1 Introduction

In these notes we aim to do the following

1. Lay out the foundations for geometric formulations of Hamitonian and Lagrangian mechanics using symplectic geometry and covariant derivatives,
2. Explain why the cotangent bundle is the "canonical" choice for the symplectic formulation of Hamiltonian mechanics,
3. Show how the Legendre transformation relates the symplectic formulation of Lagrangian mechanics to the symplectic formulation of Hamiltonian mechanics.

2 Newtonian and Hamilton Mechanics

Our starting assumption is that a point particle at a given time t occupies a certain point in $\mathbb{R}^3$. We denote this position as $r(t) = (x^1(t), x^2(t), x^3(t)) \in \mathbb{R}^3$. Therefore if we are given N particles their collective coordinate can be written as $r(t) = (r^1(t), ..., r^N(t)) = (x^1(t), x^2(t), x^3(t), ..., x^{3N-2}(t), x^{3N-1}(t), x^{3N}(t)) \in \mathbb{R}^{3n}$. Newton‘s fundamental discovery was that (in today‘s language) such systems obey a second order differential equation:

$$m\frac{d^2r}{dt^2} = F(r(t), t)$$

where $F : \mathbb{R}^{3n+1} \to \mathbb{R}^{3n}$ is a vector valued quantity called the force field and m is the weight of the particles (in the case when particles are have different masses, m is replaced by a diagonal matrix whose elements represent masses of the particles). Therefore once the force field and some initial conditions $r(0), \dot{r}(0)$ are known, theoretically one can find the solutions of this system of equations. As with all higher order ordinary differential equations, we can turn this equation into a first order differential equation by introduction of an extra variable, say v. Then

$$\frac{dv}{dt} = \frac{1}{m}F(r(t), t) \qquad \frac{dr}{dt} = v(t) \tag{2.1}$$

so that this becomes a first order differential equation posed on $\mathbb{R}^{6n}$ and one can solve it if an initial condition $(r(0), v(0)) \in \mathbb{R}^{6n}$ is fixed. This space of extended variables is usually called a phase space. Although inside this phase space any weird curve of the form $(a(t), b(t))$ is allowed, due to the form of the ODE (2.1), we are interested in curves of the form $(a(t), \frac{da}{dt}(t))$. That is curves whose extended variables are precisely the tangent vectors of the curve in $\mathbb{R}^{3N}$. Such curves are usually denoted as dynamical curves and for a given force field F, the solutions of the ODE (2.1) lie inside this family. Therefore, the solutions $(r(t), v(t))$ of this ordinary differential equation are simply the integral curves of the vector-field $Z : \mathbb{R}^{3n+1} \to \mathbb{R}^{6n}$ given by $Z(t, r, v) = (v, F(r, t))$. This is the first important step in geometrizing the Newtonian mechanics.

As may be known, flows often come equipped with their conserved quantities or in physical terms their constants of motion, that allow to understand the qualitative and quantitative behaviour of these flows. In the case of Newtonian mechanics, it is an established fact that the total energy of an isolated system is a conserved quantity. That is what we are going to assume about our system. This in physical terms means that the force term $F(r,t)$ only appears due to certain interactions between the particles themselves and does not arise from any outside source. Even an external force is present, it is always possible to enlarge your system (which in the limit boils down to taking the whole universe as your system!) so that there are no external interactions. We will make an additional assumption that greatly simplifies the formulation of Hamiltonian mechanics. We will assume that F actually does not depend on time and can be given as the differential of some potential energy function $U : \mathbb{R}^{3n} \to \mathbb{R}$, that is $F = dU$. Such systems are usually called conservative in physical terms. Mathematically this means that F is divergence free and preserves volume. If such a U exists then the total energy of the system is defined as

$$H(r,v) = \frac{mv^2}{2} + U(r) \tag{2.2}$$

where v^2 denotes inner product of the vector v with itself. Note that the vector-field of the flow of the ODE above is given by

$$X(r(t), v(t)) = \frac{dr}{dt}\frac{\partial}{\partial r} + \frac{dv}{dt}\frac{\partial}{\partial v}.$$

