Chapter 1

Mechanics

Goal: To describe the basic structures of mechanics as geometrically as possible without making a mess.

We work in the category of smooth manifolds and functions.

1.1 Review some geometric concepts

Let M be a smooth manifold and $q \in M$.

Tangent vectors are given by paths $\gamma : \mathbb{R} \to M$. If $\gamma(0) = q$ then $\dot{\gamma}(0) \in T_q M$. Covectors are given by functions $f : M \to \mathbb{R}$, and $(df)_q \in T_q M$. Composition induces the dual pairing, since $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ and $(f \circ \gamma)'(0) = (df)_q \cdot \dot{\gamma}(0) \in \mathbb{R}$.

A time-invariant first-order ODE on M is equivalent to a vector field X on M. The ODE associated to X is $\dot{\gamma} = X$. The solution to such an ODE is given by the flow $\Phi : M \times \mathbb{R} \to M \cup \{\infty\}$ such that $\gamma(t) = \Phi(\gamma(0), t)$ is a solution for each initial condition $\gamma(0)$.

1.2 Newtonian mechanics

$$F = ma, \quad a = \ddot{q} \quad \implies \quad \ddot{q} = F(q, \dot{q})/m.$$

This is a second-order ODE. How can we express it geometrically? Reduce it to a first-order ODE:

$$\begin{aligned} \dot{q} &= v, \\ \dot{v} &= F(q, \dot{q})/m. \end{aligned}$$

This is a first-order ODE on TM. Thus it is described by some vector field X on TM.

Note: This vector field X on TM is a *second-order vector field*, meaning that $\dot{q} = v$. Geometrically, this means that $(d\pi)_{(q,v)}(X) = v$. If we pick coordinates $\{q^i\}$ on M and corresponding fiberwise coordinates $\{v^i\}$ on fibers of TM, this condition amounts to

$$X(q, v) = (v, F(q, v)/m).$$

1.3 Lagrangian mechanics

Lagrangian mechanics gives us a way to generate second-order vector fields using a variational method.

Given a function $L : TM \to \mathbb{R}$, define a functional S on the path space PM by

$$S(\gamma) := \int_{t_0}^{t_1} L(\gamma(t), \dot{\gamma}(t))\, dt.$$

Whenever L satisfies the appropriate nondegeneracy condition, the Euler-Lagrange equation $\delta S = 0$ defines a second-order ODE.

Specifically, consider a variation of γ by $\delta\gamma$ which holds constant the position and velocity of both endpoints. It's helpful to consider the problem in coordinates, for then we can write

$$\begin{aligned}
\delta S : &= \frac{d}{du}|_{u=0} S(\gamma + u\,\delta\gamma) \\
&= \frac{d}{du}|_{u=0} \int_{t_0}^{t_1} L(\gamma(t) + u\,\delta\gamma(t), \dot{\gamma}(t) + u\,\dot{\delta\gamma}(t))\, dt \\
&= \int_{t_0}^{t_1} \frac{\partial L}{\partial q^i}(\gamma(t), \dot{\gamma}(t)) \cdot (\delta\gamma)^i + \frac{\partial L}{\partial v^i}(\gamma(t), \dot{\gamma}(t)) \cdot \frac{d}{dt}(\delta\gamma)^i(t)\, dt \\
&= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial q^i}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt}\frac{\partial L}{\partial v^i}(\gamma(t), \dot{\gamma}(t)) \right) \cdot (\delta\gamma)^i(t)\, dt.
\end{aligned}$$

This vanishes for all variations $\delta\gamma$ precisely when

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

This is known as the Euler-Lagrange equation.

To solve for $\ddot{\gamma}$ we use the chain rule

$$\frac{\partial L}{\partial q^i}(\gamma(t), \dot{\gamma}(t)) = \frac{\partial^2 L}{\partial v^i \partial q^j}(\gamma(t), \dot{\gamma}(t))\, \dot{\gamma}^j(t) + \frac{\partial^2 L}{\partial v^i \partial v^j}(\gamma(t), \dot{\gamma}(t))\, \ddot{\gamma}^j(t).$$

We get a nondegenerate second-order ODE precisely when the matrix-valued function over TM given by

$$\frac{\partial^2 L}{\partial v^i \partial v^j}(q, v)$$

is nondegenerate. We assume that L was chosen so that this is nondegenerate.

