TEX-Notes-I

Symplectic Geometry
on
Tangent and Cotangent Bundles

Larry K. Norris
Department of Mathematics
North Carolina State University
Spring Semester, 2003
Updated Fall, 2006

Contents

1 Background Material.

Before we begin our study of symplectic geometry we need to first review the essentials of the Lagrangian and Hamiltonian formulations of classical mechanics. This review will serve as a motivation for the introduction of symplectic geometry. Our initial goal here is to gain a broad overview of classical mechanics, and to see how the mathematical theory of symplectic geometry unifies and clarifies the classical picture.

Classical mechanics – a simple picture

The non-relativistic classical universe is composed of point masses, and composite systems made up of point masses. Each point mass is characterized by its mass m and the Cartesian coordinates (x, y, z) of the location of the point mass. The **configuration** of a system of N point masses $m_1, \ldots, m_N$ is given by the 3N Cartesian coordinates

$$(x_1, y_1, z_1), (x_2, y_2, z_2), \ldots, (x_N, y_N, z_N)$$

of the masses.

Often it is convenient to specify the configuration of the system in terms of other parameters, called **generalized coordinates**, instead of using Cartesian coordinates. For example it is often convenient to use spherical coordinates (ρ, θ, ϕ), cylindrical coordinates (r, θ, z), or any other coordinate system that is well-suited to the symmetry of the system. Moreover, when the system is subject to contraints so that the 3N Cartesian coordinates are not all independent, it is often necessary to use generalized coordinates. As we will see the proper way to model such configuration spaces is to consider them as differentiable manifolds.

A system is said to have **n degrees of freedom** if n is the least number of parameters necessary to specify the configuration of the system. The configuration of a system with n degrees of freedom is specified by n generalized position coordinates (q^i) , $i = 1, 2, \ldots, n$, and all quantities, including the Cartesian coordinates (x_A, y_A, z_A) , $A = 1, 2, \ldots, N$, are to be expressed in terms of the generalized coordinates. Thus the **kinetic energy T** of the system,

$$\begin{aligned} T &= (1/2) \sum_{A=1}^{N} m_A \delta_{\mu\nu} \frac{dx_A^\mu}{dt} \frac{dx_A^\nu}{dt} \\ &= (1/2) \sum_{A=1}^{N} m_A \left\{ \left(\frac{dx_A}{dt} \right)^2 + \left(\frac{dy_A}{dt} \right)^2 + \left(\frac{dz_A}{dt} \right)^2 \right\} \qquad (1) \end{aligned}$$

is transformed into a function

$$T(q^i, \dot{q}^i) \tag{2}$$

of the generalized coordinates. Similarly the **potential energy** V of the system becomes

$$V = V(q^i, \dot{q}^i) \tag{3}$$

In these equations we are using the standard notation $\dot{q}^i = \frac{dq^i}{dt}$ to denote differentiation with respect to time.

The **Lagrangian** of the system is defined by

$$L = T - V \ , \tag{4}$$

and the equations of motion of the system are the **Euler-Lagrange equations**

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0 \ , \ i = 1, 2, \ldots, n \ . \tag{5}$$

Alternatively, we may describe the dynamics of the system in the Hamiltonian formalism. From (refkenetic energy)-(4), we **define** the **generalized momenta** by

$$\begin{aligned} p_i &= \frac{\partial L}{\partial \dot{q}^i} \\ &= \frac{\partial}{\partial \dot{q}^i}(T - V) \ , \quad i = 1, 2, \ldots, n \end{aligned} \tag{6}$$

This definition is the generalization of the definition in Cartesian coordinates

$$p_{x_A} = m_A \dot{x}_A = \frac{\partial T}{\partial \dot{x}_A} \ , \quad A = 1, 2, \ldots, N \ , \tag{7}$$

for systems for which the potential energy V is velocity-independent and thus satisfies $\frac{\partial V}{\partial \dot{q}^i} = 0 \ , \ i = 1, 2, \ldots, n$.

We now **define** the **Hamiltonian H** of the system as the energy function $T + V$ expressed in terms of the $2n$ generalized position and momentum coordinates,

$$H = H(q^i, p_j) \ . \tag{8}$$

(More generally one constructs H using the Legendre transformation; see Section 4.)

In order to construct the Hamiltonian H from the data given in equations (1)-(4) we must be able to solve equations (6) for the p_i in terms of the q^i

and $\dot{q}^i$. The condition that guarantees that this may be done is that the so-called Hessian determinant associated with a Lagangian be non-zero. When this condition is satisfied by a particular Lagrangian we say the Lagrangian is **regular** (see section 4).

In terms of the Hamiltonian function H the equations of motion, **Hamilton's equations**, become

$$\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i} \\ i &= 1, 2, \ldots, n \end{aligned} \tag{9}$$

Hamilton's equations are a generalization of Newton's equations of dynamics, and when the coordinates are Cartesian equations (9) of Hamilton reduce to Newton's dynamical equations.

Hamilton's equations form a system of $2n$ ordinary differential equations of the first order for the $2n$ unknown functions $q^i(t), p_j(t)\ ,\ i, j = 1, 2, \ldots, n$. Given the initial values of these $2n$ functions at some instant t_0, the standard existence and uniqueness theorems in the theory of ordinary differential equations guarantee that Hamilton's equations possess a unique solution on some time interval containing t_0. That is to say, given the initial values of the generalized coordinates and momenta Hamilton's equations (9) completely specify the position and momentum coordinates at all other times $t \in (t_0, t_0 + c)$ for some constant $c > 0$.

In **classical mechanics** dynamical quantities such as energy, angular momentum, etc., are called dynamical **observables**, and when the Lagrangian is regular (or by fiat) classical observables are well-defined functions of the generalized coordinates. Thus if we are given the initial values of the dynamical observables at t_0, the subsequent behavior and properties of the system are completely determined by Hamilton's equations. The fundamental laws of classical mechanics are therefore completely deterministic.

REMARKS:

• The above description of classical mechanics is clearly non-relativistic. We will shortly consider the transition to a relativistic formulation.

• Note that in the above description the "Lagrangian" formalism is de-emphasized while the "Hamiltonian" formalism is emphasized. This is because in the traditional "canonical quantization scheme" one needs to start the quantization procedure with a Hamiltonian rather than a Lagrangian. Of course, if one has a regular Lagrangian then one can always construct a corresponding Hamiltonian, and this "regularity" condition is assumed above in passing from the Lagrangian (4) to the Hamiltonian (8). There are, however, interesting and important cases (usually associated with constraints) where the Lagrangian is not regular, and the study of such non-regular systems leads to the theory of Dirac brackets and constrained systems. In this course we will mainly avoid such systems and as far as is possible develop the theory for regular Lagrangians . There are two reasons for restricting attention to regular Lagrangian systems as much as possible. The first reason, of course, is that the theory is much simpler for regular Lagrangians, and in order to gain the broad overview we are after one wants to restrict to the simpliest situation for a first look. The second reason is that when all Lagrangians are regular, then we may immediately go over to the Hamiltonian picture for all cases. As we shall see, this corresponds to the fact that all regular Lagrangians (on, say, the tangent bundle TM of configuration space M) define **the same** symplectic structure (on the cotangent bundle T^*M). This is one of the most fundamental aspects of the symplectic geometry approach to classical mechanics in that it brings order and clarity to the entire subject. We will return to this important point later.

From the simple picture of classical mechanics presented above we can extract some very **general principles** upon which the theory is founded. From the Lagrangian point of view the dynamics of a classical system is determined, via the Lagrange equations, once the Lagrangian of the system is given . The Lagrangian $L = T - V$ for, say, N point masses always includes the same **kinetic energy** term T (equation (1)), and so different systems of N point masses will differ only in the **potential energy** term V (equation (3)) assumed for the systems. If we assume, as is usually done, that L depends only the coordinates and velocities, and that these quantities are free to vary **independently**, then the proper geometrical arena for Lagrangian mechanics is the 6N-dimensional tangent bundle TM of the 3N-dimensional configuration space M. The **velocity phase space** TM is the manifold whose points are pairs $(q, \vec{v})$ for $q \in M$ and $\vec{v}$ a tangent vector to M at q. Thus the Langrangian of a classical system should be thought of as a function

$$L : TM \longrightarrow \mathbb{R} \ . \tag{10}$$

The point I want to stress here is that classical Lagrangian mechanics can be considered geometrically as the study of functions of the type given in equation (10) together with the associated Lagrange equations on TM. We can, at least initially, forget about most of the explicit details of particular Lagrangians, and study the geometry on TM implied by the Lagrange equations. However, suppose one's goal is to study quantization. The traditional approach to quantization is "canonical quantization" which is based on Hamiltonian rather than Lagrangian mechanics. Thus rather than concentrating on the geometry associated with Lagrangian mechanics one would need to concentrate on the geometry associated with Hamiltonian mechanics.

larry

If the Lagrangian is regular then, as remarked above, the prescription given in equations (6)-(8) allows us to go over to the Hamiltonian formalism. In this picture the dynamics of a classical system is determined, via the Hamilton equations, once the Hamiltonian of the system is given . Hamiltonians (expression (8)) derived from regular Lagrangians depend on the coordinates q^i and the momenta p_i, and these quantities are assumed to be free to vary **independently**. Typically the conjugate momenta p_i correspond to **covariant vectors** whereas the velocities $\dot{q}^i$ correspond to **contravariant vectors**. The conclusion is that the proper geometrical arena for the Hamiltonian mechanics of an N-particle system is the 6N-dimensional **cotangent bundle** T^*M of the 3N-dimensional configuration space M. The **momentum phase space** T^*M is the manifold whose points are pairs (q, β) for $q \in M$ and β a covector at q. Classical Hamiltonians derived from regular Lagrangians therefore can be thought of as functions

$$H : T^*M \longrightarrow \mathbb{R} \,. \tag{11}$$

Thus classical Hamiltonian dynamics can be considered as the study of functions of the type given in (11) together with the geometry on T^*M implied by Hamilton's equations. This geometry, as we shall see, is **symplectic geometry**. Moreover, the symplectic geometry formulation of classical Hamiltonian dynamics is well-suited to study the traditional "canonical quantization" scheme, and such an approach leads to the theory of geometric quantization due to B. Kostant and J-M. Souriau.

