MECHANICS MATH 249A, WINTER 2012

Updated version of Reuben Brasher's LaTeXed notes from the 2010 offering of Math 249A

1. January 9–11: Overview and examples. The harmonic oscillator and the spherical pendulum

The course will focus on geometric mechanics and conservative systems, emphasizing mathematical models of (physical) systems with geometric structure that can be used to simplify analysis of the system. Examples: variational problems, Lagrangian systems, Hamiltonian systems (symplectic and Poisson). We typically won't be going for complete solutions—usually will seek only partial information. E.g., are the solutions constrained to lie on level sets of known functions? Are there any geometric invariants under the flow? Any symmetries?

These conservative, highly structured models are usually extreme idealizations. One strategy for analyzing more realistic models is to first analyze the idealized system, using all available machinery, then analyze the more complex model as a perturbation of the "understood" model.

The following examples will be treated bare hands, in detail, but demonstrate most of the key features of the classes of systems we'll be studying. Unmotivated observations for these examples will hopefully be completely predictable and obvious by the end of the course.

Example 1 (Harmonic oscillator). Consider block of mass m attached to a wall by a spring, sliding on a frictionless table. Assume the spring has a rest length l and is a linear spring, that is response is given by Hooke's law. Force $F(x) = k(\ell - x)$ for some $k > 0$ spring constant. Together Hooke's law and Newton's law give

$$m\,\ddot{x} = k(\ell - x).$$

Set $y = x - \ell$, displacement from rest length. Then $\ddot{y} = \ddot{x}$, so

$$m\,\ddot{y} = -k\,y.$$

If we know about trig functions, then

$$\frac{d^2}{dt^2}\cos(\omega\,t) = -\omega^2\cos(\omega\,t) \qquad \text{and} \qquad \frac{d^2}{dt^2}\sin(\omega\,t) = -\omega^2\sin(\omega\,t)$$

imply that the general solution has the form

$$y(t) = a\,\cos\omega\,t + b\,\sin\omega\,t,$$

where $\omega = \sqrt{\frac{k}{m}}$. Given initial values $y(0) = y_0$ and $\dot{y}(0) = v_0$, the particular solution is

$$y(t) = y_0\cos\omega\,t + \frac{m}{k}v_0\sin\omega\,t.$$

Pretend we don't know about trig functions, then what? Rewrite as a first order system:

$$\begin{aligned}\dot{y} &= v\\ \dot{v} &= -\frac{k}{m}y\end{aligned} \tag{1}$$

Date: February 6, 2012.

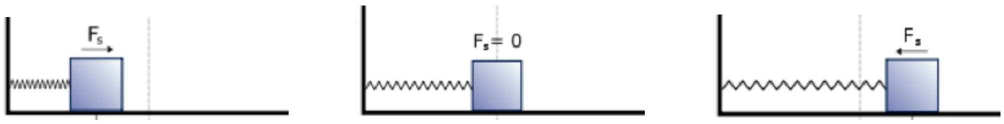


FIGURE 1. Spring–mass system: compressed ($y < 0$), at rest ($y = 0$), extended ($y > 0$) states. (Figure taken from Wikipedia article on the harmonic oscillator.)

Observation: If $(y(t), v(t))$ is a solution of (1) then

$$\frac{d}{dt}\left(\frac{m}{2}v^2 + \frac{k}{2}y^2\right) = m\,v\,\dot{v} + k\,y\,\dot{y} = (m\,\dot{v} + k\,y)v = 0.$$

Hence solutions lie on level curves of $C(y,v) := \frac{1}{2}(m\,v^2 + k\,y^2)$. Level curves of C are ellipses centered at the origin.

Given initial data (y_0, v_0), we can solve $C(y,v) = C(y_0, v_0)$ to express the speed as a function of y and the initial data

$$|\dot{y}| = \sqrt{v_0^2 + \frac{k}{m}\left(y_0^2 - y^2\right)}.$$

To compute the period of a given oscillation of the mass, we can compute the time elapsed moving from the most compressed state, $-y_{\max}$, to the maximally extended state $y_{\max}$, for that trajectory, then double that time. Since the mass is transitioning from moving left to right (or vice versa) when maximally compressed or extended, the velocity at those states will be zero, the initial data $y_0 = -y_{\max}$, $v_0 = 0$ will determine an orbit with the desired trace. If y is increasing, then

$$\frac{dy}{dt} = \sqrt{v_0^2 + \frac{k}{m}\left(y_0^2 - y^2\right)} = \omega\sqrt{y_{\max}^2 - y^2}.$$

Solve by separation of variables, time elapsed moving from $-y_{\max}$ to $+y_{\max}$,

$$t = \frac{1}{\omega}\int_{-y_{\max}}^{y_{\max}} \frac{dy}{\sqrt{y_{\max}^2 - y^2}} = \frac{\pi}{\omega}.$$

Hence the period is $\frac{2\pi}{\omega}$. Note that the period is the same for all orbits; this property of the harmonic oscillator is very special.

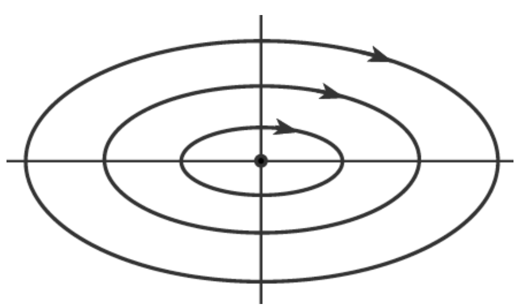


FIGURE 2. Sample phase portrait for the spring–mass system. The level curves of $C(y,v) = \frac{1}{2}(m\,v^2 + k\,y^2)$ are ellipses; the arrows indicate the direction of motion along level curves.

Example 2 (Spherical Pendulum). Point mass on sphere of radius ℓ in $\mathbb{R}^3$ acted on by gravity. Let $\mathbf{x} \in \mathbb{R}^3$ denote the position of the point mass. $\ell^2 = \|\mathbf{x}\|^2$ implies $0 = 2\,\mathbf{x}\cdot\dot{\mathbf{x}}$ implies that $\dot{\mathbf{x}}$ is tangent to the sphere and

$$0 = \|\dot{\mathbf{x}}\|^2 + \mathbf{x}\cdot\ddot{\mathbf{x}}$$

that is radial component of determined by the velocity.

Newton's law for an unconstrained particle influenced by gravity is $m\,\ddot{\mathbf{x}} = -g\,m\,\mathbf{e}_3$ where $\mathbf{e}_3$ denotes the upward unit vertical vector, $m > 0$ is the mass of the particle, and g is the strength of gravity. The constraint $\|\mathbf{x}\| \equiv \ell$ determines the radial component of the acceleration; the tangential component satisfies

$$\mathbb{P}_{\mathbf{x}}(\ddot{\mathbf{x}} + g\,\mathbf{e}_3) = 0, \tag{2}$$

where $\mathbb{P}_{\mathbf{x}}$ denotes orthogonal projection onto the plane perpendicular to $\mathbf{x}$ (identified with tangent plane $T_{\mathbf{x}}S^2_\ell$), given by

$$\mathbb{P}_{\mathbf{x}}\mathbf{w} = \mathbf{w} - \frac{\mathbf{w}\cdot\mathbf{x}}{\|\mathbf{x}\|^2}\mathbf{x}. \tag{3}$$

Hence

$$\begin{aligned} 0 &= \ddot{\mathbf{x}} + g\,\mathbf{e}_3 - \frac{1}{\ell^2}(\mathbf{x}\cdot\ddot{\mathbf{x}} + g\,\mathbf{e}_3\cdot\mathbf{x})\mathbf{x} \\ &= \ddot{\mathbf{x}} + g\,\mathbf{e}_3 + \frac{1}{\ell^2}(\|\dot{\mathbf{x}}\|^2 - g\,\mathbf{e}_3\cdot\mathbf{x})\mathbf{x} \end{aligned} \tag{4}$$

Claim: If $\mathbf{x}(t)$ is a solution of (2), then

$$\frac{d}{dt}\left(\tfrac{1}{2}\,\|\dot{\mathbf{x}}\|^2 + g\,\mathbf{e}_3\cdot\mathbf{x}\right) = 0.$$

Hence $\|\dot{\mathbf{x}}\|^2$ can be expressed in terms of $\mathbf{x}\cdot\mathbf{e}_3$ and the initial data.

Proof:

$$\frac{d}{dt}\left(\tfrac{1}{2}\,\|\dot{\mathbf{x}}\|^2 + g\,\mathbf{e}_3\cdot\mathbf{x}\right) = \dot{\mathbf{x}}\cdot\ddot{\mathbf{x}} + g\,\mathbf{e}_3\cdot\dot{\mathbf{x}} = \dot{\mathbf{x}}\cdot(\ddot{\mathbf{x}} + g\,\mathbf{e}_3).$$

We know that the radial component of $\dot{\mathbf{x}}$ is 0, since $0 = \mathbf{x}\cdot\dot{\mathbf{x}}$. Equation (2) implies that $\ddot{\mathbf{x}} + g\,\mathbf{e}_3$ is radial (that is, a rescaling of $\mathbf{x}$). Hence the dot product of $\dot{\mathbf{x}}$ and $\ddot{\mathbf{x}} + g\,\mathbf{e}_3$ is zero. □

Let $\mathbf{x}(t)$ be a solution of (2) and let $z(t) := \frac{1}{\ell}\mathbf{x}(t)\cdot\mathbf{e}_3$. Taking the inner product of (4) with $\mathbf{e}_3$ yields

$$0 = \ddot{\mathbf{x}}\cdot\mathbf{e}_3 + g\,\mathbf{e}_3\cdot\mathbf{e}_3 + \frac{1}{\ell^2}\left(\|\dot{\mathbf{x}}\|^2 - g\,\mathbf{e}_3\cdot\mathbf{x}\right)\mathbf{x}\cdot\mathbf{e}_3 = \ddot{z} + g + f(z)z,$$

where $f(z) := \frac{1}{\ell^2}\left(\|\mathbf{v}_0\|^2 + g(2\,\mathbf{e}_3\cdot\mathbf{x}_0 - 3\,z)\right)$. Thus, if $\mathbf{x}(t)$ is a solution of (2), then $z(t)$ satisfies the second order, one degree of freedom ODE

$$\ddot{z} + f(z)z + g = 0. \tag{5}$$

On the other hand, a solution $z(t)$ of (5) describes a solution of the full spherical pendulum problem modulo rotation about the vertical axis. (This is a special case of *singular reduction* with respect to a symmetry group of a conservative system.) Define

$$H(z, v) := \tfrac{1}{2}v^2 + \int^z (f(u)u + g)du.$$

If $z(t)$ is a solution of (5), then $H(z(t), \dot{z}(t))$ is constant with respect to t. If $\dot{z} \neq 0$ for $t_0 \leq t \leq t_1$, we can find an implicit algebraic equation relating t and z by solving $H(z, \dot{z}) = h_0 := H(z(t_0), \dot{z}(t_0))$ for $\dot{z}$, then solving the resulting separable first order ODE $\dot{z} = X(z)$. (See previous example.)

