x(u), \nabla x(u))\right) - \frac{\partial}{\partial u^2}\left(\frac{\partial L_\Phi}{\partial v^k_{(2)}}(x(u), \nabla x(u))\right) \\
&= \sum_{i,j=1}^n \frac{\partial \Phi_{ij}}{\partial q^k}(x(u))v^i_{(1)}(u)v^j_{(2)}(u) - \frac{\partial}{\partial u^1}\left(-\sum_{i=1}^n \Phi_{ik}(x(u))v^i_{(2)}(u)\right) - \frac{\partial}{\partial u^2}\left(\sum_{i=1}^n \Phi_{ik}(x(u))v^i_{(1)}(u)\right) \\
&= \sum_{i,j=1}^n \frac{\partial \Phi_{ij}}{\partial q^k}(x(u))v^i_{(1)}(u)v^j_{(2)}(u) + \sum_{i,j=1}^n \frac{\partial \Phi_{ik}}{\partial q^j}(x(u))v^j_{(1)}(u)v^i_{(2)}(u) + \sum_{i=1}^n \Phi_{ik}(x(u))\frac{\partial v^i_{(2)}}{\partial u^1}(u) \\
&\quad - \sum_{i,j=1}^n \frac{\partial \Phi_{ik}}{\partial q^j}(x(u))v^j_{(2)}(u)v^i_{(1)}(u) - \sum_{i=1}^n \Phi_{ik}(x(u))\frac{\partial v^i_{(1)}}{\partial u^2}(u) \\
&\overset{(*)}{=} \sum_{i,j=1}^n \left(\frac{\partial \Phi_{ij}}{\partial q^k}(x(u)) + \frac{\partial \Phi_{jk}}{\partial q^i}(x(u)) - \frac{\partial \Phi_{ik}}{\partial q^j}(x(u))\right) v^i_{(1)}(u)v^j_{(2)}(u),
\end{aligned}$$

where in $(*)$ we use that

$$\frac{\partial v^i_{(1)}}{\partial u^2} = \frac{\partial^2 q^i}{\partial u^1 \partial u^2} = \frac{\partial v^i_{(2)}}{\partial u^1}$$

to make the single sums cancel. The last expression is precisely the k-th component of the 1-form

$$\mathrm{d}\Phi_{x(u)}\left(\frac{\partial x}{\partial u^1}(u), \frac{\partial x}{\partial u^2}(u), \cdot\right) = 0,$$

along the image $x[\Omega]$, so we're done. Like before, it also follows that

$$\mathrm{d}(\mathscr{A}_{L_\Phi})_x(\widetilde{x}) = \int_\Omega \mathrm{d}\Phi_{x(u)}\left(\frac{\partial x}{\partial u^1}(u), \ldots, \frac{\partial x}{\partial u^m}(u), \frac{\partial \widetilde{x}}{\partial s}(0,u)\right)\,\mathrm{d}u,$$

for any variation $\widetilde{x}$ of x with fixed boundary. The proof for general m is then just an exercise in notation.

(3) Consider again a pseudo Riemannian manifold (Q, g), and the fiberwise linear Lagrangian associated to the metric g itself, that is, $L_g\colon TQ^{\oplus 2} \to \mathbb{R}$ given by $L_g(x, \boldsymbol{v}_1, \boldsymbol{v}_2) = g_x(\boldsymbol{v}_1, \boldsymbol{v}_2)$. Let's compute the Euler-Lagrange equations for L_g, using the coordinate expression

$$L_g(x, \boldsymbol{v}_1, \boldsymbol{v}_2) = \sum_{i,j=1}^n g_{ij}(x)v^i_{(1)}v^j_{(2)}.$$

First, we compute

$$\frac{\partial L_g}{\partial q^k}(x, \boldsymbol{v}_1, \boldsymbol{v}_2) = \sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial q^k}(x)v^i_{(1)}v^j_{(2)} \text{ and } \frac{\partial L_g}{\partial v^k_{(\ell)}}(x, \boldsymbol{v}_1, \boldsymbol{v}_2) = \sum_{i=1}^k g_{ik}(x)\big(\delta^\ell_2 v^i_{(1)} + \delta^\ell_1 v^i_{(2)}\big).$$