To study the foundations of Hamiltonian mechanics we can temporarily ignore most of the explicit details of particular Hamiltonians and study the general geometrical features implied by the Hamilton equations. We can, moreover, forget about deriving Hamiltonians from regular Lagrangians and simply start with functions of the type given in (11). We will not be able to do this completely (e.g. Hamilton's principle requires us to start with a Lagrangian), but this will be our initial approach.

If we are going to work with arbitrary Hamiltonians on T^*M we will eventually need to select particular Hamiltonians in order to test and apply the theory. Thus we need to consider what types of Hamiltonian systems one should consider in both classical and quantum particle mechanics. Although there is no exhaustive list of allowable Hamiltonian systems, the following "short list" will suffice.

1. The Free particle in flat spacetime.

2. The single particle non-relativistic harmonic oscillator.

3. The Free particle in a curved spacetime (the gravitational interaction).

4. The charged particle in flat spacetime (the electromagnetic interaction).

5. The charged particle in Einstein-Maxwell spacetimes (combined gravitational and electromagnetic interaction).

These test cases are the fundamental classical models that one needs to consider. Presumably, more complicated models can be built up from these (and other) simple models.

Finally, we need to make explicit an assumption that is implicit in the structure of classical mechanics. We assume that the points of the classical universe can be modeled as the points of a **differentiable manifold**. This is the most general model that has been devised that also agrees well with classical experience.

EXERCISES:

1. The Hamiltonian function of a system is

$$H = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + V(q^1, q^2, q^3) \ .$$

Obtain Hamilton's equations and interpret them. In particular, if the coordinates (q^i) are Cartesian, how are Hamilton's equations related to Newton's equations of motion?

2. The motion of a point mass m in a central potential field $V(r)$ is described in spherical polar coordinates (r, θ, ϕ). If these are taken as generalized coordinates, define the corresponding generalized momenta, the Hamiltonian function, and derive Hamilton's equations.

3. Work out the Lagrangian and Hamiltonian descriptions of the 1-dimensional non-relativistic harmonic oscillator.

4. Find the Lagrangian for a non-relativistic charged particle moving under the influence of combined electric and magnetic fields $\vec{E}$ and $\vec{B}$. Obtain the Hamiltonian and show that the Hamilton's equations are equivalent to the Lorentz force law.

5. Show that the configuration space for the double pendulum (moving in a single vertical plane) is a torus.

2 Hamilton's Principle.

We have seen that in the Lagrangian formulation of classical mechanics one studies Lagrangians

$$\mathcal{L} : TM \longrightarrow \mathbb{R}$$

and the associated geometry and dynamics implied by the Lagrange equations. Moreover, the basic mathematical assumption is that configuration space M is a differentiable manifold. Let us be more precise about the underlying framework.

NOTATION:

$$\begin{aligned} M &= \textbf{configuration space} \\ &= \text{set of all allowed configurations} \\ &= \text{an n-dim differentiable manifold} \end{aligned} \tag{12}$$

$$\begin{aligned} \mathcal{A} &= \text{the complete } C^\infty \text{ atlas for } M \\ &= \{(U_\alpha, \mu_\alpha) \mid \alpha \in J\} \end{aligned} \tag{13}$$

where, for each $\alpha \in J$, J an index set,

$$\begin{aligned} &(i)\ U_\alpha \subset M \\ &(ii)\ M = \cup_{\alpha \in J} U_\alpha \\ &(iii)\ \mu_\alpha : U_\alpha \to \mu_\alpha(U_\alpha) \overset{open}{\subset} \mathbb{R}^n \ \text{ is a bijection} \\ &(iv) \text{ if } \ U_\alpha \cap U_\beta \neq \emptyset \text{ then the maps} \end{aligned} \tag{14}$$

$$\mu_\beta \circ \mu_\alpha^{-1} \quad \text{and} \quad \mu_\alpha \circ \mu_\beta^{-1}$$

are C^∞ Cartesian maps.

Each pair $(U_\alpha, \mu_\alpha) \in \mathcal{A}$ is a **local chart** for M, and each $x \in M$ is contained in at least 1 chart domain U_α. Given (U_α, μ_α) we define the associated $\mathbb{R}$-valued **coordinate functions** x^i_α , $i = 1, 2, \ldots, n$ by

$$\mu_\alpha(x) = (x^1_\alpha(x), x^2_\alpha(x), \ldots, x^n_\alpha(x)) \ .$$

Thus $x^i_\alpha = \text{proj}^i \circ \mu_\alpha$ where proj^i is the projection onto the i^{th} coordinate in $\mathbb{R}^n$. Generally we will drop the subscript labeling which local chart and write simply (U, μ) for a chart and (x^i) for the coordinate functions.

$$
\begin{aligned}
TM &= \{(x,\vec{t})|x \in M\ , \vec{t} \in T_x M\ \} \\
&= \textbf{velocity phase space} \\
&= \text{set of kinematically possible states of motion} \\
&= \text{the 2n-dim } \textbf{tangent bundle} \text{ of } M
\end{aligned} \tag{15}
$$

The tangent bundle TM has a projection map $\pi : TM \to M$ which is defined by

$$\pi(x,\vec{t}) = x \quad .$$

The local coordinates (x^i) are generalized coordinates for the configuration space M. We use these coordinates to define **standard velocity phase space coordinates** (q^i, v^j) on $TM \xrightarrow{\pi} M$ as follows.

Let U be the chart domain and set $\hat{U} = \pi^{-1}(U) \subset TM$. Define coordinate functions (q^i, v^j) , $i, j = 1, 2, \ldots, n$ by

$$
\begin{aligned}
q^i(x,\vec{t}) &= x^i \circ \pi(x,\vec{t}) \\
&= x^i(x)
\end{aligned} \tag{16}
$$

$$v^j(x,\vec{t}) = dx^j(\vec{t}) \tag{17}$$

for each point $(x,\vec{t}) \in \hat{U} \subset TM$. Thus the coordinates q^i are the coordinates x^i pulled up to TM and assign local coordinates to the first factor in $(x,\vec{t})$; the functions v^j assign as coordinates of the second factor $\vec{t}$ in $(x,\vec{t})$ the components of $\vec{t}$ with respect to the coordinated linear frame field $(\frac{\partial}{\partial x^i})$. For such standard coordinates on TM the functions q^i are the **generalized coordinates** while the funcitons v^j are **generalized velocities**.

Now that we have the basic mathematical model $TM \xrightarrow{\pi} M$, if TM is the set of **kinematically possible** states of motion, we next want to know how to determine the **dynamically possible** states of motion for a classical system with configuration space M.

Hamilton's Principle: The dynamical behavior of a classical system is completely determined by the Lagrangian $\mathcal{L} : TM \to \mathbb{R}$ ($\mathcal{L} \in C^\infty(TM)$ and $\mathcal{L} = T - V$ for conservative systems.) The dynamical trajectories are the solutions of the variational equation $\delta I[\gamma] = 0$ where $I[\gamma]$ is the functional

$$I[\gamma] = \int_{t_1}^{t_2} \mathcal{L}(q^i, \frac{dq^i}{dt})dt$$

where the variation is over all curves $\gamma : [t_1, t_2] \to M$ with fixed endpoints $\gamma(t_1)$, $\gamma(t_2)$.