How to get the "angle" information, given $z(t)$? Given a solution $z(t)$ of (5) satisfying given initial conditions

$$\begin{aligned} z(0) &= \mathbf{x}_0\cdot\mathbf{e}_3, \\ \dot{z}(0) &= \mathbf{v}_0\cdot\mathbf{e}_3, \end{aligned}$$

let $\mathbf{s}(t) = \left(\sqrt{1 - z(t)^2}, 0, z(t)\right) \in \mathbb{S}^2 \cap x\text{–}z$ plane. (The intersection of the unit sphere and the x–z plane is an example of a "slice" for the circle action on the sphere given by rotations about $\mathbf{e}_3$).

Let R_θ denote the rotation matrix

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

Multiplication by R_θ implements rotation about $\mathbf{e}_3$ through the angle θ. Seek curve $\theta(t)$ such that $\mathbf{x}(t) = \ell\, R_{\theta(t)} s(t)$ satisfies the spherical pendulum equation.

Observe that if $\mathbf{x}(t)$ is a solution of the pendulum equation, then

$$\begin{aligned} \frac{d}{dt}(\mathbf{x} \times \dot{\mathbf{x}}) &= \dot{\mathbf{x}} \times \dot{\mathbf{x}} + \mathbf{x} \times \ddot{\mathbf{x}} \\ &= \mathbf{x} \times \mathbb{P}_{\mathbf{x}} \ddot{\mathbf{x}} \\ &= \mathbf{x} \times (-\mathbb{P}_{\mathbf{x}} g\, \mathbf{e}_3) \\ &= g\, \mathbf{e}_3 \times \mathbf{x}. \end{aligned}$$

In particular, $(\mathbf{x} \times \dot{\mathbf{x}}) \cdot \mathbf{e}_3$ is constant along solutions, i.e. the vertical component of the angular momentum is a constant of the motion.

Now plug in $\mathbf{x} = R_\theta \mathbf{s}$.

$$\begin{aligned} \dot{\mathbf{x}} &= \left(\frac{d}{dt} R_\theta\right) \mathbf{s} + R_\theta \dot{\mathbf{s}} = R_\theta(\dot\theta\, \mathbf{e}_3 \times \mathbf{s}) + R_\theta \dot{\mathbf{s}} \\ &= R_\theta(\dot\theta\, \mathbf{e}_3 \times \mathbf{s} + \dot{\mathbf{s}}). \end{aligned}$$

(Here we've used the fact that $\mathbf{y}(t) = R_t \mathbf{y}_0$ is the solution of the linear IVP $\dot{\mathbf{y}} = \mathbf{e}_3 \times \mathbf{y}$, $\mathbf{y}(0) = \mathbf{y}_0$. See upcoming notes on matrix groups, matrix exponentiation, and linear systems of ODEs if this material is unfamiliar or has gotten a little rusty.) Inserting the expression for $\dot{\mathbf{x}}$ into the vertical component of the angular momentum gives

$$\begin{aligned} (\mathbf{x} \times \dot{\mathbf{x}}) \cdot \mathbf{e}_3 &= \big(R_\theta(s) \times R_\theta(\dot\theta \mathbf{e}_3 \times \mathbf{s} + \dot{\mathbf{s}})\big) \cdot \mathbf{e}_3 \\ &= (\mathbf{s} \times (\dot\theta \mathbf{e}_3 \times \mathbf{s} + \dot{\mathbf{s}})) \cdot \mathbf{e}_3, \end{aligned}$$

since $R_\theta \mathbf{x} \cdot \mathbf{e}_3 = \mathbf{x} \cdot R_{-\theta} \mathbf{e}_3 = \mathbf{x} \cdot \mathbf{e}_3$ for any $\theta \in \mathbb{R}$ and and $\mathbf{x} \in \mathbb{R}^3$. Since $\mathbf{s}$ is in the x–z plane for all t, $\mathbf{s} \times \dot{\mathbf{s}}$ is a scalar multiple of $\mathbf{e}_2$, and hence perpendicular to $\mathbf{e}_3$, it follows that

$$(\mathbf{x} \times \dot{\mathbf{x}}) \cdot \mathbf{e}_3 = \dot\theta(\mathbf{s} \times (\mathbf{e}_3 \times \mathbf{s})) \cdot \mathbf{e}_3 = \dot\theta \left\| \mathbf{e}_3 \times \mathbf{s} \right\|^2.$$

Hence $\dot\theta \left\| \mathbf{e}_3 \times \mathbf{s} \right\|^2 \equiv \mu$ for some constant μ if $R_\theta \mathbf{s}$ is a solution of the pendulum equations.

Case 1: If $\mu \neq 0$, then $\mathbf{s}$ is never vertical and

$$\dot\theta = \frac{\mu}{\left\| \mathbf{e}_3 \times \mathbf{s} \right\|^2} = \frac{\mu}{\ell^2 - z^2}.$$

Hence, since we already know $z(t)$,

$$\theta(t) - \theta_0 = \mu \int_0^t \frac{du}{\ell^2 - z(u)^2}.$$

Case 2: $\mu = 0$. If $\mathbf{s}$ is ever vertical, then the vertical component of the angular momentum doesn't determine $\dot\theta$—the circle group action fixes $\pm\mathbf{e}_3$. (This can only occur if $\mu = 0$. Pop quiz: must this occur if $\mu = 0$?) Go back to the expression for the vertical component of the angular momentum in terms of $\mathbf{x} = (x, y, z)$, namely $0 = (\mathbf{x} \times \dot{\mathbf{x}}) \cdot \mathbf{e}_3 = x\dot y - \dot x y$. This ODE is separable:

$$\frac{\dot x}{x} = \frac{\dot y}{y}, \qquad \text{with solution} \quad \ln x = \ln y + C,$$

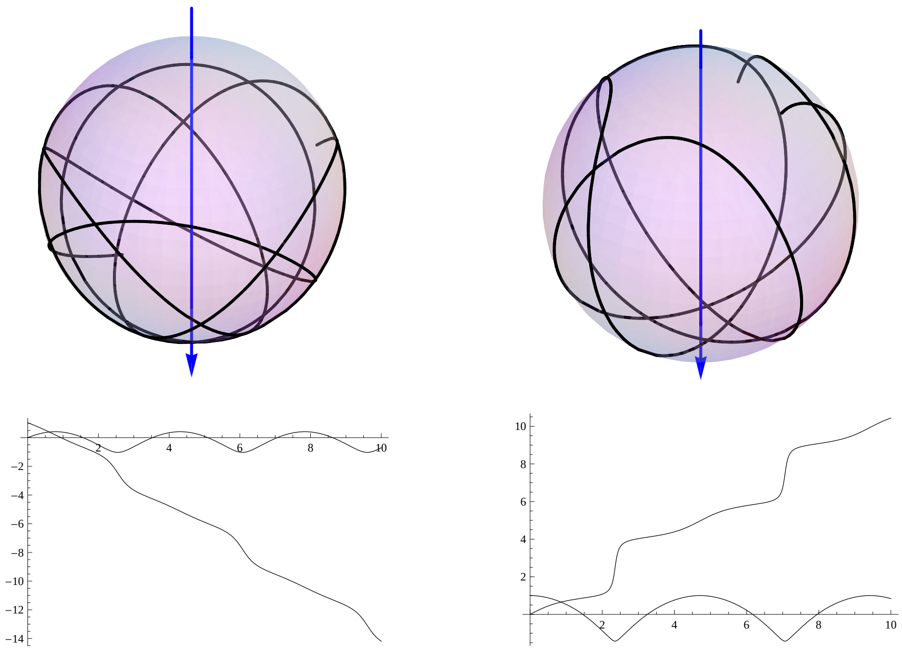


FIGURE 3. Two trajectories of the spherical pendulum. Top: traces of trajectories. Bottom: graphs of the spherical coordinates of the solutions, as functions of time.

so x and y are proportional. Thus there exists θ_0 such that $\mathbf{x}(t) = R_{\theta_0}\mathbf{s}(t)$ for all t. That is, if $\mu = 0$, the solution remains in a fixed vertical plane.

Exercise 1. Compute the full evolution of equations in spherical coordinates, that is plug $\mathbf{x}(\theta, \phi) = (\cos\theta\,\cos\phi, \cos\theta\,\sin\phi, \sin\theta)$, where θ and ϕ are functions of t, into $\mathbb{P}_{\mathbf{x}}(\ddot{\mathbf{x}} + g\,\mathbf{e}_3) = 0$ and derive the evolution equations for θ and ϕ. (Use of a symbolic computation language, e.g. *Mathematica* or SAGE, is strongly recommended—lots of trig identities need to be invoked to get something tolerable.)

Exercise 2. Determine the equations of motion and conservation laws for a point mass in $\mathbb{R}^3$ connected to the origin by a spring with rest length ℓ moving under the influence of gravity. Assume a linear spring (Hooke's Law) with spring constant k; the spring cannot bend, i.e. it must always be radial. Let m denote the mass and g denote the strength of gravity.

2. January 11–18: Calculus of Variations

Consider a manifold Q, two points q_0, q_1 in Q, and a closed interval $I = [t_0, t_1]$ in $\mathbb{R}$. Let

$$\mathcal{C}(I, q_0, q_1) = \{\gamma \in \mathcal{C}(I, Q) : \gamma(q_0) = q_0 \text{ and } \gamma(q_1) = q_1\}.$$

Very formal/nonrigorous assertion: $\mathcal{C}(I, q_0, q_1)$ is an infinite-dimensional manifold. What are the tangent vectors?

Consider a smooth curve γ_ε where $|\varepsilon| < \varepsilon_0$ in $\mathcal{C}(I, q_0, q_1)$. If $t_0 < t < t_1$, then $\gamma_\varepsilon(t)$ is a curve parameterized by ε. Define

$$\delta\gamma(t) := \frac{d}{d\varepsilon}\gamma_\varepsilon(t)\bigg|_{\varepsilon=0} \in T_{\gamma_0(t)}Q.$$

Since the endpoints are fixed, $\delta\gamma(t_j) = 0$, $j = 0, 1$. $\delta\gamma$ is a vector field over γ_0, that is, $\delta\gamma$ assigns to each point in I a tangent vector to Q with $\delta\dot{\gamma}(t) \in T_{\gamma_0(t)}Q$. Since the endpoints are fixed, $\delta\gamma(t_j) = 0$, $j = 0, 1$.