So H being conserved means $X(H) = 0$ or

$$\frac{dr}{dt}\frac{\partial H}{\partial r} + \frac{dv}{dt}\frac{\partial H}{\partial v} = 0.$$

Using the fact that our solutions satisfy $\frac{dr}{dt} = v$, one gets:

$$v\frac{\partial H}{\partial r} + \frac{dv}{dt}\frac{\partial H}{\partial v} = 0.$$

A sufficient condition for this equation to hold true is

$$\frac{\partial H}{\partial r} = -c\frac{dv}{dt} \qquad \frac{\partial H}{\partial v} = c\frac{dr}{dt}. \tag{2.3}$$

for some constant $c > 0$. This derivation does not assume a specific form of H. In general we could call any smooth function H which is a constant of motion and which satisfies the relations (2.3) a Hamiltonian of the flow given by the vector-field (2.2). Note that if we are not given a vector-field but just a smooth function H, we could go the reverse direction and we could construct a vector field $(v, \frac{dv}{dt})$ using the relations (2.3) for which H becomes the Hamiltonian. That is any smooth function H generates a vector-field whose flows preserve H and H is a Hamiltonian of this flow. This generality will be explored more in detail later.

We leave it as an exercise to check that H given in (2.2) certainly satisfies these relations with our chosen vector-field (2.1) with $c = m$. So the "classical" Hamiltonian of a Newtonian system is given by (2.2). If we slightly change our coordinates so that $p = mv$ and write $H(r,p) = \frac{p^2}{2m} + U(r)$, the equations take the much more pleasant form

$$\frac{\partial H}{\partial r} = -\frac{dp}{dt} \qquad \frac{\partial H}{\partial p} = \frac{dr}{dt}.$$

We have actually discovered a remarkable fact:

$$X(r,p) = (\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial r}). \tag{2.4}$$

This relation between the conserved quantity H and the vector-field that generates the solutions of the Newtonian equation is known as the Hamiltonian mechanics. This vector-field is called the Hamiltonian vector-field associated to H. Note again that any smooth function H ,not necessarily of the form (2.2), can be used to construct a Hamiltonian vector-field by the rule (2.4). In general we denote such a vector-field by X_H.

Yet we are just starting to discovering amazing stuff about H. Note that the right hand-side of the equation almost looks like dH but some how reversed. To write this more concisely, consider the $6n \times 6n$ matrix ω of the form

$$\omega = \begin{pmatrix} 0_{3n\times 3n} & I_{3n\times 3n} \\ -I_{3n\times 3n} & 0_{3n\times 3n} \end{pmatrix}$$

Then one can write

$$\omega \circ X = dH.$$

In the language of differential forms dH is a differential 1-form while X is a vector field, therefore M is actually a map from vector fields to differential 1-forms. Indeed M is anti-symmetric matrix which is a differential 2-form. One can check that in coordinates (r,p) it is given by

$$\omega = \sum_{i=1}^{3n} dr^i \wedge dp^i. \tag{2.5}$$

With this point of view, the Hamilton's equations can be written as

$$X \lrcorner \omega = dH. \tag{2.6}$$

At this point we know that ω is a differential 2-form on $\mathbb{R}^{6n}$. It makes it a potential candidate for a differential 2-form on either the tangent bundle or the cotangent bundle of $\mathbb{R}^{3n}$. This can be checked by verifying whether if ω transforms according to coordinate changes on the tangent bundle or the cotangent bundle. Assume we are given a coordinate change $\phi : \mathbb{R}^{3n} \to \mathbb{R}^{3n}$. This induces the transformations $\phi_1 = (\phi, D\phi)$ and $\phi_2 = (\phi, D\phi^T)$ respectively on the tangent bundle and the cotangent bundle with their usual differentiable structure. These also induce transformation on differential 1-forms, 2-forms and vector fields defined on the tangent bundle and the cotangent bundle.