As is well-known the variational equations are equivalent to the Euler-Lagrange equations

$$\frac{dq^i}{dt} = v^i \tag{18}$$

$$\frac{d}{dt}(\frac{\partial \mathcal{L}}{\partial v^i}) - \frac{\partial \mathcal{L}}{\partial q^i} = 0 \tag{19}$$

REMARK: In the statement of Hamilton's principle the notation $\mathcal{L}(q^i, \frac{dq^i}{dt})$ is shorthand notation for

$$\mathcal{L}(q^i, v^i) \circ \tilde{\gamma} \equiv \tilde{\gamma}^*(\mathcal{L}) \quad ,$$

where the curve

$$t \to \tilde{\gamma}(t) \equiv (\gamma(t), \dot{\gamma}(t))$$

is the **lift** of the curve $t \to \gamma(t)$ on M to TM. Similarly, the equations (18) and (19) are equations **on the tangent bundle** TM. They are often abbreviated as

$$\frac{d}{dt}(\frac{\partial \mathcal{L}}{\partial \dot{q}^i}) - \frac{\partial \mathcal{L}}{\partial q^i} = 0 \ ,$$

but we prefer equations (18) and (19) since they emphasize that the domain is TM. After solving equations (18) and (19) on TM for a curve $\gamma(t)$, the dynamical trajectory on M is then given by $\pi \circ \gamma(t) = (x^i(t))$.

REMARK: Note that the equations (18) and (19) depend explicitly on the choice of coordinates (q^i, v^j). We now indicate how we may remedy this situation. The goal is to find an invariantly defined vector field $X_\mathcal{L}$ on TM such that the differential equations for the integral curves of $X_\mathcal{L}$ are precisely the Lagrange equations (18) and (19) in local coordinates. Our discussion will initially be a local one based on the coordinates (q^i, v^j). At the end of the discussion we will indicate briefly how one obtains the invariant global generalization.

2.1 The Euler-Lagrange Equations Are Covariant

Since the Euler-Lagrange equations are derived from a coordinate independent variational principle, the equations themselves must be coordinate independent, which of course means "covariant" in the language of tensors. We show how this works by considering two special cases.

1. We suppose the Lagrangian is given in Cartesian coordinates (y^i) in $\mathbf{R}^3$ by

$$\mathcal{L} = \frac{1}{2} m g_{ij} \dot{y}^i \dot{y}^j - V(y) \tag{20}$$

   where here $g_{ij} = \delta_{ij} = \text{diag}(1,1,1)$ are the components of Euclidean metric tensor in rectangular coordinates. In this case the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}^i}\right) = \frac{\partial \mathcal{L}}{\partial y^i} \tag{21}$$

   reduce to

$$\frac{d}{dt}\left(m\delta_{ij}\dot{y}^j\right) = -\frac{\partial V}{\partial y^i} \tag{22}$$

   In this "conservative system" the vector force on the particle is $\vec{F} = F^i \frac{\partial}{\partial y^i} = -\delta^{ij} \frac{\partial V}{\partial y^j} \frac{\partial}{\partial y^j}$. So using the inverse metric δ^{ij} we may rewrite equation (22) as

$$\frac{d}{dt}\left(m\dot{y}^j\right) = F^i \tag{23}$$

   or, in vector notation as

$$m\vec{a} = \vec{F} \tag{24}$$

   Hence we have the result that the Euler-Lagrange equations, when applied to the Lagrangian of a conservative system written in rectangular coordinates, reproduces Newton's second law of motion relative to the rectangular coordinate system that one identifies with the inertial frame defined by the distant stars.

2. We suppose the Lagrangian is given in an arbitrary curvilinear coordinate system (x^i) in $\mathbf{R}^3$ by

$$\mathcal{L} = \frac{1}{2} m g_{ij} \dot{x}^i \dot{x}^j - V(x) \tag{25}$$

where here g_{ij} are the NON-CONSTANT components of the Euclidean metric tensor in the curvilinear coordinates. Since the metric tensor transforms as a second rank tensor field, the components g_{ij} are given by

$$g_{ij} = \delta_{ab} \frac{\partial y^a}{\partial x^i} \frac{\partial y^b}{\partial x^j} \tag{26}$$

In this case the Euler-Lagrange equations (21) reduce to

$$\frac{d}{dt} \left(m g_{ij} \dot{x}^j \right) = -\frac{\partial V}{\partial x^i} + \frac{1}{2} \frac{\partial g_{ab}}{\partial x^i} \dot{x}^a \dot{x}^b \tag{27}$$

Substituting $\frac{d}{dt}(g_{ij}) = \dot{x}^k \frac{\partial g_{ij}}{\partial x^k}$ for the directional derivative of the components of the metric tensor in this last equation we obtain

$$m g_{ij} \frac{d}{dt} \left(\dot{x}^j \right) + m \dot{x}^k \frac{\partial g_{ij}}{\partial x^k} \dot{x}^j = -\frac{\partial V}{\partial x^i} + \frac{m}{2} \frac{\partial g_{ab}}{\partial x^i} \dot{x}^a \dot{x}^b \tag{28}$$

Consider only the terms involving the derivatives of the metric tensor. Arranging them both on the left hand side of the equation we have the expression

$$m \left(\dot{x}^k \frac{\partial g_{ij}}{\partial x^k} \dot{x}^j - \frac{1}{2} \frac{\partial g_{ab}}{\partial x^i} \dot{x}^a \dot{x}^b \right) \tag{29}$$

Next relabel the velocity vectors in the first term to agree with the labeling scheme in the second term, and factor out $\frac{1}{2} \dot{x}^a \dot{x}^b$:

$$\frac{m}{2} \left(2 \frac{\partial g_{ib}}{\partial x^a} - \frac{\partial g_{ab}}{\partial x^i} \right) \dot{x}^a \dot{x}^b \tag{30}$$

Notice that the factor $\dot{x}^a \dot{x}^b$ is symmetric in the indices a and b. Hence the terms multiplying it must also be symmetric in a and b, since for any object A_{ab} we have

$$\begin{aligned}
2\dot{x}^a \dot{x}^b A_{ab} &= \dot{x}^a \dot{x}^b A_{ab} + \dot{x}^a \dot{x}^b A_{ab} \\
&= \dot{x}^a \dot{x}^b A_{ab} + \dot{x}^b \dot{x}^a A_{ba} \quad \boxed{\text{relabel indices in 2nd term}} \\
&= \dot{x}^a \dot{x}^b A_{ab} + \dot{x}^a \dot{x}^b A_{ba} \quad \boxed{\text{use symmetry of } \dot{x}^a \dot{x}^b} \\
&= \dot{x}^a \dot{x}^b \left(A_{ab} + A_{ba} \right) \quad \boxed{\text{factor out } \dot{x}^a \dot{x}^b}
\end{aligned}$$

Applying this formula to the term $2 \frac{\partial g_{ib}}{\partial x^a} \dot{x}^a \dot{x}^b$ in (30) above we may rewrite (30) as

$$\frac{m}{2} \left(\frac{\partial g_{ib}}{\partial x^a} + \frac{\partial g_{ia}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^i} \right) \dot{x}^a \dot{x}^b \tag{31}$$

Solving the definition

$$\Gamma^j_{ab} = \frac{1}{2} g^{ji} \left(\frac{\partial g_{ib}}{\partial x^a} + \frac{\partial g_{ai}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^i} \right)$$

for the terms in parentheses on the right hand side we find

$$\frac{1}{2} \left(\frac{\partial g_{ib}}{\partial x^a} + \frac{\partial g_{ai}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^i} \right) = g_{ij} \Gamma^j_{ab}$$

Substituting this result into the expression (31) we find that those terms may be rewritten as

$$\frac{m}{2} \left(\frac{\partial g_{ib}}{\partial x^a} + \frac{\partial g_{ai}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^i} \right) \dot{x}^a \dot{x}^b = m g_{ij} \Gamma^j_{ab} \dot{x}^a \dot{x}^b \tag{32}$$

Finally substituting this expression for the two terms involving the derivatives of the metric tensor components in (28) above we arrive at the equation

$$m g_{ij} \frac{d}{dt} \left(\dot{x}^j \right) + m g_{ij} \Gamma^j_{ab} \dot{x}^a \dot{x}^b = -\frac{\partial V}{\partial x^i} \tag{33}$$

Factoring out $m g_{ij}$ on the left hand side of this equation we find

$$m g_{ij} \left(\frac{d}{dt} \dot{x}^j + \Gamma^j_{ab} \dot{x}^a \dot{x}^b \right) = -\frac{\partial V}{\partial x^i}$$

Using the contravariant form of the metric tensor to eliminate the factor g_{ij} on the left we find this equation can be put into the form

$$m \left(\frac{d}{dt} \dot{x}^j + \Gamma^j_{ab} \dot{x}^a \dot{x}^b \right) = -g^{ji} \frac{\partial V}{\partial x^i}$$

We recognize this equation as Newton's second law of motion

$$m\vec{a} = \vec{F}$$

where the components of the acceleration vector in the curvilinear coordinates x^i are given by

$$a^j = \frac{d}{dt} \dot{x}^j + \Gamma^j_{ab} \dot{x}^a \dot{x}^b$$

and the components of the force $\vec{F}$ are $F^j = -g^{ji} \frac{\partial V}{\partial x^i}$.

