Three views of mechanics
John Hubbard, in L. Gross’s course
February 10, 2011

1 Introduction

A mechanical system is manifold with a Riemannian metric $K : TM \to \mathbb{R}$ called kinetic energy and a function $V : M \to \mathbb{R}$ called potential energy. Everything is required to be smooth, or at least of class $C^2$.

There are then equations of motion, which are a flow in $TM$, i.e., a map $\Phi : TM \times \mathbb{R} \to TM$. Intuitively, $\Phi((q, \xi), t)$ should be the position at time $t$ of the system starting out at $(q, \xi)$ at time $0$. This should be the flow of a first order differential equation on $TM$, or of a second-order differential equation on $M$.

I know of three ways of setting up this equation; all of them are a lot harder than one might expect:

1. The Newtonian approach: $F = ma$.

2. The Lagrangian approach: principle of least action.

3. The Hamiltonian approach: simplectic manifolds and the symplectic gradient.

I will examine these one at a time.

1.1 The Newtonian approach

Here $F = -\nabla V$. Already, we need to be careful: $dV$ is naturally a 1-form on $M$, but turning it into the vector-field $\nabla V$ requires an identification of tangent spaces and cotangent space. We have one at hand: $K$ is a non-degenerate quadratic form on $T_q M$ at every point $q \in M$, providing us with such an identification. Still, it might be better to write $\nabla_K V$, to remember that even to define $\nabla V$ we need to know $K$.

But the real problem is $a$, the acceleration. If we have a curve $\gamma : I \to M$, the velocity vector $\gamma'(t)$ is a perfectly well defined element of $T_{\gamma(t)} M$. But

$$
\gamma''(t) = \lim_{h \to 0} \frac{\gamma'(t + h) - \gamma'(t)}{h}
$$
is NOT well defined: the two vectors being subtracted belong to different spaces.

The differential geometers know how to deal with this: they define the covariant derivative $D_{\gamma'(t)}\gamma''(t)$ given by the Levi-Civita connection associated to $K$. This isn’t immensely hard, but it is non-trivial. There is one case where it can be understood without any technicalities: when $M \subset \mathbb{R}^n$ and $K(\xi) = \frac{m^2}{2} |\xi|^2$ the ordinary inner product.

In that case, the formula above for $\gamma''(t)$ makes sense, but it doesn’t give a tangent vector to $M$. We define (this is the Levi-Civita connection in this case)

$$D_{\gamma'(t)}\gamma'(t) = \text{the orthogonal projection of } \gamma''(t) \text{ onto } T_{\gamma(t)}M.$$ 

This corresponds to the “high-school” component of the force in the direction of motion, and I will consider the differential equation

$$mD_{\gamma'(t)}\gamma'(t) = -\nabla_K V(\gamma(t))$$

to be the “known” approach to mechanics.

### 1.2 The Lagrangian approach

Define the Lagrangian function $L : TM \to \mathbb{R}$ to be

$$L(q, \dot{q}) = K(\dot{q}) - V(q).$$

Let $P_{(a,b)}(M)$ be the space of $C^2$ curves $\gamma : [0, 1] \to M$ with $\gamma(0) = a$, $\gamma(1) = b$. Then the principle of least action says:

A path $\gamma \in P_{(a,b)}(M)$ is a trajectory of classical mechanics if and only if it is a critical points of the action $A : P_{(a,b)} \to \mathbb{R}$ given by

$$A(\gamma) = \int_0^1 L(\gamma(t), \gamma'(t)) dt.$$ 

Note the special case where $V = 0$; in that case a “least action curve” is a critical point of the energy, well known to be the geodesics on $(M, K)$.

We need to transform the principle of least action into a differential equation: this is a standard topic in the calculus of variations.

To be rigorous about “critical points” of functions on “infinite dimensional manifolds like $P_{(a,b)}(M)$, we need to be a bit more specific about just what they are. Let $P^k_{(a,b)}(M)$ be the space of maps $\gamma : [0, 1] \to M$ with $\gamma(0) = a$ and $\gamma(1) = b$ of class $C^k$. It is a standard result of “global analysis” that $P^k_{(a,b)}(M)$ is a Banach manifold, and its tangent space $T_{\gamma}P^k_{(a,b)}(M)$ at $\gamma \in P^k_{(a,b)}(M)$ is the space of maps $\delta : C^k([0, 1], TM)$ such that

$$T_{\gamma}P^k_{(a,b)}(M) = \{ \delta \in C^k([0, 1], TM) \mid \delta(t) \in T_{\gamma(t)}M, \delta(0) = \delta(1) = 0 \}.$$
Thus it makes sense ask whether $A$ is differentiable, and if so what its derivative is. Note that its derivative should be a linear functional on $T_T^*\mathbb{P}^k(M)$. Again “global analysis” says that when $k \geq 1$ the function $A$ is differentiable, and

$$[DA(\gamma)]\delta = \int_0^1 \left( \frac{\partial L}{\partial q} (\gamma(t), \gamma'(t)) \delta(t) + \frac{\partial L}{\partial \dot{q}} (\gamma(t), \gamma'(t)) \dot{\delta}'(t) \right) dt = 0$$

for all $\delta \in P(a,b)(M)$.