Exercise 2.1. Check that under the transformation

$$\tilde{r} = \phi(r) \qquad \tilde{p} = D(\phi^{-1})^T p,$$

one has that

$$\sum_{i=1}^{3n} dr^i \wedge dp^i = \sum_{i=1}^{3n} d\tilde{r}^i \wedge d\tilde{p}^i.$$

Also check that this is not the case for the transformation

$$\tilde{r} = \phi(r) \qquad \tilde{p} = D(\phi)p.$$

We remind that in physics a canonical transformation is any transformation of the coordinates (r, p) that preserves the form of the Hamilton's equations of motion that is:

$$(\frac{d\tilde{r}}{dt}, \frac{d\tilde{p}}{dt}) = (\frac{\partial H(\phi)}{\partial p}, -\frac{\partial H(\phi)}{\partial r}).$$

This exercise gives us the remarkable fact that the coordinate changes of the cotangent bundle are canonical transformations and also that if we were actually working on a manifold instead of $\mathbb{R}^n$ then the locally defined collection of differential 2-forms $\sum_{i=1}^{3n} dr^i \wedge dp^i$ match on the overlap since the transition functions (which are the coordinate change functions of the cotangent bundle described earlier) take one to another. This tells us that one can realize the 2-form ω as a differential 2-form on the cotangent bundle of a manifold. More so one sees that ω becomes a closed, non-degenerate 2-form which in literature is known as a symplectic 2-form.

Exercise 2.2. Check that ω is a closed, non-degenerate differential form.

So lets isolate the elements which has allowed us to geometrically write Newton's equations of motion: A function $H : \mathbb{R}^{2n} \to \mathbb{R}$, a non-degenerate closed two form ω. In this case if one considers the equation (2.6) for an unknown vector-field X, due to the fact that ω is non-degenerate, there exists a unique vector-field, which we denote as $X = X_H$, that satisfies this equation. X_H is known as the Hamiltonian vector-field associated to H and ω. We note that closedness of ω was not necessary in neither writing the equation (2.6) nor solving it for an unknown X_H. It was a by product of the form of the Newton's equations of motion and the fact that the potential function $V(x)$ depends only on x coordinate. However it does have a physical significance:

Exercise 2.3. Show that ω being closed is necessary and sufficient for $\mathcal{L}_{X_H}\omega = 0$ where $\mathcal{L}_{X_H}$ denotes Lie derivative with respect to X_H.

The result of this exercise tells us that if ω is closed then the phase space volume $\omega \wedge ... \wedge \omega$ (3n times) is preserved under the the flow of X_H. Since this fact is quite important both in classical and statistical physics, we also preserve this property.

So with these elements we can generalize the notion of Hamiltonian mechanics as follows: Let E be any manifold, ω a symplectic 2-form and $H : E \to \mathbb{R}$ any smooth function. Then then there exists a unique vector-field X_H which satisfies the equation (2.6) with $X = X_H$. X_H is called the Hamiltonian vector-field of H and preserves both H and ω, that is $\mathcal{L}_{X_H}(\omega) = 0$ and $\mathcal{L}_{X_H}(H) = 0$.

Of more physical importance is the case when $E = T^*M$ that is E is the cotangent bundle of some smooth manifold M and ω is the canonical symplectic 2-form on T^*M. If T^*M is given the local trivialization coordinates (r, p) then ω is precisely the 2-form that looks like (2.5) in these coordinates. Then given a $H : T^*M \to \mathbb{R}$, equation (2.6) describes the evolution of a particle's (or a system's) trajectory on the manifold M with Hamiltonian H. To be even more physical we can put a metric g on T^*M and impose that H is of the form

$$H(r, p) = \frac{g(p, p)}{2} + U(r)$$

for some smooth function $U : M \to \mathbb{R}$ which becomes the potential energy of the system. This then describes the trajectory of a particle subjected to a force $grad(V)$ on a manifold with metric g

We have shown that the Hamiltonian equations of motion can be described on $T * M$ using a symplectic form that naturally occurs with out any reference to H, physics, Newton or Hamilton. This is the reason why most people say that canonical choice for "symplectic mechanics" is $T * M$. This still leaves some to be desired. Indeed can we not formulate mechanics using the tangent bundle using some symplectic 2-form? We can, but we need some bit of extra information. In particular an isomorphism between the cotangent bundle and the tangent bundle called the Legendre transformation will allow us to define the mechanics using the bundle using and a symplectic 2-form on the tangent bundle. This isomorphism and the symplectic 2-form that we get will however depend on our choice of what is called a Lagrangian function. We will see that attempting to define a symplectic mechanics on tangent bundle in this form will result in what are called the Lagrange's equations of motion. Therefore before carrying out this task in section 4 we first discuss Lagrange's equations of motion in the next section.