These same equations would hold on a curved manifold with a non-flat metric tensor with components g_{ij} in an arbitrary coordinate system. The difference between the two situations is that the curvature tensor of the flat metric tensor will vanish identically, but the curvature tensor of the non-flat spacetime metric tensor will be non-zero.

2.2 Symplectic geometry on TM defined by a regular Lagrangian

Thus consider the velocity phase space TM of a configuration space M for a classical system with Lagrangian $\mathcal{L} = T - V : TM \to \mathbb{R}$. Let (q^i, v^j) be generalized coordinates on a subset $\hat{U}$ of TM. Define the 1-form $\theta_{\mathcal{L}}$ locally by

$$\begin{aligned} \theta_{\mathcal{L}} &= p_i dq^i \\ p_i &= \frac{\partial \mathcal{L}}{\partial v^i} \end{aligned} \tag{34}$$

Then the 2-form $\omega_{\mathcal{L}} = d\theta_{\mathcal{L}}$ has the local coordinate expression

$$\omega_{\mathcal{L}} = dp_i \wedge dq^i \tag{35}$$

We seek the conditions that guarantee that $\omega_{\mathcal{L}}$ is **non-degenerate** in the sense that

$$X \lrcorner\, \omega_{\mathcal{L}} = 0 \Longleftrightarrow X = 0 \tag{36}$$

for X a vector field on TM. The "hook" symbol is defined for a 2-form ω by the equation

$$(X \lrcorner\, \omega)(Y) = 2\omega(X, Y) \tag{37}$$

for vector fields X, Y.

Now if the functions (q^i, p_j) actually define a coordinate system on $\hat{U}$ then clearly $\omega_{\mathcal{L}}$ would be non-degenerate. For if (q^i, p_j) are coordinates then any vector field X can be expanded as

$$X = X^i \frac{\partial}{\partial q^i} + X_i \frac{\partial}{\partial p^i} \tag{38}$$

where the functions X^i and X_i are the components of the vector field relative to the coordinate system (q^i, p_j). Then from equations (35) and (38) we find

$$X \lrcorner\, \omega_{\mathcal{L}} = X_i dq^i - X^i dp_i \ . \tag{39}$$

Then $\omega_{\mathcal{L}} \lrcorner\, X = 0 \ \Rightarrow X^i = X_i = 0$ for all $i = 1, 2, \ldots, n$, which implies that $X = 0$.

Hence, if (q^i, p_j) form a coordinate system on $\hat{U}$ then $\omega_{\mathcal{L}}$ defined in equation (35) will be non-degenerate. Given coordinates (q^i, v^j) and a Lagrangian $\mathcal{L}$, the functions (q^i, p_j), where the p_j are defined in (34), will defined a coordinate system on $\hat{U}$ if the Jacobian of the coordinate transformations

$$(q^i, v^j) \longrightarrow (q^i, p_j)$$

is non-singular. The Jacobian of this transformation is

$$\begin{aligned} J &= \frac{\partial(q^i, p_j)}{\partial(q^a, v^b)} \\ &= \begin{pmatrix} \frac{\partial q^i}{\partial q^a} & \frac{\partial p_j}{\partial q^a} \\ \frac{\partial q^i}{\partial v^b} & \frac{\partial p_j}{\partial v^b} \end{pmatrix} \\ &= \begin{pmatrix} \delta^i_a & \frac{\partial p_j}{\partial q^a} \\ 0 & \frac{\partial p_j}{\partial v^b} \end{pmatrix} \\ &= \begin{pmatrix} \delta^i_a & \frac{\partial^2 \mathcal{L}}{\partial q^a \partial v^j} \\ 0 & \frac{\partial^2 \mathcal{L}}{\partial v^a \partial v^b} \end{pmatrix} \end{aligned} \tag{40}$$

Therefore J is non-singular if and only if

$$\det\left(\frac{\partial^2 \mathcal{L}}{\partial v^a \partial v^b}\right) \neq 0 \ . \tag{41}$$

Definition # 1: A Lagrangian $\mathcal{L} : TM \to \mathbb{R}$ is **regular** if it satisfies equation (41).

We assume that $\mathcal{L}$ is regular. Thus $\mathcal{L}$ defines the forms $\theta_{\mathcal{L}}$ and $\omega_{\mathcal{L}}$ and a new coordinate system (q^i, p_j). The coordinate p_i is referred to as the **momentum** coordinate **canonically conjugate** to q^i.

Next we define the energy function $h_{\mathcal{L}}$ by

$$h_{\mathcal{L}} = v^i \frac{\partial \mathcal{L}}{\partial v^i} - \mathcal{L} = v^i p_i - \mathcal{L} \tag{42}$$

Then the exterior derivative of $h_{\mathcal{L}}$ reduces to

$$dh_{\mathcal{L}} = v^i dp_i - \left(\frac{\partial \mathcal{L}}{\partial q^i}\right) dq^i \ . \tag{43}$$

We now have all the necessary pieces of the puzzle. Since $\omega_{\mathcal{L}}$ is non-degenerate it gives a 1-1 correspondence between vector fields and 1-forms on $\hat{U} \subset TM$. Thus if X is a vector field on TM then we may define a unique 1-form $\tilde{X}$ by the equation

$$\tilde{X} = -X \lrcorner\, \omega_{\mathcal{L}} \ . \tag{44}$$

Conversely, if λ is a 1-form on TM then we obtain a unique vector field $\tilde{\lambda}$ from the equation

$$\lambda = -\tilde{\lambda} \lrcorner\ \omega_{\mathcal{L}}\ . \tag{45}$$

We may, if we wish, think of $\omega_{\mathcal{L}}$ as a sort of "skew-symmetric" metric tensor field, although we will not emphasize this interpretation.

In particular, given the 1-form field $dh_{\mathcal{L}}$ we define a unique vector field $X_{\mathcal{L}}$ by the equation

$$dh_{\mathcal{L}} = -X_{\mathcal{L}} \lrcorner\ \omega_{\mathcal{L}}\ . \tag{46}$$

At the end of this section we will **derive** this equation from Hamilton's principle.

Express $X_{\mathcal{L}}$ in local coordinates as

$$X_{\mathcal{L}} = X^i \frac{\partial}{\partial q^i} + X_i \frac{\partial}{\partial p_i}$$

and expand both sides of equation (refx sub L defined) to obtain

$$v^i dp_i - \frac{\partial \mathcal{L}}{\partial q^i} dq^i = -X_i dq^i + X^i dp_i \tag{47}$$

Equating coefficients of the independent 1-forms we obtain the pair of equations

$$X^i = v^i \tag{48}$$

$$X_i = \frac{\partial \mathcal{L}}{\partial q^i} \tag{49}$$

Thus the vector field $X_{\mathcal{L}}$ determined by equation (46) is

$$X_{\mathcal{L}} = v^i \frac{\partial}{\partial q^i} + \frac{\partial \mathcal{L}}{\partial q^i} \frac{\partial}{\partial p_i} \tag{50}$$

expressed in the new coordinate system (q^i, p_j).

Now consider the integral curves of the vector field $X_{\mathcal{L}}$. Along an integral curve $\gamma(t)$ we have

$$\begin{aligned} X_{\mathcal{L}}(\gamma(t)) &= \gamma_*(t) \\ &= \dot{q}^i \frac{\partial}{\partial q^i} + \dot{p}_i \frac{\partial}{\partial p_i} \end{aligned} \tag{51}$$

Evaluating (refexplicit form of x sub L) along $\gamma(t)$ and equating the result with equation (refaaa) we obtain the differential equations

$$\dot{q}^i = v^i \tag{52}$$

$$\dot{p}_i = \frac{\partial \mathcal{L}}{\partial q^i} \tag{53}$$

Since $p_i = \frac{\partial \mathcal{L}}{\partial v^i}$ we see that these equations are the Euler-Lagrange equations of motion (18) and (19).

SUMMARY: If $\mathcal{L}$ is a regular Lagrangian on TM then one may define locally a closed, non-degenerate 2-form $\omega_{\mathcal{L}}$ and an energy function $h_{\mathcal{L}}$. This data then determines a unique vector field $X_{\mathcal{L}}$ on TM via the equation

$$dh_{\mathcal{L}} = -X_{\mathcal{L}} \rfloor\, \omega_{\mathcal{L}} \ , \tag{54}$$

and the differential equations for the integral curves of $X_{\mathcal{L}}$ are the Lagrange equations.

REMARK: Our construction of $\theta_{\mathcal{L}}$, $\omega_{\mathcal{L}}$, $h_{\mathcal{L}}$ and $X_{\mathcal{L}}$ depends explicitly on the coordinates (q^i, v^j). However, the structure of the tangent bundle TM of a manifold is such that it allows an **invariant** definition of all of these quantities so that equation (54) actually defines $X_{\mathcal{L}}$ globally and invariantly on TM if the Lagrangian is regular. The structure of equation (54) is highly geometrical and is intrinsically tied to the differentiable structure of M and TM. Note, however, that if we fix M and then change from one system with Lagrangian $\mathcal{L}_1$ to another system with Lagrangian $\mathcal{L}_2$, then the basic structure of the equations change since the Lagrangians will define distinct 2-forms $\omega_{\mathcal{L}_1}$ and $\omega_{\mathcal{L}_2}$.