1.4 Legendre transform

We note that the function L induces a map $P_L : TM \to T^*M$ given by

$$P_L(q, v) = (dL)^{\text{vert}}_{(q,v)} \in T^*_{(q,v)} T_q M = (T_{(q,v)} T_q M)^* \cong (T_q M)^* = T^*_q M.$$

In coordinates,

$$P_L(q, v) = \frac{\partial L}{\partial v^i}(q, v)\,(dq^i)_q.$$

The condition that P_L defines a local diffeomorphism, i.e. $dP_L \neq 0$, is equivalent to the nondegeneracy condition

$$\frac{\partial^2 L}{\partial v^i \partial v^j}(q, v),$$

which we assume.

Suppose $\{p_i\}$ are the fiberwise coordinates on T^*M associated with the coordinates $\{q^i\}$. Since P_L is a local diffeomorphism, the $\{p_i\}$ pull back to $\{P_L^*(p_i)\}$ which (together with the $\{q^i\}$) give local coordinates on TM. By definition, the $\{p_i\}$ read off the coefficient of dq^i, so

$$P_L^*(p_i) = \frac{\partial L}{\partial v^i}.$$

Since P_L^* is a local diffeomorphism, we identify p_i with its pullback $P_L^*(p_i)$ to simplify notation.

This allows us to rewrite the Euler-Lagrange equations as

$$\dot{q} = v, \quad \dot{p}_i = \frac{\partial L}{\partial q^i}, \tag{1.1}$$

where $\dot{p}_i$ and $\dot{v}$ determine each other by

$$\left(\frac{\partial^2 L}{\partial v^i \partial v^j}\right) \dot{v}^i = \dot{p}_j.$$

Just as the Euler-Lagrange equation produces dynamics given a function $L : TM \to \mathbb{R}$, there is a dual way to get dynamics from a function $H : T^*M \to \mathbb{R}$.

Theorem. *Suppose L is a nondegenerate Lagrangian. Then the Hamiltonian dynamics of a Hamiltonian H agree with the Lagrangian dynamics of L if and only if*

$$P_L^*(H) = P_L(q, v) \cdot v - L + C \in \Omega^0(TM),$$

where C is any constant.

To make sense of this theorem, we must first explain Hamiltonian mechanics.

1.5 Hamiltonian mechanics

The Poincare one-form $\alpha \in \Omega^1(T^*M)$ is defined by $\alpha(v) = p(\pi_*(v))$ for $v \in T_{(q,p)}T^*M$. In coordinates,

$$\theta = p_i\, dq^i.$$

Taking the exterior derivative,

$$\omega := d\theta = dp_i \wedge dq^i.$$

Note that ω is closed since it is exact. It is an antisymmetric analogue of a metric. In fact, ω is nondegenerate, so we can use it to convert between one-forms and vector fields over T^*M.

We define the *symplectic gradient* X_f of a function $f \in \Omega^0(T^*M)$ by the relation

$$\omega(X_f, Y) = df \cdot Y \quad \text{for all vector fields } Y.$$

In components,

$$\begin{aligned} df &= \frac{\partial f}{\partial q^i} dq^i + \frac{\partial f}{\partial p_i} dp_i, \\ X_f &= \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial}{\partial p_i}. \end{aligned}$$

Any Hamiltonian function $H \in \Omega^0(T^*M)$ produces dynamics given by the flow of the vector field X_H. This leads to the equations of motion

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{1.2}$$

1.6 Proof of the Legendre transform formula

We now prove Theorem 1.4.