This computation definitely makes sense in coordinates. If we have chosen coordinates $q_1, \ldots, q_m$ on $\mathbb{M}$ so that

$$(q, \dot{q}) \mapsto (q_1, \ldots, q_m, \dot{q}_1 \partial/\partial q_1, \ldots, \dot{q}_m \partial/\partial q_m)$$

identifies an open subset $\mathbb{R}^{2m}$ to an open subset of $TT\mathbb{M}$. In these coordinates, $DL$ is a line matrix of length $2m$ whose entries are functions of $q, \dot{q}$; the first $m$ are $\partial L/\partial q$ and the last $m$ are $\partial L/\partial \dot{q}$.

Thus

$$[DA(\gamma)]\delta = \int_0^1 \left( \frac{\partial L}{\partial q} (\gamma(t), \gamma'(t)) \delta(t) + \frac{\partial L}{\partial \dot{q}} (\gamma(t), \gamma'(t)) \dot{\delta}'(t) \right) dt = 0$$

for all $\delta \in P(a,b)(M)$.

In the standard way, we transform this using integration by parts to

$$\int_0^1 \left( \frac{\partial L}{\partial q} (\gamma(t), \gamma'(t)) - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} (\gamma(t), \gamma'(t)) \right) \delta(t) dt = 0$$

which implies that

$$\frac{\partial L}{\partial \dot{q}} (\gamma(t), \gamma'(t)) = \frac{d}{dt} \frac{\partial L}{\partial q} (\gamma(t), \gamma'(t)).$$

This is a second-order differential for $\gamma$, called the Euler-Lagrange equation.

### 1.3 The Hamiltonian approach

The Hamiltonian formalism applies in greater generality than our mechanical systems: it applies to a function on a symplectic manifold. Let $(N, \sigma)$ by a symplectic manifold, and $H : N \to \mathbb{R}$ be a function. The form $\sigma$ is a non-degenerate closed 2-form on $N$; in particular it induces an isomorphism between the tangent space and the cotangent space at every point, and we can specify a vector-field on $N$ by the formula

$$\sigma(\xi, \nabla_\sigma H) = dH(\xi)$$

for all tangent vectors $\xi$ to $N$, called the symplectic gradient of $H$. 

3
The differential equation of Hamiltonian mechanics is
\[ \dot{x} = \nabla_\sigma H(x). \]

In our case, the symplectic manifold is the cotangent bundle \( T^\top M \). This, like all cotangent bundles, has a natural symplectic structure, defined as follows.

There is a natural 1-form \( \omega \) on \( T^\top M \):

if \((q, p) \in T^\top M\), i.e., \( q \in M \) and \( p \in T_q^\top M \), and if \( \xi \in T_{(q,p)}T^\top M \), then
\[ \omega(\xi) = p(\pi_\ast \xi) \]
where \( \pi : T^\top M \to M \) is the natural projection, and \( \pi_\ast \) is its derivative.
Then \( \sigma = d\omega \).

Suppose \( q_1, \ldots, q_m \) are local coordinates on \( M \). Then \( dq_1, \ldots, dq_m \) are 1-forms on \( M \), i.e., sections of the cotangent bundles, so any point of \( T^\top M \) above this coordinate patch can be written
\[ \sum_i p_i dq_i, \]
where \( q_1, \ldots, q_m, p_1, \ldots, q_m \) are coordinate functions on this region.

Then
\[
\omega(q, p) \left( \sum_i \alpha_i \frac{\partial}{\partial q_i} + \sum_i \beta_i \frac{\partial}{\partial p_i} \right) = \sum_j p_j dq_j \left( \sum_i \alpha_i \frac{\partial}{\partial q_i} \right)
\]
\[ = \sum_j p_j \sum_i \alpha_i \frac{\partial}{\partial q_i} + \sum_i \beta_i \frac{\partial}{\partial p_i} \]
\[ = \sum_j p_j \alpha_j = \left( \sum_j p_j dq_j \right) \left( \sum_i \alpha_i \frac{\partial}{\partial q_i} + \sum_i \beta_i \frac{\partial}{\partial p_i} \right). \tag{1} \]