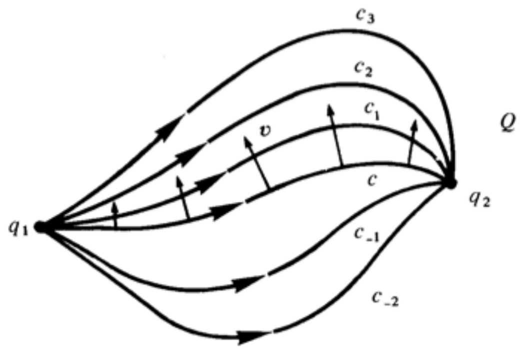


FIGURE 4. A family of parameterized curves on a manifold Q. (From *Foundations of Mechanics*.)

On the other hand, if V is a smooth vector field on a neighborhood U in Q containing $\gamma(I)$, the trace of γ, for some curve $\gamma \in \mathcal{C}(I, q_0, q_1)$, with $V(\gamma(t_j)) = 0$, $j = 0, 1$, then there is a curve $\gamma_\varepsilon \in \mathcal{C}(I, q_0, q_1)$, such that $\gamma = \gamma_0$ and $\delta\gamma = V \circ \gamma$. Sketch of construction: Let $\mathcal{F}_\varepsilon$ be the flow of V. Since for each $t_* \in I$, $\mathcal{F}_\varepsilon(t_*)$ is defined for $|\varepsilon| < \varepsilon_0(t_*)$ for some $\mathcal{E} : I \to \mathbb{R}^+$ and I is compact, so there is ε_0 such that $\mathcal{F}_\varepsilon(t)$ is defined for all $|\varepsilon| < \varepsilon_0$. Set $\gamma_\varepsilon = \mathcal{F}_\varepsilon \circ \gamma$. Hence

$$\frac{d}{d\varepsilon}\gamma_\varepsilon = \frac{d}{d\varepsilon}\mathcal{F}_\varepsilon \circ \gamma = V \circ \mathcal{F}_\varepsilon \circ \gamma = V \circ \gamma_\varepsilon.$$

It follows that $T_\gamma\mathcal{C}(I, q_0, q_1)$ is, roughly speaking, the set of smooth vector fields over γ equalling 0 at the end points.

Now consider a smooth function $L : TQ \to \mathbb{R}$ and define $\mathcal{L} : \mathcal{C}(I, q_0, q_1) \to \mathbb{R}$ by

$$\mathcal{L}(\gamma) = \int_I L(\dot\gamma(t))\, dt.$$

Seek curve $\gamma \in \mathcal{C}(I, q_0, q_1)$ such that γ is a critical point of $\mathcal{L}$. (Historical focus was on curves minimizing $\mathcal{L}$.) γ is a critical point of $\mathcal{L}$ if and only if $\frac{d}{dt}\mathcal{L} \circ \gamma_\varepsilon\big|_{\varepsilon=0} = 0$ for all smooth $\gamma_\varepsilon \in \mathcal{C}(I, q_0, q_1)$ satisfying $\gamma_0 = \gamma$.

$$\begin{aligned}
0 = \frac{d}{d\varepsilon}\mathcal{L}(\gamma_\varepsilon)\bigg|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\int_I L(\dot\gamma_\varepsilon(t))\, dt\bigg|_{\varepsilon=0} \\
&= \int_I \frac{\partial}{\partial\varepsilon} L(\dot\gamma_\varepsilon(t))\bigg|_{\varepsilon=0} dt \\
&= \int_I dL(\dot\gamma_0(t))\, \frac{\partial}{\partial\varepsilon}\dot\gamma_\varepsilon(t)\bigg|_{\varepsilon=0} dt \\
&= \int_I dL(\dot\gamma_0(t))(\delta\dot\gamma_0(t))\, dt,
\end{aligned}$$

since mixed partials commute.

$\delta\dot\gamma(t) \in TTQ$—it involves position, velocity, variation of position, and variation of velocity information. Focus on two special cases— vector spaces and matrix groups—where these pieces of information can easily be teased apart to further simplify the expression $\frac{d}{d\varepsilon}\mathcal{L}(\gamma_\varepsilon)\big|_{\varepsilon=0}$.

Special case 1: $Q = X$, a vector space. Identify TX with X^2, so $L : X^2 \to \mathbb{R}$. Define $\frac{\partial L}{\partial q}, \frac{\partial L}{\partial v} : X^2 \to X^*$ by

$$\frac{\partial L}{\partial q}(q, v) \cdot u + \frac{\partial L}{\partial v}(q, v) \cdot w = dL(q, v)\,((q, u), (v, w)) \tag{6}$$

for $q, v, u, w \in X$. (Here $\cdot$ denotes the natural pairing between X and X^*.)

In a slight abuse of notation, let $\frac{d}{dt}\gamma(t) = (\gamma(t), \dot\gamma(t))$. Then

$$\frac{\partial}{\partial\varepsilon}\frac{\partial}{\partial t}\gamma_\varepsilon(t)\bigg|_{\varepsilon=0} = \frac{\partial}{\partial\varepsilon}\left(\gamma_\varepsilon(t), \dot\gamma_\varepsilon(t)\right)\bigg|_{\varepsilon=0}$$

and

$$\begin{aligned}\frac{\partial}{\partial\varepsilon}\mathcal{L}(\gamma_\varepsilon)\bigg|_{\varepsilon=0} &= \int_I dL(\gamma(t), \dot\gamma(t))(\gamma(t), \delta\gamma(t), \dot\gamma(t), \delta\dot\gamma(t))\, dt \\ &= \int_I \left(\frac{\partial L}{\partial q}(\gamma, \dot\gamma)\cdot\delta\gamma + \frac{\partial L}{\partial v}(\gamma, \dot\gamma)\cdot\delta\dot\gamma\right) dt.\end{aligned}$$

Using the product rule and the fundamental theorem of calculus,

$$\begin{aligned}\int_I \left(\frac{d}{dt}\left(\frac{\partial L}{\partial v}(\gamma, \dot\gamma)\right)\cdot\delta\gamma + \frac{\partial L}{\partial v}(\gamma, \dot\gamma)\cdot\delta\dot\gamma\right) dt &= \int_I \frac{d}{dt}\left(\frac{\partial L}{\partial v}(\gamma, \dot\gamma)\cdot\delta\gamma\right) dt \\ &= \left(\frac{\partial L}{\partial v}(\gamma, \dot\gamma)\cdot\delta\gamma\right)\bigg|_{t_0}^{t_1} \\ &= 0,\end{aligned}$$

since $\delta\gamma(t_j) = 0$, $j = 0, 1$. Hence

$$\frac{d}{d\varepsilon}\mathcal{L}(\gamma_\varepsilon)\bigg|_{\varepsilon=0} = \int_I \left(\frac{\partial L}{\partial q}(\gamma, \dot\gamma) - \frac{d}{dt}\left(\frac{\partial L}{\partial v}(\gamma, \dot\gamma)\right)\right)\cdot\delta\gamma\, dt. \tag{7}$$

If γ is a smooth curve such that (7) equals zero for any smooth $\delta\gamma$ satisfying $\delta\gamma(t_0) = \delta\gamma(t_1) = 0$, then

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}(\gamma, \dot\gamma)\right) = \frac{\partial L}{\partial q}(\gamma, \dot\gamma) \tag{8}$$

on I. (8) is called the **Euler-Lagrange equation** determined by the **Lagrangian** L. On more general function spaces, a curve γ such that (7) equals zero for any smooth $\delta\gamma$ is called a *weak solution* of the Euler-Lagrange equations, and need not satisfy (8) pointwise. Curves satisfying (8) at all points on I are called *strong solutions* of the Euler-Lagrange equations.

Example 3. Let X be an inner-product space, $V : X \to \mathbb{R}$ a smooth function, and $m \in \mathbb{R}^+$. If

$$L(q, v) = \frac{m}{2}\|v\|^2 - V(q),$$

then

$$\frac{\partial L}{\partial q}(q, v) = -dV(q) \qquad \text{and} \qquad \frac{\partial L}{\partial v}(q, v) = mv^\flat$$

where $v^\flat \in X^*$ is the element of the dual space X^* satisfying $v^\flat \cdot w = \langle v, w\rangle$ for all $w \in X$. Why?

$$\begin{aligned}\frac{d}{d\varepsilon}L(q, v + \varepsilon w)\bigg|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\left(\frac{m}{2}\|v + \varepsilon w\|^2 - V(q)\right)\bigg|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}\left(\frac{m}{2}\langle v + \varepsilon w,\, v + \varepsilon w\rangle\right)\bigg|_{\varepsilon=0} \\ &= \frac{m}{2}\,\frac{d}{d\varepsilon}\left(\|v\|^2 + 2\varepsilon\langle v,\, w\rangle + \varepsilon^2\|w\|^2\right)\bigg|_{\varepsilon=0} \\ &= m\langle v,\, w\rangle.\end{aligned}$$

Thus the Euler-Lagrange equation is

$$\frac{d}{dt}\left(m\, v^\flat\right) = -dV(\gamma).$$

If we let $\nabla V : X \to X$ denote the map satisfying

$$\langle\nabla V(q),\, w\rangle = dV(q)\cdot w$$

for $q, w \in X$, then the Euler-Lagrange is equivalent to $m\,\ddot{\gamma} = -\nabla V(\gamma)$.

Note: to have a well defined second order ODE on X, we need to have a non-degenerate dependence of $\frac{d}{dt}\left(\frac{\partial L}{\partial v}(\gamma, \dot{\gamma})\right)$ on $\ddot{\gamma}$, ie. we need a non-degenerate $\partial^2 L/\partial v^2$. A Lagrangian L satisfying this nondegeneracy condition is called a **regular Lagrangian**.

3. January 18 and 23, the calculus of variations on matrix groups

What's special about vector spaces, that we started with that case? The tangent bundle of a vector space is trivial: $TX \equiv X \times X$, so we can define the partial derivatives (6) of a smooth function L on TX and derive the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial v}(\gamma, \dot{\gamma}) = \frac{\partial L}{\partial q}(\gamma, \dot{\gamma})$$

as the condition of criticality for the functional $\mathcal{L}(\gamma) = \int_0^T L(\gamma, \dot{\gamma})\,dt$.