Putting all of this together, we have

$$
\begin{aligned}
0 &= \frac{\partial}{\partial u^1}\left(\frac{\partial L}{\partial v^k_{(1)}}(x(u),\nabla x(u))\right) + \frac{\partial}{\partial u^2}\left(\frac{\partial L}{\partial v^k_{(2)}}(x(u),\nabla x(u))\right) - \frac{\partial L}{\partial q^k}(x(u),\nabla x(u)) \\
&= \frac{\partial}{\partial u^1}\left(\sum_{i=1}^n g_{ik}(x(u))v^i_{(2)}(u)\right) + \frac{\partial}{\partial u^2}\left(\sum_{i=1}^n g_{ik}(x(u))v^i_{(2)}(u)\right) - \sum_{i,j=1}^n \frac{\partial g_{ij}}{\partial q^k}(x(u))v^i_{(1)}(u)v^j_{(2)}(u) \\
&= 2\sum_{i=1}^n g_{ik}(x(u))\frac{\partial^2 q^i}{\partial u^1 \partial u^2}(u) + \sum_{i,j=1}^n \left(\frac{\partial g_{ik}}{\partial q^j}(x(u)) + \frac{\partial g_{jk}}{\partial q^i}(x(u)) - \frac{\partial g_{ij}}{\partial q^k}(x(u))\right) v^i_{(1)}(u)v^j_{(2)}(u),
\end{aligned}
$$

which is equivalent (by raising k) to

$$
\frac{\partial^2 q^k}{\partial u^1 \partial u^2}(x(u)) + \sum_{i,j=1}^n \Gamma^k_{ij}(x(u))\frac{\partial q^i}{\partial u^1}(u)\frac{\partial q^j}{\partial u^2}(u) = 0,
$$

for all $k = 1,\ldots,n$. We may write such condition as $\nabla_{\partial x/\partial u^1}(\partial x/\partial u^2) = \mathbf{0}$, where ∇ is the Levi-Civita connection of (Q,g), or simply

$$
\frac{D}{\partial u^1}\left(\frac{\partial x}{\partial u^2}\right) = \mathbf{0},
$$

which of course equals $(D/\partial u^2)(\partial x/\partial u^1)$. Again, if you want to be completely rigorous here, here's what is going on: we use the map $x\colon \Omega \to Q$ to pull back ∇ to a connection $x^*\nabla$ in the bundle $x^*(TQ)$ over Ω, and then we have that $(x^*\nabla)_{\partial/\partial u^1}(\partial x/\partial u^2) = \mathbf{0}$.

The notion of fiber derivative and what comes with it has a somewhat direct generalization:

Definition 36. Let $L\colon TQ^{\oplus m} \to \mathbb{R}$ be a Lagrangian on m parameters.

(i) The ℓ-th partial fiber derivative of L is the map $(\mathbb{F}L)^\ell\colon TQ^{\oplus m} \to T^*Q$ given by

$$
(\mathbb{F}L)^\ell(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m)\boldsymbol{w} = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} L(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{\ell-1},\boldsymbol{v}_\ell + t\boldsymbol{w},\boldsymbol{v}_{\ell+1},\ldots,\boldsymbol{v}_m).
$$

(ii) The total fiber derivative of L is the map $\mathbb{F}L\colon TQ^{\oplus m} \to T^*Q^{\oplus m}$ given by

$$
\mathbb{F}L(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m)(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_m) = \sum_{\ell=1}^m (\mathbb{F}L)^\ell(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m)\boldsymbol{w}_\ell.
$$

(iii) We'll say that L is regular or hyperregular according whether $\mathbb{F}L$ is a local diffeomorphism or a global diffeomorphism.

(iv) The *energy map* of L is the map $E_L\colon TQ^{\oplus m} \to \mathbb{R}$ given by

$$
E_L(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m) = \mathbb{F}L(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m)(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m) - L(x,\boldsymbol{v}_1,\ldots,\boldsymbol{v}_m).
$$

Remark. In item (ii) of the above definition we are using the natural identification between fibers $(T_xQ^{\oplus m})^* \cong T_x^*Q^{\oplus m}$. And just like calculus in $\mathbb{R}^n$, we have that

$$(\mathbb{F}L)^\ell(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)\boldsymbol{w} = \mathbb{F}L(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)(\boldsymbol{0}, \ldots, \boldsymbol{w}, \ldots, \boldsymbol{0}),$$

where obviously $\boldsymbol{w}$ in the right side enters in the ℓ-th slot.