Proof. Recall that Lagrangian dynamics are given by (1.1), while Hamiltonian dynamics are given by (1.2). Since they represent possibly distinct systems, we label the dynamic variables with either L or H:

$$\dot{q}^i_L = v^i, \quad \dot{p}^L_i = \frac{\partial L}{\partial q^i}, \quad \dot{q}^i_H = \frac{\partial H}{\partial p_i}, \quad \dot{p}^H_i = -\frac{\partial H}{\partial q^i}. \tag{1.3}$$

The equation $P_L^*(H) = P_L(q, v) \cdot v - L + C$ is equivalent to

$$d(L + P_L^*(H)) = d(P_L(q, v) \cdot v) \tag{1.4}$$

since $d : \Omega^0(TM) \to \Omega^1(TM)$ is linear with kernel equal to the constant functions. The left hand side expands to

$$\frac{\partial L}{\partial q^i} dq^i + \frac{\partial L}{\partial v^i} dv^i + P_L^* \left(\frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial p_i} dp_i \right).$$

Now we substitute the definition $p_i = \frac{\partial L}{\partial v^i}$ and the dynamic equations (1.3) to get

$$\begin{aligned}
&\dot{p}_i^L dq^i + p_i dv^i + P_L^*(-\dot{p}_i^H dq^i + \dot{q}_H^i dp_i) \\
=&(\dot{p}_i^L - \dot{p}_i^H) dq^i + (\dot{q}_H^i - v^i) dp_i + dp_i\, v^i + p_i dv^i \\
=&(\dot{p}_i^L - \dot{p}_i^H) dq^i + (\dot{q}_H^i - \dot{q}_L^i) dp_i + d(p_i v^i).
\end{aligned}$$

Thus (1.4) is satisfied if and only if $\dot{p}_i^L - \dot{p}_i^H = 0$ and $\dot{q}_H^i - \dot{q}_L^i = 0$, i.e. if the dynamics are identical. □

1.7 Homework

1.7.1 Example of a Lagrangian

Let M be $\mathbb{R}^2$. Consider the Lagrangian $L(q^1, q^2, v^1, v^2) = \frac{1}{2}m((v^1)^2 + (v^2)^2) - V(q^1, q^2)$ for some function $V : M \to \mathbb{R}$. In this case, V is called the *potential energy*. Write out the Euler-Lagrange equations, and solve for $\ddot{q}^1$ and $\ddot{q}^2$. Next, compute the Legendre transform and find the corresponding Hamiltonian $H : T^*M \to \mathbb{R}$. Write out the corresponding equations for $\dot{q}^1$, $\dot{q}^2$, $\dot{p}_1$, and $\dot{p}_2$.

1.7.2 Rederive the Euler-Lagrange equation

Without looking, rederive the Euler-Lagrange equations in the case of a single variable. So let $\gamma : [t_0, t_1] \to \mathbb{R}$ represent any curve. Let $\delta\gamma : [t_0, t_1] \to \mathbb{R}$ represent any function satisfying

$$\delta\gamma(t_0) = \delta\gamma(t_1) = \dot{\delta\gamma}(t_0) = \dot{\delta\gamma}(t_1) = 0.$$

Consider the family of curves parameterized by u, given by $\gamma + u\,\delta\gamma : [t_0, t_1] \to \mathbb{R}$. For small u, this represents a small perturbation of γ.

Solve the equation $\delta S(\gamma) = 0$ as follows. Start with the equation

$$0 = \delta S(\gamma) := \frac{d}{du}|_{u=0} S(\gamma + u\,\delta\gamma),$$

where

$$S(\gamma) := \int_{t_0}^{t_1} L(\gamma(t), \dot{\gamma}(t))\, dt.$$

Expand out using the chain rule. Then you get something of the form

$$\int_{t_0}^{t_1} \underline{\hspace{2em}}\, \delta\gamma(t)\, dt + \int_{t_0}^{t_1} \underline{\hspace{2em}}\, \dot{\delta\gamma}(t)\, dt.$$

Use integration by parts on the second term to put it in the form of the first term. (What happens to the integration-by-parts boundary term?)

Now since $\delta\gamma(t)$ is arbitrary (except at the endpoints), the other factor must vanish identically. If all went well, you should now have the Euler-Lagrange equation.

1.7.3 Twin paradox

Suppose there are two twins who begin on a planet which is stationary with respect to the reference frame. One twin is an astronaut, who goes on a voyage with coordinates given by (t, x) starting at $(0, 0)$, flying out to $(2, 1)$, then flying back to $(4, 0)$. The other twin begins at $(0, 0)$ and stays put, reuniting with his brother at $(4, 0)$. How do the ages of the twins compare when they reunite at $(4, 0)$?