Given a smooth function on an arbitrary tangent bundle TQ, we can define the partial with respect to the velocity, $\frac{\partial L}{\partial v}(v_q) : T_qQ \to T_q^*Q$, at $v_q \in TQ$ by

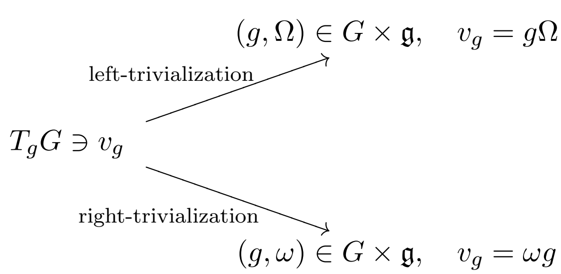
$$\frac{\partial L}{\partial v}(v_q) \cdot w_q := \left.\frac{d}{d\varepsilon} L(v_q + \varepsilon\, w_q)\right|_{\varepsilon=0} \qquad \text{for all } w_Q \in T_qQ,$$

but the partial with respect to position, $\frac{\partial L}{\partial q}(v_q)$, cannot be defined without a trivialization or connection. If we want to 'freeze' the velocity while varying the position, we need some way of propagating the frozen value from fiber to fiber in the tangent bundle.

A Lie group is the next best thing after a vector space in this regard. The spaces of left and right invariant vector fields on Lie groups (see Lie groups mini-tutorial) provide natural and convenient trivializations of the tangent bundles (and induce corresponding trivializations of the cotangent bundle etc.). The maps $\tau_\ell, \tau_r : TG \to G \times \mathfrak{g}$ given by

$$\tau_\ell(v_g) := (g, {L_{g^{-1}}}^* v_g) \qquad \text{and} \qquad \tau_r(v_g) := (g, {R_{g^{-1}}}^* v_g)$$

are global trivializations of the tangent bundle of G, with inverses $\tau_\ell^{-1}(g, \xi) = {L_g}^* \xi$ etc.; for a matrix group, $\tau_\ell(V_A) = (A, A^{-1}V_A)$ and $\tau_\ell^{-1}(A, \xi) = A\,\xi$.



Given a smooth function $L : TG \to \mathbb{R}$ on the tangent bundle of a matrix group, define the left-trivialized partial derivatives $\frac{\delta L}{\delta g}, \frac{\delta L}{\delta \Omega} : G \times \mathfrak{g} \to \mathfrak{g}^*$ by

$$\frac{\delta L}{\delta g}(g, \Omega) \cdot \eta + \frac{\delta L}{\delta \Omega}(q, \Omega) \cdot \xi = dL(g\,\Omega) \cdot (g\,\eta\,\Omega + g\,\xi)$$

for all $g \in G$ and $\Omega, \eta, \xi \in \mathfrak{g}$. (The right trivialized partial derivatives are defined analogously.) If the Lagrangian is explicitly given as a function on $G \times \mathfrak{g}$, then

$$\frac{\delta L}{\delta g}(g, \Omega) \cdot \eta + \frac{\delta L}{\delta \Omega}(g, \Omega) \cdot \xi = dL(g, \Omega) \cdot (g\,\eta, (\Omega, \xi)).$$

(Here $T(G \times \mathfrak{g}) \sim TG \times (\mathfrak{g} \times \mathfrak{g})$.) The definitions for an arbitrary Lie group are analogous, but involve the tangent maps of left or right multiplication.

In the vector space case, commutation of partial derivatives, $\frac{\partial}{\partial \varepsilon}\frac{\partial}{\partial t} = \frac{\partial}{\partial t}\frac{\partial}{\partial \varepsilon}$, played a crucial role in the derivation of the Euler-Lagrange equations. In the Lie group setting, how does the trivialization of the tangent bundle interact with the exchange of order of differentiation? For notational simplicity, consider curves in a matrix group G.

Given a smooth curve $\gamma_\varepsilon \in \mathcal{C}(I, g_0, g_1)$, let $\Omega_\varepsilon : I \to \mathfrak{g}$ be the parametrized family of maps such that

$$\frac{\partial}{\partial t}\gamma_\varepsilon(t) = \gamma_\varepsilon(t)\Omega_\varepsilon(t),$$

that is, $(\gamma_\varepsilon, \Omega_\varepsilon) = \tau_\ell(\dot\gamma_e)$ is the left trivialization of γ_ε, and let $\eta : I \to \mathfrak{g}$ denote the algebra component of the left trivialization of $\frac{d}{d\varepsilon}\gamma_\varepsilon(t)\big|_{\varepsilon=0}$. If we let $'$ denote the evaluation of $\frac{\partial}{\partial\varepsilon}$ at $\epsilon = 0$, and let $\gamma = \gamma_0$, then $\tau_\ell(\gamma') = (\gamma, \eta)$.

$$\begin{aligned}
\frac{\partial}{\partial\varepsilon}\frac{\partial}{\partial t}\gamma_\varepsilon(t)\bigg|_{\varepsilon=0} &= \frac{\partial}{\partial\varepsilon}\left(\gamma_\varepsilon(t)\Omega_\varepsilon(t)\right)\bigg|_{\varepsilon=0} \\
&= \left(\frac{\partial}{\partial\varepsilon}\gamma_\varepsilon(t)\bigg|_{\varepsilon=0}\right)\Omega_0(t) + \gamma_0(t)\left(\frac{\partial}{\partial\varepsilon}\Omega_\varepsilon(t)\bigg|_{\varepsilon=0}\right) \\
&= \left(\gamma'\,\Omega + \gamma\,\Omega'\right)(t) \\
&= \left(\gamma\,\eta\,\Omega + \gamma\,\Omega'\right)(t),
\end{aligned}$$

where $\Omega = \Omega_0$. On the other hand

$$\frac{\partial}{\partial t}\frac{\partial}{\partial\varepsilon}\gamma_\varepsilon(t)\bigg|_{\varepsilon=0} = \frac{d}{dt}(\gamma(t), \eta(t)) = \left(\dot\gamma\,\eta + \gamma\,\dot\eta\right)(t) = \left(\gamma\,\Omega\,\eta + \gamma\,\dot\eta\right)(t).$$

Hence, by equality of mixed partials

$$\gamma\left(\eta\,\Omega + \Omega'\right) = \gamma(\Omega\,\eta + \dot\eta).$$

Solving for Ω' (γ is invertible) yields

$$\Omega' = \dot\eta + \Omega\,\eta - \eta\,\Omega = \left(\frac{d}{dt} + \mathrm{ad}_\Omega\right)\eta,$$

where $\mathrm{ad}_\Omega : \mathfrak{g} \to \mathfrak{g}$ is given by $\mathrm{ad}_\Omega\,\eta := [\Omega, \eta] = \Omega\,\eta - \eta\,\Omega$.

Now consider $L : G \times \mathfrak{g} \to \mathbb{R}$ and $\mathcal{L}(\gamma) := \int_I L(\tau_\ell(\dot\gamma))\,dt$.

$$\begin{aligned}
\frac{d}{d\varepsilon}\mathcal{L}(\gamma_\varepsilon)\bigg|_{\varepsilon=0} &= \frac{d}{d\varepsilon}\int_I L(\gamma_\varepsilon, \Omega_\varepsilon)\,dt\bigg|_{\varepsilon=0} \\
&= \int_I dL(\gamma, \Omega)\cdot(\gamma\,\eta, (\Omega, \Omega'))\,dt \\
&= \int_I \left(\frac{\delta L}{\delta g}(\gamma, \Omega)\cdot\eta + \frac{\delta L}{\delta\Omega}(\gamma, \Omega)\cdot\Omega'\right)dt \\
&= \int_I \left(\frac{\delta L}{\delta g}(\gamma, \Omega)\cdot\eta + \frac{\delta L}{\delta\Omega}(\gamma, \Omega)\cdot(\dot\eta + \mathrm{ad}_\Omega\,\eta)\right)dt.
\end{aligned}$$

Integrate the term involving $\dot\eta$ by parts to get

$$\int \frac{\delta L}{\delta\Omega}(\gamma, \Omega)\cdot\dot\eta\,dt = -\int \frac{d}{dt}\left(\frac{\delta L}{\delta\Omega}(\gamma, \Omega)\right)\cdot\eta\,dt;$$

the boundary terms are zero, since $\gamma_\varepsilon(t_j) \equiv g_j$ implies $\eta(t_j) = 0$. We can 'flip ad_Ω across the pairing' using the dual operator $\mathrm{ad}^*_\Omega : \mathfrak{g}^* \to \mathfrak{g}^*$, defined by $\mathrm{ad}^*_\Omega\,\mu\cdot\eta = \mu\circ\mathrm{ad}_\Omega\,\eta$ for all $\eta \in \mathfrak{g}$ and $\mu \in \mathfrak{g}^*$, to get

$$\frac{d}{d\varepsilon}\mathcal{L}(\gamma_\varepsilon)\bigg|_{\varepsilon=0} = \int_I \left(\frac{\delta L}{\delta g}(\gamma, \Omega) + \left(\mathrm{ad}^*_\Omega - \frac{d}{dt}\right)\frac{\delta L}{\delta\Omega}(\gamma, \Omega)\right)\cdot\eta\,dt.$$

Thus γ is a critical point of $\mathcal{L}$ with respect to smooth variations η satisfying $\eta(t_0) = \eta(t_1) = 0$ if and only if

$$\frac{d}{dt}\left(\frac{\delta L}{\delta\Omega}(\gamma, \Omega)\right) = \frac{\delta L}{\delta g}(\gamma, \Omega) + \mathrm{ad}^*_\Omega\frac{\delta L}{\delta\Omega}(\gamma, \Omega).$$

The above construction generalizes naturally to arbitrary Lie groups, even infinite dimensional ones. (Just use the appropriate linearizations of left or right multiplication in place of matrix multiplication.) Since any vector space is an additive Lie group, the vector space case can be regarded as a special case of the Lie group one.

Conservation of energy. Given a Lagrangian $L : G \times \mathfrak{g} \to \mathbb{R}$, define the associated *energy* $E : G \times \mathfrak{g} \to \mathbb{R}$ by

$$E(g, \Omega) := \frac{\delta L}{\delta \Omega}(\gamma, \Omega) \cdot \Omega - L(g, \Omega).$$

If $g(t), \Omega(t))$ is a solution of the Euler-Lagrange equations, then

$$\frac{d}{dt} E(g, \Omega) =$$

Example 4. $G = \mathrm{SO}(3)$.