3 A View at the Geometric Nature of Lagrange‘s Equations in the Classical Case

Unlike the Hamiltonian function H which is a physical quantity that describes the total energy of a system, there is not so much of a physical significance to the function L called the Lagrangian function (except that of course it will be used to derive the classical mechanics equations in another way). For a classical system evolving on a manifold of metric g subjected to a potential V, the Lagrangian is given by

$$L(x, v) = \frac{m}{2} g(v, v) - V(x). \tag{3.1}$$

In the simplest case where the space and the metric is Euclidean it is given by

$$L(x,v) = \frac{mv^2}{2} - V(x).$$

One can check that this quantity is not conserved by the flow of the vector-field (2.1). However as before it can be used to generate this vector-field easily by noting

$$\frac{\partial L}{\partial x} = m\frac{dv}{dt} \qquad \frac{\partial L}{\partial v} = mv. \tag{3.2}$$

Unlike the Hamiltonian case this can not be turned into a symplectic formalism directly since if we made such an attempt the ω that we would get would not be anti-symmetric and in particular would not be a two form. It will be the job of the next section to turn this into a symplectic formalism. In this section we will show that actually the equations 3.2 make global sense on any manifold. More precisely we will show that these equations are just local manifestations of a geometric equality posed on any manifold. To do this note first that the equations (3.2) are equivalent to requiring

$$\frac{d}{dt}\frac{\partial L}{\partial v} = \frac{\partial L}{\partial x}.$$

So now lets derive the same equations on an arbitrary manifold with a metric g. Note that Newton‘s equations of motion with a potential $V(x)$ in Euclidean space were

$$\frac{dx^i}{dt} = \frac{dv^i}{dt}, \quad \frac{dv^i}{dt} = -\frac{\partial V}{\partial x^i}.$$

The first equation means that $v(t)$ is the curve of vectors that is tangent to the curve $x(t)$. In particular they are inside the tangent space. The second equation requires to take the derivative of a curve of tangent vectors $v(t)$ along $x(t)$. There are many ways to make sense of this in an arbitrary manifold including seeing $\frac{dv}{dt}$ as elements of TTM, which was in fact the view adopted in the previous section on Hamiltonian dynamics formalism. For the purposes of this section covariant derivative associated to a connection will lead to a neat generalization, and therefore we will employ that. So working on the base manifold M and we will replace the usual time derivatives of Newtonian mechanics by directional covariant derivatives of the Levi-Civita connection of g.

More precisely we replace $\frac{dv}{dt}$ by $\nabla_{x(t)}v(t)$ where ∇ is the covariant derivative associated to the Levi-Civita connection of the metric g. In this case then the Newtonian equations of motion become

$$\nabla_{x(t)}v(t) = -\sharp(dV), \tag{3.3}$$

where $dV = \frac{\partial V}{\partial x^i}dx^i$ and $\sharp : T^*M \to TM$ is the isomorphism of the cotangent space and the tangent space given by g which acts as:

$$\sharp(dx^j) = g^{ij}\frac{\partial}{\partial x^i}.$$

Then in coordinates this equation is:

$$(\frac{dv^i}{dt} + \Gamma^i_{k\ell} v^k v^\ell)\frac{\partial}{\partial x^i} = g^{ij}\frac{\partial V}{\partial x^j}\frac{\partial}{\partial x^i}.$$

where Γ^i_{jk} are the Christofel symbols of the Levi-Civita connection associated to g. Denoting the inverse of the metric by $(g^{-1})^{ij} = g_{ij}$ and looking at the above equality component wise, we get

$$g_{ji}(\frac{dv^i}{dt} + \Gamma^i_{k\ell} v^k v^\ell) = \frac{\partial V}{\partial x^j}.$$