Our goal in the next few sections will be to show that if we reformulate everything on the cotangent bundle T^*M, using the Legendre transformation, then the geometrical structure simplifies considerably in the sense that all regular Lagrangians define the **same** fundamental 2-form, the **canonical symplectic** 2-form, on T^*M.

To derive equation (54) we first reformulate Hamilton's principle as follows. With the Lagrangian $\mathcal{L} : TM \to \mathbb{R}$ given, the action functional is

$$I[\gamma] = \int_0^1 \mathcal{L}(q^i, \frac{dq^i}{dt})dt = \int_0^1 \tilde{\gamma}^*(\mathcal{L})dt \tag{55}$$

where the curve

$$t \to \tilde{\gamma}(t) = (\gamma(t), \dot{\gamma}(t)) \tag{56}$$

is the **lift** of a curve $t \to \gamma(t)$ on M to TM. Note also that for convenience we have normalized the limits of the integral in (55).

We define a **variation of** $t \to \gamma(t)$ **with fixed endpoints** to be a 1-parameter family of smooth curves

$$s \to \gamma_s(t) \tag{57}$$

satisfying the additional conditions

$$\gamma_s(0) = \gamma(0) \ , \quad \gamma_s(1) = \gamma(1) \tag{58}$$

for all appropriate values of the parameter s. We call the original curve $t \to \gamma(t)$ the **base curve**. Associated with each such variation is a vector field $X(t)$ defined along $\gamma(t)$ by

$$X(t) = \frac{\partial}{\partial s}(\gamma_s(t))|_{s=0} \quad . \tag{59}$$

REMARK: Note that the individual members $\gamma_s(t)$ of each such 1-parameter family of curves are **points** in an infinite dimensional manifold, namely the manifold of all smooth paths on M. Given the "point" $t \to \gamma(t)$, a 1-parameter family of curves (57) in M satisfying (58) is a "curve through the point $\gamma(t)$". The vector field $X(t)$ along $\gamma(t)$ defined in (59) above can then be interpreted as a "tangent vector at the point $t \to \gamma(t)$" in the infinite dimensional space. Ideas of this type are central in the study of path and loop spaces, loop groups, etc.. The topology on the infinite dimensional path space is relatively complicated and we will not discuss it here.

Note that by lifting a 1-parameter family of curves $\gamma_s(t)$ to TM according to (56), we may then define a vector field $\tilde{X}(t)$ along the lifted base curve $\tilde{\gamma}(t)$ by

$$\tilde{X}(t) = \frac{\partial}{\partial s}(\tilde{\gamma}_s(t))|_{s=0} \quad . \tag{60}$$

EXERCISE: Show that in standard coordinates (q^i, v^j) on TM that

$$\tilde{X}(t) = X^i(t)\frac{\partial}{\partial q^i} + \dot{X}^i(t)\frac{\partial}{\partial v^i} \quad . \tag{61}$$

where $X(t)$ is defined in (59).

REMARK: The fact that the vertical components of the vector field $\tilde{X}(t)$ are the time derivatives of the "horizontal components" $X^i(t)$ in (61) is the analog of the formula

$$\delta\dot{q}^i = \frac{d}{dt}(\delta q^i)$$

that occurs in classical derivations of Lagrange's equations.

We now define the (first) variation of the action functional given in (55) by

$$\delta I[\gamma] = \frac{d}{ds}I[\gamma_s]|_{s=0} = \int_0^1 \frac{\partial}{\partial s}(\tilde{\gamma}_s^*(\mathcal{L}))|_{s=0}dt \tag{62}$$

EXERCISE: With the definitions as above, show that

$$\frac{\partial}{\partial s}(\tilde{\gamma}_s^*(\mathcal{L}))|_{s=0} = \tilde{\gamma}^*(\tilde{X}(\mathcal{L})) \quad . \tag{63}$$

Using the result of this exercise in the first variation (62) we obtain

$$\delta I = \int_0^1 \tilde{\gamma}^*(\tilde{X}(\mathcal{L}))dt = \int_{\tilde{\gamma}} \tilde{X}(\mathcal{L})dt \quad . \tag{64}$$

Thus Hamilton's principle can be restated as: A curve $t \rightarrow \gamma(t)$ in M is a dynamical trajectory iff

$$\int_{\tilde{\gamma}} \tilde{X}(\mathcal{L})dt = 0 \tag{65}$$

where $\tilde{X}(t) =< X(t), \dot{X}(t) >$ is any vector field along $\tilde{\gamma}(t) = (\gamma(t), \dot{\gamma}(t))$ such that $X(0) = 0$ and $X(1) = 0$.

To derive the Langrange equations in the form (54) we need to reformulate the integrand in (65). The crucial point is that although we require, for example, $X(0) = 0$, we **do not** know anything about $\dot{X}(0)$. The only restriction is that the endpoints be fixed. To see what to do, not that

$$\begin{aligned} \tilde{X}(\mathcal{L}) &= \tilde{X}\lrcorner\, d\mathcal{L} \\ &= \tilde{X}\lrcorner\, (\frac{\partial \mathcal{L}}{\partial q^i}dq^i + \frac{\partial \mathcal{L}}{\partial v^i}dv^i) \end{aligned} \tag{66}$$

We need a way to eliminate the term involving dv^i, since if we have

$$\int_{\tilde{\gamma}} \tilde{X}\lrcorner\, (\text{a horizontal 1-form})dt = 0 \quad ,$$

then we may conclude that the horizontal 1-form must vanish since its coefficients would be the arbitrary $X^i(t)$. What we need is an invariant version of "integration by parts" that occurs at this step in the classical variational principle. This brings us to the natural geometrical structures carried by the tangent bundle of a manifold. General references: Crampin, M. and Thompson,G., *Affine bundles and integrable almost tangent structures*, Math. Proc. Camb. Phil. Soc. (1985), **98**, pp 61-71; Woodhouse, N. *Geometric Quantization*, Oxford University Press, Oxford (1980).

We define a type (1,1) tensor field S on TM as follows. We consider S as a **vector-valued 1-form** and define it by giving its values on vectors tangent to TM. Thus if X is a tangent vector to TM at (p, w) then

$$S(X) := \frac{d}{dt}(p, w + td\pi(X))|_{t=0} \quad . \tag{67}$$

EXERCISE: Show that in standard coordinates (q^i, v^j) on TM that S has the local form

$$S = \frac{\partial}{\partial v^i} \otimes dq^i \quad . \tag{68}$$

The significance of this canonically defined tensor field on TM is the following. For each function $f : TM \to \mathbb{R}$ we use S to assign a unique 1-form θ_f defined by

$$Y \lrcorner\, \theta_f = S(Y) \lrcorner\, df \tag{69}$$

for all tangent vectors Y on TM.

EXERCISE: Show that in standard coordinates (q^i, v^j) on TM that θ_f has the explicit form

$$\theta_f = \frac{\partial f}{\partial v^i} dq^i \quad . \tag{70}$$

LEMMA: A non-zero tangent vector $X \in T_{(p,w)}(TM)$ is vertical iff $S(X) = 0$.
PROOF: Left as an exercise.

We also need the **Euler vector field** on TM which is defined as follows. First we define the **vertical lift** of a tangent vector $Y \in T_pM$ by the formula

$$Y^v(p, w) := \frac{d}{dt}(p, w + tY)|_{t=0} \quad . \tag{71}$$

DEFINITION: The Euler vector field $\mathcal{E}$ on TM is the vector field defined at $(p, Y) \in TM$ by

$$\mathcal{E}(p, Y) = Y^v(p, Y) \quad . \tag{72}$$

EXERCISE: Show that in local coordinates (q^i, v^j) on TM that the Euler vector field has the expression

$$\mathcal{E} = v^i \frac{\partial}{\partial v^i} \quad . \tag{73}$$

We use S and $\mathcal{E}$ in various ways. One such use is to characterize curves on TM that are lifts of curves on M.

LEMMA: A curve $t \to \sigma(t)$ in TM is the lift of a curve $t \to \gamma(t)$ on M iff

$$S(\sigma_*(t)) = \mathcal{E}(\sigma(t)) \quad . \tag{74}$$

Proof: Left as an exercise.

DEFINITION: For each function $f : TM \to \mathbb{R}$ define the associated **Hamiltonian function** h_f by

$$h_f := \mathcal{E}(f) - f \quad . \tag{75}$$

The following lemma is the crucial one needed to complete the derivation of equations (54).
LEMMA: If X is a vector field on TM such that $S(X) = \mathcal{E}$, then for each $f : TM \to \mathbb{R}$ the 1-form

$$dh_f + X \lrcorner\ d\theta_f \tag{76}$$

is horizontal.

Proof: See Woodhouse, *Geometric Quantization*, pp. 18.