The algebra $so(3)$ of $SO(3)$ is isomorphic to $\mathbb{R}^3$; $\xi \in \mathbb{R}^3$ maps to $\hat{\xi}$, the skew symmetric real matrix satisfying $\hat{\xi}\mathbf{y} = \xi \times \mathbf{y}$ for all $\mathbf{y} \in \mathbb{R}^3$.

$$\begin{aligned}
[\hat{\xi}, \hat{\zeta}]\mathbf{y} &= (\hat{\xi}\,\hat{\zeta} - \hat{\zeta}\,\hat{\xi})\mathbf{y} \\
&= \xi \times (\zeta \times \mathbf{y}) - \zeta \times (\xi \times \mathbf{y}) \\
&= (\xi \times \zeta) \times \mathbf{y} \\
&= \widehat{\xi \times \zeta}\,\mathbf{y}
\end{aligned}$$

implies that the $\mathbb{R}^3$ implementation of ad is $\mathrm{ad}_\xi \zeta = \xi \times \zeta$. It follows that

$$\begin{aligned}
(\mathrm{ad}^*_\Omega \mu) \cdot \eta = \mu \cdot \mathrm{ad}_\Omega \eta &= \mu \cdot (\Omega \times \eta) \\
&= (\mu \times \Omega) \cdot \eta.
\end{aligned}$$

Hence $\mathrm{ad}^*_\Omega \mu = \mu \times \Omega$.

Consider

$$L(g, \Omega) = \frac{1}{2}\Omega \cdot \mathbb{I}\Omega - V(g),$$

where $\mathbb{I}$ is a positive definite symmetric matrix (the inertia tensor) and $V : \mathrm{SO}(3) \to \mathbb{R}$ is smooth.

What are the trivialized derivatives? $\frac{\delta L}{\delta g}(g, \Omega) = -\frac{\delta V}{\delta g}(g) \in \mathbb{R}^3$, which is determined by

$$\frac{\delta V}{\delta g}(g) \cdot \eta = dV(g) \cdot (g\,\hat{\eta})$$

for all $\eta \in \mathbb{R}^3$. Until we choose a specific V, we can't simplify that any further.

$$\begin{aligned}
\frac{\delta L}{\delta \Omega}(g, \Omega) \cdot \zeta &= \left.\frac{d}{d\varepsilon} L(g, \Omega + \varepsilon\,\zeta)\right|_{\varepsilon=0} \\
&= \left.\frac{d}{d\varepsilon}\left(\frac{1}{2}(\Omega + \varepsilon\,\zeta) \cdot \mathbb{I}(\Omega + \varepsilon\,\zeta) - V(g)\right)\right|_{\varepsilon=0} \\
&= \left.\frac{d}{d\varepsilon}\left(\frac{1}{2}\big(\Omega \cdot \mathbb{I}\Omega + \varepsilon(\zeta \cdot \mathbb{I}\Omega + \Omega \cdot \mathbb{I}\zeta) + \epsilon^2\zeta \cdot \mathbb{I}\zeta\big)\right)\right|_{\varepsilon=0} \\
&= \frac{1}{2}(\zeta \cdot \mathbb{I}\Omega + \Omega \cdot \mathbb{I}\zeta) \\
&= (\mathbb{I}\Omega) \cdot \zeta,
\end{aligned}$$

since $\mathbb{I}$ is symmetric. Thus $\frac{\delta L}{\delta \Omega}(g, \Omega) = \mathbb{I}\Omega$ and the Euler-Lagrange equations are

$$\frac{d}{dt}\mathbb{I}\Omega = -\frac{\delta V}{\delta g}(g) + \mathrm{ad}^*_\Omega(\mathbb{I}\Omega).$$

Since $\mathbb{I}$ is constant, we have

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega - \frac{\delta V}{\delta g}(g).$$

4. January 25, the free rigid body and heavy top. Action and energy.

Special case 1: $V \equiv 0$. The free rigid body.

The evolution equations are

$$\dot{g} = g\hat{\Omega} \qquad \text{and} \qquad \mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega.$$

This system has one scalar and one vector-valued conserved quantity: Define

$$E(g,\Omega) := \tfrac{1}{2}\Omega \cdot \mathbb{I}\Omega \qquad \text{and} \qquad \alpha(g,\Omega) := g\mathbb{I}\Omega.$$

E is the *energy* (more about that soon) and α is the *spatial angular momentum.* If $(g(t),\Omega(t))$ is a solution of the Euler-Lagrange equations for $V \equiv 0$, then

$$\begin{aligned}\frac{d}{dt}E(g(t),\Omega(t)) &= \frac{1}{2}(\dot{\Omega} \cdot \mathbb{I}\Omega + \Omega \cdot \mathbb{I}\dot{\Omega}) \\ &= \Omega \cdot \mathbb{I}\dot{\Omega} \\ &= \Omega \cdot (\mathbb{I}\Omega) \times \Omega \\ &= 0\end{aligned}$$

and

$$\begin{aligned}\frac{d}{dt}\alpha(g(t),\Omega(t)) &= \dot{g}\mathbb{I}\Omega + g\mathbb{I}\dot{\Omega} \\ &= g\hat{\Omega}\mathbb{I}\Omega + g((\mathbb{I}\Omega) \times \Omega) \\ &= g(\Omega \times (\mathbb{I}\Omega) + (\mathbb{I}\Omega) \times \Omega) \\ &= 0.\end{aligned}$$

Use of the conserved quantities in analysing the system: Given initial data (g_0,Ω_0), if we set set $\mu = \alpha(g_0,\Omega_0) = g_0\mathbb{I}\Omega_0$, then

$$\alpha(g(t),\Omega(t)) = \alpha(g_0,\Omega_0) = \mu \qquad \Longrightarrow \qquad \Omega(t) = \mathbb{I}^{-1}g(t)^T\mu.$$

($g(t) \in SO(3)$ implies $g(t)^T = g(t)^{-1}$; I've used g^T, since it's faster to compute.) The evolution equation for g reduces to the first order IVP

$$\dot{g} = g\,\widehat{\mathbb{I}^{-1}g^T\mu} \qquad\qquad g(0) = g_0$$

on $SO(3)$. Given a solution of that first order system, the trivialized velocity can be found by differentiation: $\hat{\Omega}(t) = g(t)^T\dot{g}(t)$, or algebraically: $\Omega(t) = \mathbb{I}^{-1}g(t)^T\mu$.

Sneak preview of reduction, part I. We can also solve the equations on some interval $[0,T]$ by first solving the first order IVP

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega \qquad\qquad \Omega(0) = \Omega_0$$

on $\mathbb{R}^3$, then using that solution to define a linear, time dependent IVP $\dot{g}(t) = g(t)\hat{\Omega}(t)$ on $SO(3)$ and solving for $g(t)$. Why would we do this? This approach involves solving two systems of first order IVPs in succession, rather than just one. If you're using numerical approximations, you're probably better off just solving the first order system for $g(t)$, but the two-step approach arguably offers some visual intuition that isn't quite so accessible in the 'just work on $SO(3)$' approach.

For the sake of convenience and tradition, introduce the *body angular momentum* $\mathbf{M} = \mathbb{I}\Omega = g\alpha(g,\Omega)$ [I need to find boldface Greek letters—I'd like to have all of my vectorial quantities be boldface.]. The body angular momentum satisfies the evolution equation

$$\dot{\mathbf{M}} = \mathbf{M} \times \mathbb{I}^{-1}\mathbf{M}.$$

Figure 5. Heavy top. Left: stolen off the web (some labels aren't relevant to our formulation); center: familiar implementation with fixed point (also stolen off the web); right: level sets of the conserved quantities in body angular momentum space ($\mathbf{M}$ space), from *Mechanics and Symmetry*, Marsden & Ratiu.

We can express the energy and spatial angular momentum in terms of g and $\mathbf{M}$:

$$E(g, \mathbb{I}^{-1}\mathbf{M}) = \tfrac{1}{2}(\mathbb{I}^{-1}\mathbf{M}) \cdot \mathbb{I}(\mathbb{I}^{-1}\mathbf{M}) = \tfrac{1}{2}\mathbf{M} \cdot \mathbb{I}^{-1}\mathbf{M}$$

and

$$\alpha(g, \mathbb{I}^{-1}\mathbf{M}) = g\mathbb{I}(\mathbb{I}^{-1}\mathbf{M}) = g\mathbf{M}.$$

Whether we use Ω or $\mathbf{M}$, the energy doesn't depend on g; the spatial angular momentum does, but its magnitude doesn't. Conservation of $\|\mathbf{M}\|$ implies $\mathbf{M}(t)$ stays on a sphere centered at 0 (for now, this is where the convenience of $\mathbf{M}$ relative to Ω comes in—I find spheres easier to visualize than level sets of $\|\mathbb{I}\Omega\|$). Conservation of $\frac{1}{2}\mathbf{M} \cdot \mathbb{I}^{-1}\mathbf{M}$ implies that $\mathbf{M}(t)$ stays on an ellipsoid determined by $\mathbb{I}$ and $\mathbf{M}_0$. Intersections of these sets determine the traces of solutions $\mathbf{M}(t)$, unless $\mathbb{I}$ is a scalar multiple of the identity (in that case, the two scalar conservation laws are equivalent). The right hand (yellow background) plot in Figure 4 shows the intersections of representative level sets of the energy expressed as a function of $\mathbf{M}$ with a representative sphere in $\mathbf{M}$–space.

Solution curves for the evolution of $\mathbf{M}$ are relatively easy to visualize; *if* you are good at inferring the actual rotation from the body angular momentum, you can guesstimate the motion of the body from $\mathbf{M}(t)$. WARNING: do *not* confuse $\mathbf{M}$ and Ω when attempting this; since the evolution equation can be written in mixed notation as $\dot{\Omega} = \mathbf{M} \times \Omega$, you can see that the only time this won't cause trouble is when there's not much going on for that particular trajectory anyway—if Ω and $\mathbf{M}$ are parallel, Ω (and hence $\mathbf{M}$) is constant and the body just rotates at a fixed rate about a fixed axis forever.