The description of $\mathbb{F}L$ in coordinates is similar as in the single-variable case. Namely, coordinates in Q induce coordinates in each copy of T^*Q inside $T^*Q^{\oplus m}$, treated as independent, and thus we may write

$$(q^1, \ldots, q^n, p_1^{(1)}, \ldots, p_k^{(1)}, \ldots, p_1^{(m)}, \ldots, p_n^{(m)})$$

as coordinates for $T^*Q^{\oplus m}$, so that if $\mathbb{F}L(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = (x, \mathsf{p}^1, \ldots, \mathsf{p}^m)$, we have

$$p_k^{(\ell)}(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \frac{\partial L}{\partial v_{(\ell)}^k}(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m),$$

leading us to

$$\mathbb{F}L(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \left(x, \sum_{k=1}^n \frac{\partial L}{\partial v_{(1)}^k}(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)\, \mathrm{d}q^k \bigg|_x, \ldots, \sum_{k=1}^n \frac{\partial L}{\partial v_{(m)}^k}(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)\, \mathrm{d}q^k \bigg|_x \right),$$

as expected. We are also justified in writing $\mathsf{p}^\ell = (\mathbb{F}L)^\ell(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$.

Example 37.

(1) For the natural Lagrangian $L\colon TQ^{\oplus m} \to \mathbb{R}$ with potential energy $V\colon Q \to \mathbb{R}$, we have

$$\begin{aligned}
&\mathbb{F}L(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m) = \\
&\quad = \sum_{\ell=1}^m \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \left(\frac{1}{2} g_x(\boldsymbol{v}_\ell + t\boldsymbol{w}_\ell, \boldsymbol{v}_\ell + t\boldsymbol{w}_\ell) + \frac{1}{2}\sum_{k \neq \ell} g_x(\boldsymbol{v}_k, \boldsymbol{v}_k) - V(x) \right) \\
&\quad = \sum_{\ell=1}^m g_x(\boldsymbol{v}_\ell, \boldsymbol{w}_\ell).
\end{aligned}$$

Thus $\mathbb{F}L(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = (x, (\boldsymbol{v}_1)_\flat, \ldots, (\boldsymbol{v}_m)_\flat)$. Clearly the energy of L is given by

$$E_L(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = \frac{1}{2}\sum_{\ell=1}^n g_x(\boldsymbol{v}_\ell, \boldsymbol{v}_\ell) + V(x).$$

(2) For $\Phi \in \Omega^m(Q)$, we have

$$\begin{aligned}
\mathbb{F}L_\Phi(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m) &= \sum_{\ell=1}^m \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} \Phi_x(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{\ell-1}, \boldsymbol{v}_\ell + t\boldsymbol{w}_\ell, \boldsymbol{v}_{\ell+1}, \ldots, \boldsymbol{v}_m) \\
&= \sum_{\ell=1}^m \Phi_x(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{\ell-1}, \boldsymbol{w}_\ell, \boldsymbol{v}_{\ell+1}, \ldots, \boldsymbol{v}_m) \\
&= \sum_{\ell=1}^m L_\Phi(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{\ell-1}, \boldsymbol{w}_\ell, \boldsymbol{v}_{\ell+1}, \ldots, \boldsymbol{v}_m).
\end{aligned}$$

Compare this with the usual formula for the total derivative of a multilinear map defined in a product of vector spaces. The energy of L_Φ is

$$\begin{aligned} E_{L_\Phi}(x, v_1, \ldots, v_m) &= mL_\Phi(x, v_1, \ldots, v_m) - L_\Phi(x, v_1, \ldots, v_m) \\ &= (m-1)L_\Phi(x, v_1, \ldots, v_m). \end{aligned}$$

In both cases above, for $m = 1$ we indeed recover the computations done in Example 18 (p. 16).

We proceed to look for a version of Noether's theorem in this setting. Based on the experience we had on the previous section, we will start with the case $G = \mathbb{R}$:

Theorem 38 (Noether). *Let $L\colon TQ^{\oplus m} \to \mathbb{R}$ be a Lagrangian on m parameters, and $(\varphi_s)_{s\in\mathbb{R}}$ a 1-parameter group of diffeomorphisms of Q leaving L invariant. Then the* Noether current *$\mathcal{J} = (\mathcal{J}^1, \ldots, \mathcal{J}^m)\colon TQ^{\oplus m} \to \mathbb{R}^m$ given by*

$$\mathcal{J}^\ell(x, v_1, \ldots, v_m) = (\mathbb{F}L)^\ell(x, v_1, \ldots, v_m)\left(\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}\varphi_s(x)\right),$$

has zero divergence along L-critical m-surfaces $x\colon \Omega \to Q$. In other words, the pull-back map $\Omega \ni u \mapsto \mathcal{J}(x(u), \nabla x(u)) \in \mathbb{R}^m$ is a vector field along Ω, and $\operatorname{div}\mathcal{J}(x(u), \nabla x(u)) = 0$.