REMARK: By the above lemma typical vector fields satisfying $S(X) = \mathcal{E}$ are tangents to lifted curves.

We are now ready to derive (54) from Hamilton's principle, which we recall now has the form:

A curve $t \to \gamma(t)$ in M is a dynamical trajectory iff

$$\int_{\tilde{\gamma}} \tilde{X}(\mathcal{L}) dt = 0 \tag{77}$$

where $\tilde{X}(t) =< X(t), \dot{X}(t) >$, and $X(t)$ is any vector field along $\gamma(t)$ such that $X(0) = 0$ and $X(1) = 0$.

In order to carry out the "integration by parts" **on TM** we first extend $\tilde{X}$ and $\tilde{T} \equiv \tilde{\gamma}_*$, originally only defined along $\tilde{\gamma}(t)$, off the curve to vector fields $\tilde{X}$ and $\tilde{T}$ in such a way that

$$[\tilde{X}, \tilde{T}] = 0 \quad , \quad S(\tilde{T}) = \mathcal{E} \quad . \tag{78}$$

We leave it as an exercise to show that this can be done. Note that for the vector field $\tilde{T}$ we may simply take the tangent vectors to the curves $s \rightarrow \gamma_s(t)$ defining a variation of the original curve.

Now consider the integrand in (65). We have

$$\begin{aligned} \tilde{X}(\mathcal{L}) &= \tilde{X} \lrcorner\, d\mathcal{L} \\ &= \tilde{X} \lrcorner\, d(\mathcal{E}(\mathcal{L}) - h_{\mathcal{L}}) \\ &= \tilde{X} \lrcorner\, d(\mathcal{E}(\mathcal{L})) - \tilde{X} \lrcorner\, dh_{\mathcal{L}} \end{aligned} \tag{79}$$

Using (69) and (78) we have

$$d(\mathcal{E}(\mathcal{L})) = d(\mathcal{E} \lrcorner\, d\mathcal{L}) = d(S(\tilde{T}) \lrcorner\, d\mathcal{L}) = d(\tilde{T} \lrcorner\, \theta_{\mathcal{L}}) \ . \tag{80}$$

Using this back in (79) we obtain

$$\tilde{X}(\mathcal{L}) = \tilde{X} \lrcorner\, d(\tilde{T} \lrcorner\, \theta_{\mathcal{L}}) - \tilde{X} \lrcorner\, dh_{\mathcal{L}} \quad . \tag{81}$$

By definition we have the identity

$$L_{\tilde{T}} \theta_{\mathcal{L}} = \tilde{T} \lrcorner\, d\theta_{\mathcal{L}} + d(\tilde{T} \lrcorner\, \theta_{\mathcal{L}}) \quad . \tag{82}$$

Using this to rewrite the first term on the right hand side in (81) we have

$$\begin{aligned} \tilde{X} \lrcorner\, d(\tilde{T} \lrcorner\, \theta_{\mathcal{L}}) &= \tilde{X} \lrcorner\, (L_{\tilde{T}} \theta_{\mathcal{L}}) - \tilde{X} \lrcorner\, (\tilde{T} \lrcorner\, d\theta_{\mathcal{L}}) \\ &= L_{\tilde{T}}(\tilde{X} \lrcorner\, \theta_{\mathcal{L}}) + [\tilde{X}, \tilde{T}] \lrcorner\, \theta_{\mathcal{L}} - \tilde{X} \lrcorner\, (\tilde{T} \lrcorner\, d\theta_{\mathcal{L}}) \\ &= \tilde{T}(\tilde{X} \lrcorner\, \theta_{\mathcal{L}}) - \tilde{X} \lrcorner\, (\tilde{T} \lrcorner\, d\theta_{\mathcal{L}}) \end{aligned} \tag{83}$$

We have used the identity (136) in Section 6 in going from line 1 to line 2 in equation (83). Substituting back into (81) we get

$$\tilde{X}(\mathcal{L}) = \tilde{T}(\tilde{X} \lrcorner\, \theta_{\mathcal{L}}) - \tilde{X} \lrcorner\, \left(\tilde{T} \lrcorner\, d\theta_{\mathcal{L}} + dh_{\mathcal{L}} \right) \quad . \tag{84}$$

Substituting into (77) from (84) we obtain

$$\int_{\tilde\gamma} \tilde X(\mathcal{L})dt = 0$$

$$\Downarrow$$

$$0 = \int_{\tilde\gamma} (\tilde T(\tilde X \lrcorner\ \theta_{\mathcal{L}})dt - \int_{\tilde\gamma} \tilde X \lrcorner\ \left(\tilde T \lrcorner\ d\theta_{\mathcal{L}} + dh_{\mathcal{L}}\right) dt\ . \tag{85}$$

For the first integral we get

$$\begin{aligned} \int_{\tilde\gamma} (\tilde T(\tilde X \lrcorner\ \theta_{\mathcal{L}})dt &= \int_0^1 \frac{d}{dt}(\tilde X \lrcorner\ \theta_{\mathcal{L}})dt \\ &= (\tilde X \lrcorner\ \theta_{\mathcal{L}})|_0^1 \\ &= 0 \end{aligned} \tag{86}$$

since $\theta_{\mathcal{L}}$ is horizontal and the horizontal parts of $\tilde X$ vanish at the endpoints. Moreover, the 1-form "hooked with X" in the remaining integral is horizontal by (48). We conclude that a curve $t \to \gamma(t)$ is a dynamical trajectory iff

$$\begin{array}{ll} (a) & S(\tilde T) = \mathcal{E}\ , \\ (b) & dh_{\mathcal{L}} + \tilde T \lrcorner\ d\theta_{\mathcal{L}} = 0 \end{array} \tag{87}$$

These equations are equivalent to equations (54) plus the condition that the dynamical trajectories are the integral curves of $X_{\mathcal{L}}$.

SUMMARY: The tangent bundle TM of a manifold M supports the canonically defined (1,1) tensor field S and the canonically defined Euler vector field $\mathcal{E}$. For each Lagrangian $f : TM \to \mathbb{R}$ we

A. first define the action potential 1-form θ_f by the formula

$$X \lrcorner\ \theta_f = S(X) \lrcorner\ df \tag{88}$$

for all vector fields X on TM. We then use θ_f to

B define the associated Lagrangian vector field X_f by the formula

$$d(\mathcal{E}(f) - f) = -X_f \lrcorner\ d\theta_f\ \ . \tag{89}$$

C The vector field X_f will be unique if f is regular. Finally,

D the equations for the integral curves of the vector field X_f are the Lagrange equations for the Lagrangian f.

3 The Cotangent Bundle of a Manifold

Before we can introduce the Legendre transformation we need some basic facts about the structure of the **cotangent bundle** T^*M of an n-dim differentiable manifold M. We suppose that M is the configuration space of some classical system.

$$
\begin{aligned}
T^*M &= \{(x,\lambda) \mid x \in M \;,\; \lambda \in T^*_x M\} \\
&= \textbf{momentum phase space} \\
&= \textbf{set of all kinematically possible states of motion} \\
&= \text{a 2n-dim differentiable manifold}
\end{aligned}
$$

The projection map $\pi : T^*M \to M$ is defined by

$$\pi(x,\lambda) = x \;.$$

3.1 Standard coordinates on T^*M

Let (U,μ) be a chart on M; $\mu = (x^i)$, $i = 1, 2, \ldots, n$. Define coordinates (q^i, p_j) on $\hat{U} = \pi^{-1}(U) \subset T^*M$ by:

$$
\begin{aligned}
q^i(x,\lambda) &= x^i \circ \pi(x,\lambda) = x^i(x) \\
p_j(x,\lambda) &= \lambda(\frac{\partial}{\partial x^j})
\end{aligned} \tag{90}
$$

Thus the functions q^i are essentially the same as the q^i defined on TM (except that the domains are different !), while the functions p_j assign the components of the covector λ, with respect to the basis of 1-forms dx^i, as the coordinates of the second factor in (x,λ). Although it is not obvious from (90), we will see shortly that coordinates on T^*M defined in this way are **canonical coordinates** for many, but not all, classical systems with M as configuration space.

The structure of the vector bundle $T^*M \xrightarrow{\pi} M$ is intimately connected with the structure of M. This follows from the fact that a complete atlas for T^*M is standardly built up from local charts defined as in (90), which in turn are based on the complete atlas $\mathcal{A}$ for M. The structure of the two manifolds M and T^*M implies the existence of a certain globally defined 1-form θ, the **canonical** 1-form, on T^*M. We define θ by defining its values on arbitrary vectors tangent to T^*M. If $X \in T_{(x,\lambda)}(T^*M)$, then

$$\boxed{\theta(X) := \lambda(d\pi(X))} \;. \tag{91}$$

Another useful way of writing this equation is

$$X \lrcorner\, \theta = (d\pi X) \lrcorner\, \lambda \ . \tag{92}$$

REMARK: The definition (91) should be compared with the definition of the soldering 1-form on the bundle of linear frames LM of M. We will see later in this course that the soldering 1-form allows one to build up a generalized symplectic geometry on LM.