Special case 2: Heavy top. Take $V(g) = g\mathbf{m} \cdot \mathbf{e}_3$, where $\mathbf{e}_3$ is the upward unit vertical vector and $\mathbf{m}$ is the vector from a fixed point in the top to the center of mass, scaled by the mass (see the middle picture in Figure 4—the top is not free to move across the supporting surface, but has a 'universal joint' at one point). Consider $V : \mathrm{SO}(3) \to \mathbb{R}$. Then

$$\begin{aligned} dV(g)(g\hat{\eta}) &= (g\hat{\eta}\mathbf{m}) \cdot \mathbf{e}_3 = (\eta \times \mathbf{m}) \cdot g^T\mathbf{e}_3 \\ &= (\mathbf{m} \times g^T\mathbf{e}_3) \cdot \eta \end{aligned}$$

Hence $\frac{\delta V}{\delta g}(g) = \mathbf{m} \times g^T\mathbf{e}_3$ and the Euler-Lagrange equations for the heavy top are

$$\dot{g} = g\,\hat{\Omega} \qquad \text{and} \qquad \mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega - \mathbf{m} \times g^T\mathbf{e}_3.$$

The conserved quantities are now:

$$\begin{aligned} E(g, \Omega) &= \frac{1}{2}\Omega \cdot \mathbb{I}\Omega + (g\mathbf{m}) \cdot \mathbf{e}_3 && \text{(energy)} \\ \alpha_{\mathrm{vert}}(g, \Omega) &= (g\mathbb{I}\Omega) \cdot \mathbf{e}_3 && \text{(vertical component of angular momentum).} \end{aligned}$$

Where does $\mathbb{I}$ come from? The kinetic energy of a rigid body is the integral over the body of the kinetic energies of all particles forming the body. If the body rotates about some fixed point, then there is a curve $g(t) \in \mathrm{SO}(3)$ such that the trajectory $x(t,X)$ of the particle with reference position X relative to the fixed point is $x(t,X) = g(t)\,X$. (If X is the initial position of the particle, then $g(0) = \mathbb{1}$.) The mass is given by a density function $\rho : B_{\mathrm{ref}} \to \mathbb{R}^+$ where B_{ref} is the reference configuration. The total kinetic energy of the rotating body is

$$
\begin{aligned}
KE = \int_{B_{\mathrm{ref}}} \frac{\rho(X)}{2} \left\|\dot{x}(t,X)\right\|^2 \, dX \\
&= \int_{B_{\mathrm{ref}}} \frac{\rho(X)}{2} \left\|\dot{g}\,X\right\|^2 \, dX \\
&= \frac{1}{2}\int_{B_{\mathrm{ref}}} \rho(X) \left\|g\hat{\Omega}X\right\|^2 \, dX \\
&= \frac{1}{2}\int_{B_{\mathrm{ref}}} \rho(X) \left\|\Omega \times X\right\|^2 \, dX,
\end{aligned}
$$

since $g \in SO(3)$ preserves the Euclidean norm. Now use the vector identity

$$
\begin{aligned}
\|\Omega \times X\|^2 &= \|x\|^2 \|\Omega\|^2 - (X \cdot \Omega)^2 \\
&= \Omega \cdot \left(\|X\|^2 \mathbb{1} - XX^T\right) \Omega,
\end{aligned}
$$

since $X \cdot \Omega = X^T \Omega$. Integrating this expression over the body, keeping in mind that Ω is independent of X, gives

$$
\begin{aligned}
KE &= \frac{1}{2}\int_{B_{\mathrm{ref}}} \rho(X)\Omega \cdot \left(\|X\|^2 \mathbb{1} - XX^T\right) \Omega \, dX \\
&= \frac{1}{2}\Omega \cdot \underbrace{\left(\int_{B_{\mathrm{ref}}} \rho(X) \left(\|X\|^2 \mathbb{1} - XX^T\right) dX\right)}_{\mathbb{I}} \Omega.
\end{aligned}
$$

Exercise 3. *Sneak preview of reduction, part II.* The evolution equation for Ω in the heavy top system depends on g only through $g^T \mathbf{e}_3$. Given a solution $(g(t), \Omega(t))$ of the heavy top equations, define $\gamma(t) := g(t)^T \mathbf{e}_3 \in \mathbb{S}^2$. Show that (γ, Ω) satisfy

$$
\begin{aligned}
\dot{\gamma} &= \gamma \times \Omega \\
\mathbb{I}\,\dot{\Omega} &= (\mathbb{I}\,\Omega) \times \Omega + \gamma \times m.
\end{aligned}
$$

Show that the functions $C_1, C_2 : \mathbb{S}^2 \times \mathbb{R}^3 \to \mathbb{R}$ given by

$$
C_1(\gamma, \Omega) := \tfrac{1}{2}\Omega \cdot \mathbb{I}\Omega + \gamma \cdot m \qquad \text{and} \qquad C_2(\gamma, \Omega) := (\mathbb{I}\Omega) \cdot \gamma
$$

are constants of the motion, i.e. are constant along solutions of the induced dynamical system on $\mathbb{S}^2 \times \mathbb{R}^3$. (*Note:* $\mathbb{S}^2 \times \mathbb{R}^3 \approx \mathrm{SO}(3)/\mathbb{S}^1 \times \mathbb{R}^3 \approx (\mathrm{SO}(3) \times \mathbb{R}^3)/\mathbb{S}^1$.)

Fiber derivatives, action, energy.

Given a Lagrangian on a tangent bundle, we can construct an associated energy function that generalizes the energies from the examples. Let Q be an arbitrary manifold, and $L : TQ \to \mathbb{R}$ be a smooth function. Define

i. the fiber derivative $\mathbb{F}L : TQ \to T^*Q$ by

$$
\mathbb{F}L(v_q) \cdot w_q := \left.\frac{d}{d\varepsilon} L(v_q + \varepsilon w_q)\right|_{\varepsilon=0}
$$

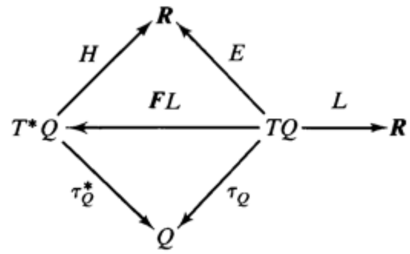


FIGURE 6. Fiber derivatives, energy, etc.; here $\tau_Q = \pi_{TQ}$ and $\tau_Q^* = \pi_{T^*Q}$. (From *Foundations of Mechanics*.)

ii. the action $A : TQ \to \mathbb{R}$ by

$$A(v_q) := \mathbb{F}L(v_q) \cdot v_q = \frac{d}{d\varepsilon} L((1+\varepsilon)v_q)\Big|_{\varepsilon=0}$$

iii. the energy $E : TQ \to \mathbb{R}$ by $E(v_q) := A(v_q) - L(v_q)$.

Example 5. Let $Q = X$ be a vector space, $v_q \sim (q, v) \in X^2$.

$$\begin{aligned}\mathbb{F}L(q,v) \cdot (q,w) &= \frac{d}{d\varepsilon} L(q, v + \varepsilon w)\Big|_{\varepsilon=0} \\ &= \frac{\partial L}{\partial v}(q,v) \cdot w\end{aligned}$$

Here $\mathbb{F}L$ is modeled as taking values in X^*.

Special case, $L(q,v) = \frac{1}{2}(mv) \cdot v - V(q)$ where $m \in \mathcal{L}(X, X^*)$ is symmetric and $V : X \to \mathbb{R}$.

$$\begin{aligned}\mathbb{F}L(q,v) \cdot (q,w) &= \frac{d}{d\varepsilon}\left(\frac{1}{2} m(v+\varepsilon w) \cdot (v + \varepsilon w) - V(q)\right)\Big|_{\varepsilon=0} \\ &= \frac{1}{2}\left((mv) \cdot w + (mw) \cdot v\right) \\ &= (mv) \cdot w.\end{aligned}$$

Thus $\mathbb{F}L(q,w) = (q, mv)$, $A(q,v) = (mv) \cdot v$, and

$$E(q,v) = (mv) \cdot v - \left(\frac{1}{2}(mv) \cdot v - V(q)\right) = \frac{1}{2}(mv) \cdot v + V(q).$$

Example 6. Let $Q = G$ a matrix group. Consider the left trivializations of TG and T^*G.

Given $\ell : G \times \mathfrak{g} \to \mathbb{R}$, define $L : TG \to \mathbb{R}$ by $L(g\Omega) := \ell(g, \Omega) =$; then

$$\mathbb{F}L(g\Omega) \cdot (g\eta) = \frac{d}{d\varepsilon} \ell(g, \Omega + \varepsilon\eta)\Big|_{\varepsilon=0} = \frac{\delta \ell}{\delta \Omega}(g, \Omega) \cdot \eta$$

for $g \in G$, $\Omega, \eta \in \mathfrak{g}$. The action and energy satisfy

$$A(g\Omega) = \frac{\delta \ell}{\delta \Omega}(g, \Omega) \cdot \Omega \qquad \text{and} \qquad E(g\Omega) = \frac{\delta \ell}{\delta \Omega}(g, \Omega) \cdot \eta - \ell(g, \Omega).$$

The solutions of the Euler-Lagrange equations on a vector space or matrix group conserve the energy. In the matrix group case, if $(g(t), \Omega(t))$ satisfies the Euler-Lagrange equations determined by $\ell : G \times \mathfrak{g} \to \mathbb{R}$, then

$$\begin{aligned}
\frac{d}{dt}E(g,\Omega) &= \frac{d}{dt}\left(\frac{\delta\ell}{\delta\Omega}(g,\Omega)\cdot\Omega - \ell(g,\Omega)\right)\\
&= \frac{d}{dt}\left(\frac{\delta\ell}{\delta\Omega}(g,\Omega)\right)\cdot\Omega + \frac{\delta\ell}{\delta\Omega}(g,\Omega)\cdot\dot{\Omega} - \frac{\delta\ell}{\delta g}(g,\Omega)\cdot\Omega - \frac{\delta\ell}{\delta\Omega}(g,\Omega)\dot{\Omega}\\
&= \left(\frac{d}{dt}\left(\frac{\delta\ell}{\delta\Omega}(g,\Omega)\right) - \frac{\delta\ell}{\delta g}(g,\Omega)\right)\cdot\Omega\\
&= \operatorname{ad}^*_\Omega \frac{\delta\ell}{\delta\Omega}(g,\Omega)\cdot\Omega\\
&= \frac{\delta\ell}{\delta\Omega}(g,\Omega)\cdot\operatorname{ad}_\Omega\Omega\\
&= 0,
\end{aligned}$$

since $\operatorname{ad}_\Omega\Omega = [\Omega,\Omega] = 0$. The vector space calculation is analogous.

Principle of least action

Recall that for a one dimensional Euler-Lagrange system, we used conservation of energy to implicitly solve the system: if $L(q,v) = \frac{m}{2}\dot{q}^2 - V(q)$, then $E(q,\dot{q}) = \frac{m}{2}\dot{q}^2 + V(q)$. To implicitly find the solution of the IVP with initial data $q(0 = q_0$ and $\dot{q}(0) = v_0$, note that conservation of energy gives

$$E(q,\dot{q}) = E(q_0,v_0) = e, \qquad \text{and hence} \qquad |\dot{q}| = \sqrt{\frac{2}{m}(e - V(q))}.$$

On an interval $[t_0,t]$ of time for which $\dot{q} \neq 0$,

$$t - t_0 = \pm\int_{q_0}^{q} \frac{dq}{\sqrt{\frac{2}{m}(e - V(q))}}.$$

Great, but if you can't get down to one dimension, how much does conservation of energy help in understanding they system?