Remark. For $m = 1$, note that the divergence of a vector field $f \in \mathfrak{X}(\mathbb{R})$, which is nothing else than a smooth map $f\colon \mathbb{R} \to \mathbb{R}$, is $\operatorname{div}(f)(t) = f'(t)$.

Proof: The proof is the same as given in the case $m = 1$, with an extra summation due to the multivariable version of the Euler-Lagrange equations. Namely, we set $x_s = \varphi_s \circ x$, observe that by invariance all of the x_s are L-critical if x is, and that

$$\frac{\partial x_s}{\partial u^\ell}(u) = \mathrm{d}(\varphi_s)_{x(u)}\left(\frac{\partial x}{\partial u^\ell}(u)\right).$$

Describe everything in coordinates as

$$(x_s(u), \nabla x_s(u)) = \big(q^1(s,u), \ldots, q^n(s,u), v^1_{(1)}(s,u), \ldots, v^n_{(1)}(s,u), \ldots, v^1_{(m)}(s,u), \ldots, v^n_{(m)}(s,u)\big),$$

and take the s-derivative on both sides of $L(x_s(u), \nabla x_s(u)) = L(x(u), \nabla x(u))$ (valid for all $s \in \mathbb{R}$), to get

$$\begin{aligned} 0 &= \sum_{k=1}^n \frac{\partial L}{\partial q^k}(x_s(u), \nabla x_s(u))\frac{\partial q^k}{\partial s}(s,u) + \sum_{\ell=1}^m\sum_{k=1}^n \frac{\partial L}{\partial v^k_{(\ell)}}(x_s(u), \nabla x_s(u))\frac{\partial v^k_{(\ell)}}{\partial s}(s,u) \\ &= \sum_{k=1}^n\sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\frac{\partial L}{\partial v^k_{(\ell)}}(x_s(u), \nabla x_s(u))\right)\frac{\partial q^k}{\partial s}(s,u) + \sum_{\ell=1}^m\sum_{k=1}^n \frac{\partial L}{\partial v^k_{(\ell)}}(x_s(u), \nabla x_s(u))\frac{\partial}{\partial u^\ell}\left(\frac{\partial q^k}{\partial s}(s,u)\right) \\ &= \sum_{\ell=1}^m \frac{\partial}{\partial u^\ell}\left(\sum_{k=1}^n \frac{\partial L}{\partial v^k_{(\ell)}}(x_s(u), \nabla x_s(u))\frac{\partial q^k}{\partial s}(s,u)\right) \end{aligned}$$

Set $s = 0$. We are done. $\square$

Before proceeding, let's see some examples:

Example 39.

(1) Let (Q, g) be a pseudo-Riemannian manifold. Consider again the fiberwise bilinear Lagrangian associated to the metric g itself, that is, $L_g \colon TQ^{\oplus 2} \to \mathbb{R}$ given by

$$L_g(x, \boldsymbol{v}_1, \boldsymbol{v}_2) = g_x(\boldsymbol{v}_1, \boldsymbol{v}_2).$$

If $\boldsymbol{\xi} \in \mathfrak{X}(Q)$ is Killing, then the L_g is invariant under the flow of $\boldsymbol{\xi}$. The associated Noether current $\mathcal{J} \colon TQ^{\oplus 2} \to \mathbb{R}^2$ is given by

$$\mathcal{J}(x, \boldsymbol{v}_1, \boldsymbol{v}_2) = \big(g_x(\boldsymbol{\xi}_x, \boldsymbol{v}_2), g_x(\boldsymbol{v}_1, \boldsymbol{\xi}_x)\big).$$

If $x \colon \Omega \to Q$ satisfies $\nabla_{\partial x/\partial u^1}(\partial x/\partial u^2) = \boldsymbol{0}$, we have that the divergence of the vector field along Ω given by

$$u \mapsto \left(g_{x(u)}\left(\boldsymbol{\xi}_{x(u)}, \frac{\partial x}{\partial u^2}(u)\right), g_x\left(\frac{\partial x}{\partial u^1}(u), \boldsymbol{\xi}_{x(u)}\right)\right)$$

is zero. That gives

$$g_{x(u)}\left(\nabla_{\partial x/\partial u^1}\boldsymbol{\xi}, \frac{\partial x}{\partial u^2}(u)\right) + g_{x(u)}\left(\frac{\partial x}{\partial u^1}(u), \nabla_{\partial x/\partial u^2}\boldsymbol{\xi}\right) = 0.$$