DEFINITION #2: A vector $X \in T_{(x,\lambda)}(T^*M)$ is **vertical** if $d\pi(X) = 0$.

DEFINITION #3: A 1-form μ on T^*M is **horizontal** if $\mu(X) = 0$ whenever X is vertical.

Let us now work out the local coordinate expression for θ in a standard coordinate chart (90). Note first that the basis vectors $\frac{\partial}{\partial p_j}$ are vertical, since

$$\begin{aligned} d\pi(\frac{\partial}{\partial p_j})(x^i) &= \frac{\partial}{\partial p_j}(x^i \circ \pi) && (93) \\ &= \frac{\partial}{\partial p_j}(q^i) && (94) \\ &= 0 && (95) \\ & && (96) \end{aligned}$$

Expressing θ in the chart (q^i, p_j) we have

$$\theta = \theta_i dq^i + \theta^i dp_i \ . \tag{97}$$

Then working at an arbitrary point $(x, \lambda) \in \hat{U}$ we find for the components θ_i

$$\begin{aligned} \theta_i &= \theta(\frac{\partial}{\partial q^i} \mid_{(x,\lambda)}) \\ &= \lambda(d\pi(\frac{\partial}{\partial q^i})) \\ &= \lambda(\frac{\partial}{\partial x^i}) \\ &= p_i(x, \lambda) && (98) \\ & && (99) \end{aligned}$$

where the last line follows from the definition (90). Similarly, for the components θ^i we find

$$\begin{aligned}
\theta^i &= \theta(\frac{\partial}{\partial p_i} \mid_{(x,\lambda)}) \\
&= \lambda(d\pi(\frac{\partial}{\partial p_i})) \\
&= 0
\end{aligned} \tag{100}$$

since the $\frac{\partial}{\partial p_i}$ are vertical. Thus in any local chart defined as in (90) we have

$$\theta = p_i dq^i \ . \tag{101}$$

DEFINITION #4: The **canonical symplectic 2-form** ω on T^*M is $\omega := d\theta$. In local coordinates (90) it has the coordinate expression

$$\boxed{\omega = dp_i \wedge dq^i} \ . \tag{102}$$

REMARK: It is clear from the definition that ω is **closed**. Moreover, from our earlier discussion it is also clear that ω is **non-degenerate**.

As we will see later when we use ω to write down Hamilton's equations on T^*M, it is the simple form (102) of ω when expressed in the coordinates (q^i, p_j) that implies the specific $\dot{q}^i = \frac{\partial \mathcal{H}}{\partial p_i}$ and $\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q^i}$ **form** of Hamilton's equations. This is the reason that the coordinates (q^i, p_j) are called canonical coordinates. Generalizing we have

DEFINITION #5: Coordinates (u^i, w_j) on T^*M are **canonical coordinates** if the symplectic 2-form ω has the form

$$\omega = dw_i \wedge du^i \tag{103}$$

when expressed in these coordinates.

REMARK: This definition leads to the question: What is the group of tranformations that transforms canonical coordinates to canonical coordinates? That is to say, what is the group that preserves the **form** of equation (103)? This group is the **symplectic group SP(n,$\mathbb{R}$)**, and although we will not discuss this group in detail now we can gain some insight into its structure by recalling the definition of Killing vector fields.

If g is a metric tensor field on a manifold M then a vector field ξ is a **Killing vector field** if

$$L_\xi(g) = 0 \ , \tag{104}$$

where the symbol L_ξ denotes Lie differentiation. Generally Killing vector fields do not exist, but when they do they express a certain symmetry of the particular metric tensor that admits the Killing vector field. In particular, in flat Minkowski spacetime the metric tensor admits a 10-parameter family of Killing vector fields. This family of Killing vectors forms a Lie algebra under the Lie bracket, the Poincaré Lie algebra $p(4) = o(1,3) \oplus \mathbb{R}^4$ of the Poincaré group $P(4) = O(1,3) \otimes \mathbb{R}^4$. As is well known the $P(4)$ transformations on Minkowski spacetime are precisely those that preserve the **form**

$$(\eta_{ij}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

of the metric tensor when expressed in a Lorentzian chart on M. Those transformation are composed of 3-space rotations, 4-space translations, and the so-called "boosts" relating observers moving with constant relative velocity.

As remarked above symplectic group transformations preserve the explicit form (103) of ω under change of coordinates. Thus one might expect that the symplectic group can be defined, at least locally, by requiring in analogy with (104) the existence of vector fields X such that

$$L_X(\omega) = 0 \ . \tag{105}$$

A vector field X that satisfies (105) is called a **locally Hamiltonian vector field**, and as we will see later the flow of X does indeed define a local 1-parameter group of local canonical transformations, and the set of all locally Hamiltonian vector fields forms an infinite dimensional Lie algebra which is fundamentally related to the Poisson brackets of Hamiltonian dynamics. We will return to these ideas, and in particular equation (105), in detail later.

4 The Legendre Transformation

Let M be the configuration space of a classical system with regular Lagrangian $\mathcal{L}$, TM the velocity phase space and T^*M the momentum phase space. We now want to use $\mathcal{L}$ to set up a map $\Lambda_{\mathcal{L}} : TM \to T^*M$, called the **Legendre transformation**, which will allow us to make the transition from the Lagrangian to the Hamiltonian formalism. We begin by considering some general ideas.

Suppose that $f : TM \to \mathbb{R}$ is any smooth function on velocity phase space. Then f can be used to define a map

$$\Lambda_f : TM \to T^*M$$

that is fibering preserving (i.e. $\Lambda_f(T_xM) \subset T^*_xM)$). We map each point $(x, X) \in TM$ to the point $(x, \beta_{(x,X)})$, and since $\beta_{(x,X)}$ is to be a covector at $x \in M$ we define it by prescribing its values on arbitrary vectors in T_xM. Thus

$$\Lambda_f((x, X)) := (x, \beta_{(x,X)}) \tag{106}$$

where $\beta_{(x,X)}$ is defined by

$$\beta_{(x,X)}(Y) = \frac{d}{dt}(f(x, X + tY)) \mid_{t=0} \quad \forall\, Y \in\ T_xM\ . \tag{107}$$

Recalling the definition (71) of the **vertical lift** Y^v of a vector $Y \in T_xM$ to TM, we see that $\beta_{(x,X)}$ can also be defined by the formula

$$\beta_{(x,X)}(Y) = Y^v(x, X) \lrcorner\ df \quad , \quad \forall\, Y \in\ T_xM\ . \tag{108}$$

Observe that Λ_f is invariantly defined and is independent of any choice of coordinates, but it depends explicitly on the choice of the function f.

Since Λ_f can be defined for any smooth $f : TM \to \mathbb{R}$ we can evaluate it for the specific choice $f = \mathcal{L}$, and in this case $\Lambda_{\mathcal{L}}$ is call the **Legendre transformation based on $\mathcal{L}$**. Thus

$$\Lambda_{\mathcal{L}}((x, X)) := (x, \beta_{(x,X)}) \tag{109}$$

where, from equation (107),

$$\beta_{(x,X)}(Y) = \frac{d}{dt}(\mathcal{L}(x, X + tY)) \mid_{t=0} \quad \forall\, Y \in\ T_xM\ . \tag{110}$$

We evaluate $\beta_{(x,X)}$ in local coordinates (q^i, v^j) on TM.

$$\begin{aligned}
\beta_{(x,X)}(Y) &= \frac{d}{dt}(\mathcal{L}(x, X+tY)) \mid_{t=0} \\
&= \{\frac{\partial \mathcal{L}}{\partial v^i}(x, X+tY)\frac{d}{dt}(v^i(x, X+tY))\} \mid_{t=0} \\
&= \frac{\partial \mathcal{L}}{\partial v^i}(x, X)Y^i) \\
&= \frac{\partial \mathcal{L}}{\partial v^i}(x, X)dx^i(Y)
\end{aligned} \tag{111}$$

Thus the covector $\beta_{(x,X)} \in T_x^*M$ has the coordinate expression

$$\beta_{(x,X)} = \frac{\partial \mathcal{L}}{\partial v^i}(x, X)dx^i \ . \tag{112}$$

This expression is often abbreviated as

$$\beta_{(x,X)} = \frac{\partial \mathcal{L}}{\partial v^i}dx^i \ , \tag{113}$$

but it is important to remember that the function $\frac{\partial \mathcal{L}}{\partial v^i}$ on the right hand side of (113) is evaluated at the point $(x, X) \in T_xM$.

EXERCISE: Show that $\Lambda_{\mathcal{L}}$ is a **local differomorphism** iff $\mathcal{L}$ is regular.