Claim: The solutions of the variational problem of finding the critical points of

$$\mathcal{L}(\gamma) = \int_a^b L(\gamma(t),\dot{\gamma}(t))\,dt$$

over curves $\gamma : [a,b] \to Q$ with $\gamma(a) = q_0$, $\gamma(b) = q_1$ is equivalent to the variational problem of finding critical points of

$$\mathcal{A}(\gamma) = \int_a^b A(\gamma(t),\dot{\gamma}(t))\,dt$$

over curves $\gamma : [\alpha,\beta] \to Q$ with $\gamma(\alpha) = q_0$, $\gamma(\beta) = q_1$, $E\circ\gamma \equiv e$, where A is the action of L, E is the energy and $[\alpha,\beta]$ is not fixed.

We allow different durations by introducing reparametrizations of time. Given an interval $I = [a,b]$ and endpoints $q_0, q_1 \in Q$, let $\Omega(I,q_0,q_1)$ denote the set of smooth maps $\gamma : I \to Q$ satisfying the boundary conditions $\gamma(a) = q_0$ and $\gamma(b) = q_1$, and let $\Omega(I,q_0,q_1,e)$ denote the set of smooth maps (τ,γ) where $\tau : I \to \mathbb{R}$ satisfies $\dot{\tau} > 0$ for $t \in I$ and $\gamma \in \Omega(\tau(I),q_0,q_1)$ satisfies the energy constraint $E(\gamma(t),\dot{\gamma}(t)) \equiv e$ for $t \in \tau(I)$. Assume that e is a regular value of E, so that $E^{-1}(e)$ is a smooth manifold.

Given a curve Γ_ε in $\Omega(I,q_0,q_1)$ and $\tau_\varepsilon : I \to \mathbb{R}$, with $\tau_0(t) = t$ and $\dot{\tau}_\varepsilon > 0$ for all $t \in I$, find conditions on $\tau_\varepsilon, \Gamma_\varepsilon$ such that $(\tau_\varepsilon, \Gamma_\varepsilon \circ \tau^{-1}) \in \Omega(I,q_0,q_1,e)$. If we define $\gamma_\varepsilon = \Gamma_\varepsilon \circ \tau_\varepsilon^{-1}$, then

$$\dot{\Gamma}_\varepsilon(t) = \frac{d}{dt}(\gamma_\varepsilon\circ\tau_\varepsilon)(t) = \frac{d\gamma_\varepsilon}{d\tau}(\tau_\varepsilon(t))\dot{\tau}_\varepsilon(t)$$

implies

$$\frac{d\gamma_\varepsilon}{d\tau} \circ \tau_\varepsilon = \frac{1}{\dot{\tau}_\varepsilon} \dot{\Gamma}_\varepsilon.$$

Hence the energy is constant along γ_ε iff

$$\begin{aligned} 0 &\equiv \frac{d}{d\varepsilon} E\left(\dot{\gamma}_\varepsilon(t)\right)\Big|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} E\left(\frac{1}{\dot{\tau}_\varepsilon}\dot{\Gamma}_\varepsilon(t)\right)\Big|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} E(\dot{\Gamma}_\varepsilon(t))\Big|_{\varepsilon=0} + \frac{d}{d\varepsilon}\frac{1}{\dot{\tau}_\varepsilon(t)}\Big|_{\varepsilon=0} \mathbb{F}E(\dot{\gamma}(t)) \cdot \dot{\gamma}(t). \end{aligned}$$

If the fiber derivative term $\mathbb{F}E(\gamma, \dot{\gamma}) \cdot \dot{\gamma}$ is nonzero for all $t \in I$, we have the ODE

$$\dot{\delta\tau} = \frac{\frac{d}{d\varepsilon} E(\Gamma_\varepsilon, \dot{\Gamma}_\varepsilon)\Big|_{\varepsilon=0}}{\mathbb{F}E(\gamma, \dot{\gamma}) \cdot \dot{\gamma}}$$

for the linearization $\delta\tau = \frac{d}{d\varepsilon}\tau_\varepsilon\big|_{\varepsilon=0}$ of the reparametrization of time. Since γ is either constant or has zero velocity only at isolated points, if the fiber derivative of the energy is non-degenerate, the linearized energy constraint can be satisfied for any Γ_ε in $\Omega(I, q_0, q_1, e)$ except at isolated points, by appropriate choice of τ. Using $E = A - L$, and hence $A|_{\Omega(I,q_0,q_1,e)} = e + L$, we can now relate the two variational problems.

If Q is a vector space, then $E(q, v) = \frac{\partial L}{\partial v}(q, v) \cdot v - L(q, v)$ and

$$\mathbb{F}E(q,v)\cdot w = \frac{\partial E}{\partial v}(q,v)\cdot w = \frac{\partial^2 L}{\partial v^2}(q,v)(v,w) + \frac{\partial L}{\partial v}(q,v)\cdot w - \frac{\partial L}{\partial v}(q,v)\cdot w = \frac{\partial^2 L}{\partial v^2}(q,v)(v,w).$$

Hence if $\partial^2 L/\partial v^2$ is non-degenerate and $\dot{\gamma}$ is everywhere nonzero,

$$\dot{\delta\tau} = \frac{\frac{d}{d\varepsilon} E(\Gamma_\varepsilon, \dot{\Gamma}_\varepsilon)\Big|_{\varepsilon=0}}{\frac{\partial^2 L}{\partial v^2}(\gamma, \dot{\gamma})(\dot{\gamma}, \dot{\gamma})}.$$

Exercise 4. Consider the Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by $L(\mathbf{q}, \mathbf{v}) := \frac{1}{2}\|\mathbf{v}\|^2$. Given two points $\mathbf{q}_0$ and $\mathbf{q}_1 \in \mathbb{R}^n$, show that the solution of the Euler-Lagrange equations determined by L with end points $\mathbf{q}_0$ and $\mathbf{q}_1$ is the constant speed straight parametrized curve connecting those points. Then show that there is no parametrized curve $\gamma : [a, b] \to \mathbb{R}^n$ minimizing the total action $\mathcal{A}(\gamma) = \int_a^b A(\gamma(t), \dot{\gamma}(t))\, dt$ over the set of smooth curves with arbitrary domain and without an energy constraint; decreasing the speed along the straight path decreases $\mathcal{A}$. Finally, show that if a constant (nonzero) energy e is specified, the total action is minimized over the set of smooth curves from $\mathbf{q}_0$ and $\mathbf{q}_1$ with that pointwise energy by the constant speed straight parametrized curve with the energy e.

The action and the canonical one-form

Given a curve $\gamma : I \to Q$, let $\mathcal{C}(\gamma) = \{\dot{\gamma}(t) : t \in I\} \subset TQ$ denote the trace of $\dot{\gamma}$. Then

$$\mathcal{A}(\gamma) = \int_I A(\dot{\gamma}(t))\, dt = \int_I \mathbb{F}L(\dot{\gamma}(t)) \cdot \dot{\gamma}(t)\, dt = \int_{\mathcal{C}(\gamma)} \mathbb{F}L(v_q)\, dq.$$

We can regard the 1-form $v_q \mapsto \mathbb{F}L(v_q)dq$ on TQ as the pullback by $\mathbb{F}L$ of the canonical 1-form Θ_0 on T^*Q, which is defined as follows. Let $\pi_{T^*Q} : T^*Q \to Q$ denote the canonical projection on T^*Q, namely $\pi_{T^*Q}(\mu_q) = q$ for $\mu_q \in T^*Q$ and $q \in Q$. Then

$$\Theta_0(\mu_q) \cdot w_{\mu_q} := \mu_q \cdot \left({\pi_{T^*Q}}^* w_{\mu_q}\right).$$

The canonical 1-form on the cotangent bundle of a vector space satisfies

$$\Theta_0(q, p)(q, p, v, w) = (q, p) \cdot (q, v) = p \cdot v,$$

often expressed as $\Theta_0(q,p) = pdq$.

Example 7. Let $Q = G \subseteq \mathrm{GL}(n)$ be a matrix group. Define

$$\tilde{\tau}_{L,R} : G \times \mathfrak{g}^* \to T^*G$$

by

$$\tilde{\tau}_L(g,\mu) = g^{-T}\mu$$
$$\tilde{\tau}_R(g,\mu) = \mu g^{-T}$$

respectively the linearizations left and right multiplication by g^{-1}.

Then

$$\tilde{\tau}_L(g,\mu)\cdot(g\xi) = \left\langle g^{-T}\mu,\, g\xi\right\rangle = \left\langle \mu,\, g^{-1}g\xi\right\rangle = \left\langle \mu,\, \xi\right\rangle.$$

The computation for $\tilde{\tau}_R$ is analogous. Compute $\tilde{\Theta}_0 = \tilde{\tau}_L^*\Theta_0$, 1-form on $T(G\times\mathfrak{g}^*) \cong TG\times\mathfrak{g}\times\mathfrak{g}$

$$T_{(g,\mu)}\left(\pi_{T^*G}\circ\tilde{\tau}_L\right)(g\xi,\mu,\nu) = g\xi.$$

Hence

$$\begin{aligned}\tilde{\Theta}_0(g,\mu)\cdot(g\xi,\mu,\nu) &= \tilde{\tau}_L(g,\mu)\cdot T_{(g,\mu)}\left(\pi_{T^*G}\circ\tilde{\tau}_L\right)(g\xi,\mu,\nu)\\ &= \tilde{\tau}_L(g,\mu)\cdot g\xi\\ &= \left\langle\mu,\,\xi\right\rangle.\end{aligned}$$

Remark 8. Given $L : TQ \to \mathbb{R}$,

$$\left(\mathbb{F}L^*\Theta_0\right)(v_q) = \Theta_0\left(\mathbb{F}L(v_q)\right)\cdot\mathbb{F}L^*.$$

Formally,

$$\begin{aligned}\Theta_0\left(\mathbb{F}L(v_q)\right)\cdot\mathbb{F}L^* &= \mathbb{F}L(v_q)\cdot{\pi_{T^*Q}}^*\circ\mathbb{F}L^*\\ &= \mathbb{F}L(v_q)\cdot{\pi_{TQ}}^*.\end{aligned}$$

Remark 9. Given a 1-form β on Q, then $\beta^*\Theta_0 = \beta$. The proof is unwinding of definitions:

$$\begin{aligned}(\beta^*\Theta_0)(q)\cdot v_q &= \underbrace{\Theta_0(\beta(q))}_{\in T^*_{\beta(q)}T^*Q}\cdot\underbrace{\beta^*\cdot v_q}_{T_{\beta(q)}T^*Q}\\ &= \beta(q)(v_q)\end{aligned}$$

5. February 1–7. Symplectic structures and Hamilton's equations

Definition 10 (Vertical lift)**.** Given $\mu_q, \nu_q \in T^*_q$, define $\mathrm{vert}_{\mu_q} : T^*_qQ \to T_{\mu_q}T^*Q$ by

$$\mathrm{vert}_{\mu_q}(\nu_q) = \left.\frac{d}{d\varepsilon}(\mu_q + \varepsilon\nu_q)\right|_{\varepsilon=0}$$

Compute

$$T_{\mu_q}\pi_{T^*Q}\cdot\mathrm{vert}_{\mu_q}(\nu_q) = \left.\frac{d}{d\varepsilon}\pi_{T^*Q}(\mu_q+\varepsilon\nu_q)\right|_{\varepsilon=0} = \left.\frac{d}{d\varepsilon}q\right|_{\varepsilon=0} = 0$$

vert_{μ_q} is one-to-one, so $\mathrm{vert}_{\mu_q}(T^*_qQ)$ is isomorphic to T^*_qQ.