The upshot here is that Noether's theorem gives us a (horrible and exceedingly complicated) proof of Killing's equation $\langle \nabla_{\boldsymbol{X}}\boldsymbol{\xi}, \boldsymbol{Y}\rangle + \langle \boldsymbol{X}, \nabla_{\boldsymbol{Y}}\boldsymbol{\xi}\rangle = 0$. The reason is that such equation is tensorial in $\boldsymbol{X}$ and $\boldsymbol{Y}$, and given any $x_0 \in Q$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in T_xQ$, there is a L_g-critical 2-surface $x \colon \Omega \to Q$ with $x(u_0) = x_0$ and $(\partial x/\partial u^\ell)(u_0) = \boldsymbol{v}_\ell$ (for $\ell = 1, 2$), where $u_0 \in \Omega$ is fixed before defining x.

(2) Noether's theorem gives an alternative proof of the fact that functions $\psi \colon \Omega \to \mathbb{R}$ which are critical points of the Dirichlet energy are harmonic functions, without needing to compute the Euler-Lagrange equations explicitly. The Lagrangian $L \colon T\mathbb{R}^{\oplus m} \to \mathbb{R}$ was given by

$$L(x, v_1, \ldots, v_m) = \frac{1}{2}(v_1^2 + \cdots + v_m^2),$$

and we have that for all $s \in \mathbb{R}$, the translations $\tau_s \colon \mathbb{R} \to \mathbb{R}$ given by $\tau_s(x) = x + s$ leave L invariant (since $\tau_s'(x) = 1$). It is easy to see that the associated Noether current is given by

$$\mathcal{J}(x, v_1, \ldots, v_m) = (v_1, \ldots, v_m),$$

and we have that $\mathcal{J}(\psi(u), \nabla\psi(u)) = \nabla\psi(u)$. So Noether's theorem gives us again that $\triangle\psi = \operatorname{div}(\nabla\psi) = 0$. As a similar case, we could consider a Lorentzian version of the above Lagrangian, say, with $-v_m^2$ instead of $+v_m^2$, and conclude that functions ψ which are critical points of the associated action functional are solutions of the wave equation: $\Box\psi = 0$, where

$$\Box = \frac{\partial^2}{(\partial u^1)^2} + \cdots + \frac{\partial^2}{(\partial u^{m-1})^2} - \frac{\partial^2}{(\partial u^m)^2}$$

is the d'Alembertian operator.

We have seen previously that the single-variable version of Noether's theorem for $\mathbb{R}$-actions was in fact equivalent to a version with general G-actions, where G is any Lie group. There is no reason why this should be different here. Indeed, if G is a Lie group acting on Q, then we have a diagonal derivative action $G \circlearrowright TQ^{\oplus m}$, and so it makes sense to say when a Lagrangian on m parameters $L\colon TQ^{\oplus m} \to \mathbb{R}$ is G-invariant. The following theorem requires no proof:

Theorem 40 (Noether). *Let $L\colon TQ^{\oplus m} \to \mathbb{R}$ be a G-invariant Lagrangian. Then, for every element $X \in \mathfrak{g}$, the Noether current $\mathcal{J}^X = ((\mathcal{J}^X)^1, \ldots, (\mathcal{J}^X)^m)\colon TQ^{\oplus m} \to \mathbb{R}^m$ defined by*

$$(\mathcal{J}^X)^\ell(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = (\mathbb{F}L)^\ell(x, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)\boldsymbol{X}^\#_x,$$

has zero divergence along L-critical m-surfaces $x\colon \Omega \to Q$. In other words, the pull-back map $\Omega \ni u \mapsto \mathcal{J}^X(x(u), \nabla x(u)) \in \mathbb{R}^m$ is a vector field along Ω, and $\operatorname{div}\mathcal{J}^X(x(u), \nabla x(u)) = 0$.

References

[1] Bryant, R. L.; *An Introduction to Lie Groups and Symplectic Geometry*, Park City 1991, Geometry and quantum field theory, 2018. Currently available in the author's webpage.

[2] Godinho, L.; Natário, J.; *An Introduction to Riemannian Geometry With Applications to Mechanics and Relativity*, Springer-Verlag (Universitext), 2014.

[3] McDuff, D.; Salamon, D.; *Introduction to Symplectic Topology*, 3rd edition, Oxford University Press, 2017.

[4] Talpaert, Y.; *Differential Geometry: with Applications to Mechanics and Physics*, Pure and Applied Mathematics, Marcel Dekker, Inc., 2001.

[5] Warner, G.; *Lagrangian Mechanics*, University of Washington. Currently available in the author's webpage.