This can be rearranged as

$$\frac{d(g_{ji}v^i)}{dt} - \frac{dg_{ji}}{dt}v^i + g_{ji}\Gamma^i_{k\ell} v^k v^\ell = \frac{\partial V}{\partial x^j}.$$

Writing the time derivative of the metric explicitly we get

$$\frac{d(g_{ji}v^i)}{dt} + (-\frac{\partial g_{j\ell}}{\partial x^s} + g_{ji}\Gamma^i_{s\ell})v^s v^\ell = \frac{\partial V}{\partial x^j}.$$

Note that the Christofel symbols have the following form:

$$g_{ji}\Gamma^i_{s\ell} = \frac{1}{2}(\frac{\partial g_{j\ell}}{\partial x^s} + \frac{\partial g_{sj}}{\partial x^\ell} - \frac{\partial g_{s\ell}}{\partial x^j}).$$

Putting this into equation we get

$$\frac{d(g_{ji}v^i)}{dt} + \frac{1}{2}(\frac{\partial g_{sj}}{\partial x^\ell} - \frac{\partial g_{s\ell}}{\partial x^j} - \frac{\partial g_{j\ell}}{\partial x^s})v^s v^\ell = \frac{\partial V}{\partial x^j}.$$

This can again be rearranged to

$$\frac{d(g_{ji}v^i)}{dt} + \frac{1}{2}(\frac{\partial g_{sj}}{\partial x^\ell} - \frac{\partial g_{j\ell}}{\partial x^s})v^s v^\ell = \frac{\partial V}{\partial x^j} + \frac{1}{2}\frac{\partial g_{s\ell}}{\partial x^j}v^s v^\ell.$$

Since the metric is symmetric, that is $g_{j\ell} = g_{\ell j}$ we get that

$$(\frac{\partial g_{sj}}{\partial x^\ell} - \frac{\partial g_{j\ell}}{\partial x^s})v^s v^\ell = 0.$$

And so the final equation we get is

$$\frac{d(g_{ji}v^i)}{dt} = \frac{\partial V}{\partial x^j} + \frac{1}{2}\frac{\partial g_{s\ell}}{\partial x^j}v^s v^\ell.$$

But using the classical Lagrangian defined at (3.1), we see that in coordinates this equation are equivalent to what are called the Euler-lagrange equations that is

$$\frac{d}{dt}(\frac{\partial L}{\partial v^i}) = \frac{\partial L}{\partial x^i}. \tag{3.4}$$

Since our curves satisfy $\frac{dr}{dt} = v(t)$ it is easy to see that this expression is equivalent to Lagrangian equations of motion given in (3.2).

Therefore we have established that the local expression of the Lagrange's equations given in equation (3.2) are local manifestations of the geometric equation (3.3). Now in the next section we describe another way, using symplectic geometry on the tangent bundle, to describe Lagrangian formalism.

4 Legendre Transformation and Symplectic 2-forms on the Tangent Bundle

As in the case of the cotangent bundle, the most fundamental object for defining mechanics using the tangent bundle is a smooth function $L : TM \to \mathbb{R}$. Such a function allows us a transformation called the Legendre transformation.

Definition 4.1. Given a smooth manifold M with a smooth function $L : TM \to \mathbb{R}$, the Legendre transformation $\mathbb{L} : TM \to T^*M$ is defined by the following relation:

$$\mathbb{L}(x,v)(u) = (x, \frac{d}{dt}|_{t=0} L(x, v + tu))$$

$u, v \in T_xM$

Exercise 4.2. If TM is given the local coordinates (x, v) then the Legendre transformation in coordinates is given by

$$\mathbb{L}(x,v) = (x, \frac{\partial L}{\partial v}(x,v))$$

To see whether if this map gives a diffeomorphism between the tangent and cotangent bundle we can differentiate it to get

$$D\mathbb{L} =$$

$$\begin{pmatrix} Id_{n\times n} & 0_{n\times n} \\ (\frac{\partial^2 L}{\partial x \partial v})_{n\times n} & (\frac{\partial^2 L}{\partial v^2})_{n\times n} \end{pmatrix}_{2n\times 2n}. \tag{4.1}$$