Lemma 11. *If $\mu_q = \beta(q)$ for some 1-form β on Q, then given $w_{\mu_q} \in T_{\mu_q}T^*Q$, let $v_q = T_{\mu_q}\pi_{T^*Q}\cdot w_{\mu_q} \in T_qQ$. Then $w_{\mu_q} - \beta^*v_q \in \mathrm{vert}_{\mu_q}(T^*_qQ)$. That is, there is $\nu_q \in T^*_qQ$ such that*

$$w_{\mu_q} = \beta^*\nu_q + \mathrm{vert}_{\mu_q}(\nu_q)$$

Proof. It suffices to show that $\ker\left({\pi_{T^*Q}}^*\right) = \mathrm{vert}_{\mu_q}(T^*Q)$, since

$${\pi_{T^*Q}}^*\left(w_{\mu_q} - \beta^*v_q\right) = {\pi_{T^*Q}}^*w_{\mu_q} - (\pi_{T^*Q}\circ\beta)^*v_q = v_q - v_q = 0.$$

Since $\mathrm{vert}_{\mu_q}(T^*Q) \approx T^*_q(Q)$ in finite dimensions it is sufficient to argue that

$$\dim\ker\left({\pi_{T^*Q}}^*\right) = \dim Q,$$

which in turn happens when $\operatorname{rank}(\pi_{T^*Q}{}^*) = \dim Q$. Since $\pi_{T^*Q} \circ \beta$ is the identity, the restriction of $\pi_{T^*Q}{}^*$ to $\beta^*(T_qQ)$ is an isomorphism, and hence $\operatorname{rank}(\pi_{T^*Q}{}^*) \geq \dim Q$. The rank-nullity theorem completes the proof of the claim.

□

Definition 12 (Symplectic structure)**.** A *symplectic structure* ω on a manifold M is a closed, (weakly) non-degenerate 2-form on M.
(Jargon: ω is weakly non-degenerate if the kernel of $(u, w) \mapsto i_{(u,w)}\omega$ is trivial and strongly non-degenerate if the map is also onto. 'Weak' is relevant only in the infinite-dimensional setting.)

Example 13. Examples of symplectic structures.

- The canonical symplectic structure on a cotangent bundle: $\omega_0 := -d\Theta_0$.
  Note: Many sources use the opposite sign convention, but also use a different sign convention when defining Hamilton's equations. so a given Hamiltonian yields the same vector field for everyone.
- A Lagrangian $L : TQ \to \mathbb{R}$ is said to be *regular* if $\mathbb{F}L^*$ is surjective at all points of Q. The pullback of the canonical symplectic structure on T^*Q to TQ by the fiber derivative $\mathbb{F}L$ of a regular Lagrangian is a symplectic structure on TQ.
- $M = \mathbb{S}^2$, The area element $\omega(m)(u \times m, w \times m) = m \cdot (u \times w)$ for $m \in \mathbb{S}^2$, $u, w \in \mathbb{R}^2$ is non-degenerate; any 2-form on a two dimensional manifold is closed.

The separate document 'The canonical symplectic structure on a cotangent bundle in terms of a vertical-horizontal split' (canonicalsymp.pdf) shows one approach to 'componentizing' the canonical symplectic structure. Comments on those notes: Sorry about the Marsden-style tangent map notation. (I think the LaTeX source is trapped on an old machine and I haven't gotten around to retyping it.) The $\epsilon\,\alpha$ term: To clarify what is, and what isn't, being scaled by ϵ, let α_ϵ denote the parametrized family of one-forms satisfying $\alpha_\epsilon(q) = \epsilon\,(\alpha(q))$, with multiplication by ϵ indicating scalar multiplication within the vector space T_q^*Q. (In particular, ϵ does *not* rescale the base point q; if Q is nonlinear, this rescaling isn't defined.) Then the flow of the vector field $Y_\alpha(\mu_q) = \operatorname{vert}_{\alpha(q)}(\mu_q)$ is $\mathcal{F}_\epsilon(\mu_q) = \mu_q + \alpha_\epsilon(q)$, i.e. $\mathcal{F}_\epsilon = \mathrm{id} + \alpha_\epsilon \circ \pi_{T^*Q}$.

If ω is a (strongly) symplectic structure of M, then a any smooth function H on M determines a vector field X_H via

$$i_{X_H}\omega = dH.$$

X_H is called the *Hamiltonian vector field* of the *Hamiltonian* H.

Claim: H is preserved by the flow of X_H.

Proof: If $\mathcal{F}_t$ is the flow map of X_H, then

$$\begin{aligned}
\frac{d}{dt} H \circ \mathcal{F}_t &= \frac{d}{dt}\left(\mathcal{F}_t^* H\right) \\
&= \mathcal{F}_t^*\left(L_{X_H} H\right) \\
&= \mathcal{F}_t^*\left(i_{X_H} dH\right) \qquad \text{(use Cartan's formula } L_X = \iota_X d + d\iota_X\text{)} \\
&= \mathcal{F}_t^*\left(i_{X_H} i_{X_H}\omega\right) = 0.
\end{aligned}$$

Example 14. $M = T^*\mathbb{R}^n = \mathbb{R}^{2n}$.

The canonical symplectic structure on a vector space satisfies

$$\omega(\mathbf{q}, \mathbf{p})\big((\mathbf{u}, \mathbf{w}), (\tilde{\mathbf{u}}, \tilde{\mathbf{w}})\big) = \tilde{\mathbf{w}} \cdot \mathbf{u} - \mathbf{w} \cdot \tilde{\mathbf{u}}.$$

If the partial derivatives of the Hamiltonian $H : \mathbb{R}^{2n} \to \mathbb{R}$ are defined in the usual way, so that

$$dH(\mathbf{q}, \mathbf{p})(\mathbf{u}, \mathbf{w}) = \frac{\partial H}{\partial q}(\mathbf{q}, \mathbf{p}) \cdot \mathbf{u} + \frac{\partial H}{\partial p}(\mathbf{q}, \mathbf{p}) \cdot \mathbf{w}$$

for all $\mathbf{q}, \mathbf{p}, \mathbf{u}, \mathbf{w} \in \mathbb{R}^n$, then

$$dH(\mathbf{q},\mathbf{p})(\mathbf{u},\mathbf{w}) = \omega(\mathbf{q},\mathbf{p}) \left(\left(\frac{\partial H}{\partial p}(\mathbf{q},\mathbf{p}), -\frac{\partial H}{\partial p}(\mathbf{q},\mathbf{p}) \right), (\mathbf{u},\mathbf{w}) \right)$$

implies

$$X_H(\mathbf{q},\mathbf{p}) = \left(\frac{\partial H}{\partial p}(\mathbf{q},\mathbf{p}), -\frac{\partial H}{\partial q}(\mathbf{q},\mathbf{p}) \right).$$

Example 15. $M = \mathbb{S}^2$.

Let $\frac{\delta H}{\delta m} : \mathbb{S}^2 \to \mathbb{R}^3$ satisfy

$$dH(m)(u \times m) = \frac{\delta H}{\delta m}(m) \cdot (u \times m)$$

for $m \in \mathbb{S}^2$ and $u \in \mathbb{R}^3$. This defines $\frac{\delta H}{\delta m}$ modulo scalar multiples of m.

Then

$$\begin{aligned} \omega(m)\left(X_H(m), w \times m\right) &= dH(m)(w \times m) \\ &= \frac{\delta H}{\delta m}(m) \cdot (w \times m) \\ &= m \cdot \left(\frac{\delta H}{\delta m}(m) \cdot (u \times m) \right). \end{aligned}$$

Hence

$$X_H(m) = \frac{\delta H}{\delta m}(m)$$

modulo multiples of m.

Exercise 5. Recall that the pullback of the canonical symplectic structure of T^*G to $G \times \mathfrak{g}^*$ by left trivialization is

$$\omega(g,\mu)\big((L_g{}^*\xi, \mu, \nu), (L_g{}^*\eta, \mu, \tau) \big) = \tau \cdot \xi - \nu \cdot \eta + \mu \cdot [\xi, \eta],$$

where $\xi, \eta \in \mathfrak{g}$ and $\mu, \nu, \tau \in \mathfrak{g}^*$. Given smooth H on $G \times \mathfrak{g}^*$, let $\frac{\delta H}{\delta g} : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ and $\frac{\delta H}{\delta \mu} : G \times \mathfrak{g}^* \to \mathfrak{g}$ be given by

$$dH(g,\mu)(T_e L_g \xi, \mu, \nu) = \frac{\delta H}{\delta g}(g,\mu) \cdot \xi + \nu \cdot \frac{\delta H}{\delta \mu}(g,\mu)$$

for $g \in G$, $\xi \in \mathfrak{g}$ and $\mu, \nu \in \mathfrak{g}^*$.

Let $X_H^g : G \times \mathfrak{g}^* \to \mathfrak{g}$ and $X_H^\mu : G \times \mathfrak{g}^* \to \mathfrak{g}^*$ denote the trivialized components of the Hamiltonian, that is

$$X_H(g,\mu) = \left(L_g{}^* X_H^g(g,\mu), \mu, X_H^\mu(g,\mu) \right).$$

Show that

$$X_H^g(g,\mu) = \frac{\delta H}{\delta \mu}(g,\mu) \qquad \text{and} \qquad X_H^\mu(g,\mu) = -\frac{\delta H}{\delta g}(g,\mu) + \mathrm{ad}^*_{\frac{\delta H}{\delta \mu}(g,\mu)} \mu.$$