Thus
\[ \omega = \sum_i p_i dq_i, \quad \text{and} \quad \sigma = \sum_i dp_i \wedge dq_i. \]

Then the equation \( \dot{x} = \nabla_\sigma H(x) \) becomes, in these coordinates,
\[ \dot{q}_i = \frac{\partial H}{\partial p_i}, \quad i = 1, \ldots, m \]
\[ \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad i = 1, \ldots, m \tag{2} \]
Indeed,
\[
\sigma \left( \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}, \begin{bmatrix} \partial H/\partial p \\ -\partial H/\partial q \end{bmatrix} \right) = \left( \sum_i d\xi_i \wedge dq_i \right) \left( \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}, \begin{bmatrix} \partial H/\partial p \\ -\partial H/\partial q \end{bmatrix} \right) = \xi_2 \frac{\partial H}{\partial p} + \xi_1 \frac{\partial H}{\partial q} = dH \left( \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \right). \tag{3}
\]

In our mechanical system \((M, K, V)\), the function \(H\) is the total energy \(\frac{1}{2}q \cdot K \dot{q} + V(q)\).
However, this is a function on the tangent bundle \(TM\), not the cotangent bundle \(T^\top M\).
To view it as a function on \(T^\top M\) we need to use the isomorphism \(p = Kq\dot{q}\) induced by \(K\).
In our coordinates this gives
\[
\frac{1}{2}q \cdot Kq\dot{q} + V(q) = \frac{1}{2}(K^{-1}p) \cdot Kq(K^{-1}p) + V(q) = \frac{1}{2}p^\top K^{-1}p + V(q)
\]
Thus \(H(q, p) = \frac{1}{2}p^\top K^{-1}p + V(q)\).

2 Lagrangian mechanics is Newtonian mechanics
We can only check this when \(M\) is isometrically embedded in \(\mathbb{R}^n\). In that case
\[
[DA(\gamma)]\delta = \int_0^1 \left(\frac{m\dot{\gamma}'(t)}{\gamma'(t)} - [DV(\gamma(t))]\dot{\delta}(t)\right) dt.
\]
Saying that \(\gamma\) is a critical point of \(A\) is saying that its derivative vanishes, i.e., that
\[
[DA(\gamma)]\delta = 0 \quad \text{for all } \delta \in T_\gamma P_{(a,b)}^k(M).
\]
Using an integration by parts, this means that for all \(\delta \in T_\gamma P_{(a,b)}^k(M)\) we have
\[
0 = [DA(\gamma)]\delta = \int_0^1 \left(\frac{\delta'(t)}{\gamma'(t)} - [DV(\gamma(t))]\delta(t)\right) dt
= \int_0^1 \left(-\gamma''(t) - \nabla V(\gamma(t)) \cdot \delta(t)\right) dt, \tag{4}
\]
and that means that
\[
-\gamma''(t) - \nabla V(\gamma(t)
\]
is orthogonal to \(T_\gamma(t)M\) at every point. Thus the orthogonal projection of \(\gamma''(t)\) to \(T_\gamma(t)M\) is equal to \(-\nabla V(\gamma(t))\), as was to be shown.
3 Lagrangian mechanics is Hamiltonian mechanics

Recall the Euler-Lagrange equation:

$$\frac{\partial L}{\partial q}(\gamma(t), \gamma'(t)) = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}}(\gamma(t), \gamma'(t)).$$

We will turn this second order equation into a system of first order equations, in almost the standard way. We need the precise form of $L$:

$$L(q, p) = \frac{1}{2} \dot{q} \cdot K_q \dot{q} - V(q),$$

so that

$$\frac{\partial L}{\partial q} = \frac{1}{2} \dot{q} \cdot \frac{\partial K_q}{\partial q} \dot{q} - \frac{\partial V}{\partial q}$$

and

$$\frac{\partial L}{\partial \dot{q}} = K_q \dot{q}.$$ 

So we choose as our second variable $p = K_q \dot{q}$. Note that this is really the identification of $TM$ and $T^\top M$ induced by $K$. Now if $p(t) = K_{\gamma(t)} \gamma'(t)$ then the equation above leads to the first line of the following equations:

$$\frac{d}{dt} p = \frac{1}{2} \dot{q} \cdot \frac{\partial K_q}{\partial q} \dot{q} - \frac{\partial V}{\partial q}$$

$$= \frac{1}{2} (K_q^{-1} p)^\top \frac{\partial K_q}{\partial q} (K_q^{-1} p) - \frac{\partial V}{\partial q}$$

$$= \frac{1}{2} p^\top K_q^{-1} \frac{\partial K_q}{\partial q} K_q^{-1} p - \frac{\partial V}{\partial q}$$

$$= -\frac{1}{2} p^\top K_q^{-1} p - \frac{\partial V}{\partial q}$$

$$= -\frac{\partial H}{\partial q}. \quad (5)$$

This is the harder of Hamilton’s equations. For the other, recall that

$$H(q, p) = \frac{1}{2} p^\top K_q^{-1} p + V(q), \quad \text{so} \quad \frac{\partial H}{\partial p} = K_q^{-1} p.$$ 