We see that a sufficient condition for the Legendre transformation to be a diffeomorphism is that

$$det(\frac{\partial^2 L}{\partial v^2}) \neq 0.$$

Such Lagrangians are called regular and the isomorphism defined by such a Lagrangian is called Legendre duality. Now we will see that using this Legendre transformation, the Lagrangian equations of motion can be cast in the form of a symplectic geometry on TM. Note that geometrically the Hamiltonian equations of motion are

$$X_H \lrcorner \omega = dH,$$

where

$$\omega = \sum_i dx^i \wedge dp^i \qquad X_H = \frac{\partial H}{\partial p^i}\frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial p^i}.$$

Since the Lagrangian transformation defines a diffemorphism we have that

$$\omega_L = \mathbb{L}^*\omega$$

is a symplectic 2-form on TM.

Exercise 4.3. Check that in coordinates

$$\omega_L = \sum_{i,j} \frac{\partial^2 L}{\partial v^j \partial v^i} dx^i \wedge dv^j + \frac{\partial^2 L}{\partial x^j \partial v^i} dx^i \wedge dx^j$$

Now we will define a quantity called the energy that takes the place of the Hamiltonian in the symplectic formalism of mechanics on cotangent bundle. At the moment it might seem a bit out of the blue but we will later on show how this quantity is canonically related to the Hamiltonian through the Legendre transformation. Define

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

Exercise 4.4. Check that in local coordinates

$$E(x,v) = v^i \frac{\partial L}{\partial v^i} - L$$

Now the non-degeneracy condition assures that there exists a unique vector-field X_L on TM that satisfies

$$X_L \lrcorner \omega_L = dE.$$

Writing $X_L = \dot{x}_i \frac{\partial}{\partial x^i} + \dot{v}_i \frac{\partial}{\partial v^i}$ we have that

$$\omega_L(X_L, \cdot) = \dot{x}_i\left(\frac{\partial^2 L}{\partial v^j v^i} dv^j + \frac{\partial^2 L}{\partial x^j \partial v^i} dx^j - \frac{\partial^2 L}{\partial x^i \partial v^j} dx^j\right) - \dot{v}_i \frac{\partial^2 L}{\partial v^j \partial v^i} dx^j.$$

Equating this to dE we get the equations:

$$\dot{x}^i = v^i$$
$$\frac{\partial L}{\partial x^i} = \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right).$$

which are the Euler-Lagrange equations of motion which was to be proven. Using the Euler-Lagrange equations one can actually get a relation between dE and dH through the Legendre transformation.next exercise shows that E is actually given as the pullback of H by the Legendre transformation:

Exercise 4.5. Prove the following relation:

$$E = H \circ \mathbb{L}.$$

Thus we finally get that for L non-degenerate, the symplectic formalism of mechanics on the tangent bundle is given by

$$X_L \lrcorner (\mathbb{L}^* \omega) = d(H \circ \mathbb{L}) \tag{4.2}$$

which is equivalent to Lagrange equations of motion. Further this equations gives us the following:

$$(\mathbb{L}^{-1})^*(X_L \lrcorner \mathbb{L}^* \omega) = \mathbb{L}_* X_L \lrcorner \omega = dH,$$

therefore the Hamiltonian and Lagrangian vector-fields are related by

$$X_H = \mathbb{L}_* X_L.$$

On the passing we also want to point out the following neat relation:

Exercise 4.6. Given the following $L(x,v) = \frac{m}{2}g(v,v) - V(x)$, show that this Lagrangian is regular and that

$$\mathbb{L}(x,v)(\cdot) = (x, \frac{m}{2}g(v,\cdot)),$$

that is the Legendre transformation is simply the isomorphism given by the metric (up to a scalar factor).

Finally doing away with the Hamiltonian in equation (4.2) we can write the symplectic formalism of Lagrangian mechanics as

$$X_L \lrcorner (\mathbb{L}^*\omega) = d(A_L - L)$$

where A_L, sometimes called the action of L, is the function $TM \to \mathbb{R}$ defined by

$$A_L(x,v) = \mathbb{L}(x,v)(v).$$