So

$$\dot{q} = K_q^{-1} p = \frac{\partial H}{\partial p}.$$ 

Thus the critical points of $A$ are parametrized curves satisfying Hamilton’s equations.
4 The space pendulum

In this section we will set up the equations of motion for the space pendulum in all three formalisms and show that they give the same equations.

We will attach our pendulum, of length $l$ and mass $m$, at the origin in $\mathbb{R}^3$. Thus $M$ is the sphere of radius $l$ centered at the origin, and

$$K(q, \dot{q}) = \frac{1}{2} ml^2 |\dot{q}|^2.$$ 

We might want to study the pendulum in a constant gravitational field, with potential $V(q) = gmq^3$, or in space with $V = 0$. Of course the behavior of the latter is much more elaborate, but it isn’t much more difficult to set of the equations.

The set of positions can be parametrized by spherical coordinates; we will set them up so that the origin is the south pole

$$
\begin{align*}
x &= l \sin \phi \cos \theta \\
y &= l \sin \phi \sin \theta \\
z &= -l \cos \phi
\end{align*}
$$

and the potential energy

$$V(\phi, \theta) = -mgl \cos \phi.$$ 

Now we have our manifold with its Riemannian structure and its potential function, everything written in terms of local parameters.

4.1 The Hamiltonian approach

As usual, the Hamiltonian approach is the least intuitive and the easiest to put in practice. The Hamiltonian function is

$$H(\phi, \theta) = \frac{1}{2l^2m} \left( p_1^2 + \frac{p_2^2}{\sin^2 \phi} \right) - mgl \cos \phi$$

leading to

$$
\begin{align*}
\phi' &= \frac{\partial H}{\partial p_1} = \frac{p_1}{l^2 m} \\
\theta' &= \frac{\partial H}{\partial p_2} = \frac{p_2}{l^2 m \sin^2 \phi} \\
p_1' &= -\frac{\partial H}{\partial \phi} = \frac{p_2^2 \cos \phi}{l^2 m \sin^3 \phi} - mgl \sin \phi \\
p_2' &= -\frac{\partial H}{\partial \theta} = 0
\end{align*}
$$
The equation \( p_2' = 0 \) says that \( p_2 \) is a constant, which we call \( M \); the second equation of the left says \( M = l^2m \sin^2 \phi (\theta')^2 \), so it is the angular momentum. Differentiating the first equation on the left and substituting in the first equation on the right gives

\[
\phi'' = \frac{p_1'}{l^2m} = \frac{M^2 \cos \phi}{m^2 l^4 \sin^3 \phi} - \frac{g}{l} \sin \phi.
\]

### 4.2 The Euler-Lagrange approach

The Lagrangian is

\[
L \left( \left( \phi, \dot{\phi} \right), \left( \theta, \dot{\theta} \right) \right) = ml^2 \frac{1}{2} \left( \dot{\phi} \right) \cdot \begin{bmatrix} 1 & 0 \\ 0 & \sin^2 \phi \end{bmatrix} \left( \dot{\phi} \right) + mgl \cos \phi
\]

so the Euler-Lagrange equation

\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}
\]

becomes (after canceling \( ml^2 \))

\[
\frac{d}{dt} \begin{pmatrix} \phi'(t) \\ \sin^2 \phi(t) \theta'(t) \end{pmatrix} = \begin{pmatrix} \sin \phi \cos \phi (\theta'(t))^2 - \frac{g}{l} \sin \phi \\ 0 \end{pmatrix}.
\]

Again the second equation says that \( \sin^2 \phi \theta'(t) \) is constant, again we set \( ml^2 \sin^2 \phi(t) \theta'(t) = M \). Substituting this into the first equation gives

\[
\phi''(t) = \frac{\sin \phi \cos \phi M^2}{m^2 l^4 \sin^4 \phi} - \frac{g}{l} \sin \phi = \frac{M^2 \cos \phi}{m^2 l^4 \sin^3 \phi} - \frac{g}{l} \sin \phi,
\]

as it should to be consistent with the Hamiltonian approach.

### 4.3 The Newtonian approach

The computations are pretty painful: I haven’t finished them.