To see why regularity only guarantees a **local** diffeomorphism, consider the following 2-dim example:

$$\mathcal{L} = e^{v^1}\cos(v^2) \quad .$$

This Lagrangian is regular since

$$\left(\frac{\partial^2 \mathcal{L}}{\partial v^i \partial v^j}\right) = -e^{2v^1} \neq 0 \quad .$$

On the other hand the Legendre transformation for the momenta is

$$\begin{aligned}
p_1 \circ \Lambda_{\mathcal{L}} &= e^{v^1}\cos(v^2) \\
p_2 \circ \Lambda_{\mathcal{L}} &= -e^{v^1}\sin(v^2)
\end{aligned} \tag{114}$$

which is clearly many-to-one. However, we may restrict the domain in T^*M so that the functions in equation (114) are one-to-one, and on this domain $\lambda_{\mathcal{L}}$ will be a local diffeomorphism..

EXERCISE: If $\mathcal{L}$ is not regular, is $\Lambda_{\mathcal{L}}$ still well-defined?

In order to be able to distinguish between Lagrangians that induce Legendre transformations that are **diffeomorphisms**, as opposed to **local diffeomorphisms**, we introduced the following definition (cf. R. Abraham and J. Marsden, *Foundations of Mechanics*, second edition, 1978).

DEFINITION # 1-A: A Lagrangian $\mathcal{L} : TM \to \mathbb{R}$ is **hyper-regular** if it is regular and the Legendre transformation $\Lambda_{\mathcal{L}} : TM \to T^*M$ determined by $\mathcal{L}$ is a diffeomorphism.

To summarize, if $\mathcal{L}$ is a hyper-regular Lagrangian then $\Lambda_{\mathcal{L}}$ is a diffeomorphism, and in local coordinates (q^i, v^j) on TM and (q^i, p_j) on T^*M we have

$$\Lambda_{\mathcal{L}}(x, X) = (x, \beta_{(x,X)}) = (x, \frac{\partial \mathcal{L}}{\partial v^i}(x, X) dx^i) \ , \tag{115}$$

$$q^i(\Lambda_{\mathcal{L}}(x, X)) = q^i(x, \beta_{(x,X)}) = x^i(x) \ , \tag{116}$$

$$p_j(\Lambda_{\mathcal{L}}(x, X)) = \beta_{(x,X)}(\frac{\partial}{\partial x^j}) = \frac{\partial \mathcal{L}}{\partial v^j}(x, X) \ . \tag{117}$$

Therefore

$$(q^i, p_j) \circ \Lambda_{\mathcal{L}} = (q^i, \frac{\partial \mathcal{L}}{\partial v^j}) \ , \tag{118}$$

and we have recovered the classical coordinate form of the Legendre transformation.

A kinematically possible state of motion $(x, X) \in T_x M$, representing a velocity vector X at $x \in M$, is transformed into a kinematically possible state of motion $(x, \frac{\partial \mathcal{L}}{\partial v^j}(x, X) dx^j)$, which is the momentum covector at x determined by the Lagrangian and the velocity vector X.

Given a hyper-regular Lagrangian $\mathcal{L}$ the Legendre transformation is a fiber preserving diffeomorphism $\Lambda_{\mathcal{L}} : TM \to T^*M$. We can thus use $(\Lambda_{\mathcal{L}})_*$ to push vectors tangent to TM over to T^*M, and use $(\Lambda_{\mathcal{L}})^*$ to pull back forms on T^*M to forms on TM. In particular, we can pull back the canonical symplectic 2-form ω on T^*M to obtain a 2-form $(\Lambda_{\mathcal{L}})^*(\omega)$ on TM. We evaluate $(\Lambda_{\mathcal{L}})^*(\omega)$ in local coordinates (q^i, v^j) on TM and (q^i, p_j) on T^*M. (Note that we are using the **same symbols** q^i for half of the coordinates on both TM and T^*M although technically they are different functions.)

With ω expressed as in (102) we find

$$\begin{aligned}
(\Lambda_{\mathcal{L}})^*(\omega) &= (\Lambda_{\mathcal{L}})^*(dp_i \wedge dq^i) \\
&= d((\Lambda_{\mathcal{L}})^*(p_i)) \wedge d((\Lambda_{\mathcal{L}})^*(q^i)) \\
&= d(p_i \circ \Lambda_{\mathcal{L}}) \wedge dq^i \\
&= d(\frac{\partial \mathcal{L}}{\partial v^i}) \wedge dq^i
\end{aligned} \tag{119}$$

where the last line follows from (117). Then making use of equations (35) we have

$$\begin{aligned}
(\Lambda_{\mathcal{L}})^*(\omega) &= d(\frac{\partial \mathcal{L}}{\partial v^i}) \wedge dq^i \\
&= d(\frac{\partial \mathcal{L}}{\partial v^i} dq^i) \\
&= d(\theta_{\mathcal{L}}) \\
&= \omega_{\mathcal{L}}
\end{aligned} \tag{120}$$

REMARK: Suppose that $\mathcal{L} : TM \to \mathbb{R}$ is hyper-regular. What we have shown is that $(\Lambda_{\mathcal{L}})^*(\omega)$ is globally defined on TM, and it agrees with $\omega_{\mathcal{L}}$ defined in (35) in local coordinates on TM. If we therefore take

$$\omega_{\mathcal{L}} = (\Lambda_{\mathcal{L}})^*(\omega) \tag{121}$$

as the **definition** of $\omega_{\mathcal{L}}$, then

$$dh_{\mathcal{L}} = -X_{\mathcal{L}} \lrcorner\ \omega_{\mathcal{L}} \tag{122}$$

is equation (54) that was derived at the end of section 2. This equation serves to define $X_{\mathcal{L}}$ with $h_{\mathcal{L}}$ defined invariantly using the definition (75).

5 Hamiltonian Dynamics on Cotangent Bundles

Let M be the configuration space of a classical system with regular Lagrangian $\mathcal{L}$, TM the velocity phase space and T^*M the momentum phase space. In this section we use the Legendre transformation $\Lambda_{\mathcal{L}} : TM \to T^*M$, to make the transition from the Lagrangian to the Hamiltonian formalism.

In sections 2-4 we showed that the Lagrange equations can be characterized invariantly on TM as the differential equations for the integral curves of the vector field $X_{\mathcal{L}}$ defined by

$$dh_{\mathcal{L}} = -X_{\mathcal{L}} \lrcorner\, \omega_{\mathcal{L}} \quad , \tag{123}$$

where

$$\omega_{\mathcal{L}} = \Lambda^*_{\mathcal{L}}(\omega) \tag{124}$$

and

$$h_{\mathcal{L}} = A_{\mathcal{L}} - \mathcal{L} = \mathcal{E}(\mathcal{L}) - \mathcal{L} \ . \tag{125}$$

Our goal now is to reformulate these equations on T^*M. The key motivating idea is that if $\mathcal{L}_1$ and $\mathcal{L}_2$ are regular Lagrangians on TM and ω is the canonical symplectic form on T^*M, then as we have seen earlier

$$\omega_{\mathcal{L}_1} = (\Lambda_{\mathcal{L}_1})^*(\omega) \ \text{ and } \ \omega_{\mathcal{L}_2} = (\Lambda_{\mathcal{L}_2})^*(\omega) \tag{126}$$

Thus **each** symplectic form $\omega_{\mathcal{L}}$ on TM defined by a regular Lagrangian is the pull back of the canonical symplectic 2-form ω on T^*M under a Legendre transformation $\Lambda_{\mathcal{L}} : TM- > T^*M$. Before discussing Hamilton's equations explicitly we consider some general features of the symplectic geometry defined by ω.

Suppose that $f : T^*M \to \mathbb{R}$ is any smooth function on T^*M. Then using ω we can associate with the 1-form df a unique vector field on T^*M.

DEFINITION #1: The vector field X_f on T^*M defined by

$$df = -X_f \lrcorner\, \omega \tag{127}$$

is the (globally defined) **Hamiltonian vector field determined by** f.

REMARK: Recall that the " left-hook product" used in equation (127) is defined by

$$(X \lrcorner\, \omega)(Y) = 2\omega(X, Y) \tag{128}$$

for all vector fields X and Y.

REMARK: Equation (127) is the fundamental equation in symplectic geometry. Although it does not seem to have a name associated with it in the literature, I will refer to it as the **first structure equation** of symplectic geometry.

EXERCISE: Let (q^i, p_j) be canonical coordinates on T^*M defined as in equation (3.1). Show that X_f defined in equation (127) has the local coordinate expression

$$X_f = \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial}{\partial p_i} \quad . \tag{129}$$

EXERCISE: If X_f is the Hamiltonian vector field on T^*M determined by $f : T^*M \to \mathbb{R}$, show that in canonical coordinates (q^i, p_j) the differential equations for the integral curves of X_f are

$$\begin{aligned} \dot{q}^i &= \frac{\partial f}{\partial p_j} \\ \dot{p}_i &= -\frac{\partial f}{\partial q^i} \\ i &= 1, 2, \ldots, n \end{aligned} \tag{130}$$

REMARK: The reason we are considering arbitrary functions $f : T^*M \to \mathbb{R}$ is that we want to consider the dynamics associated with arbitrary classical observables (linear momentum, angular momentum, etc.) in addition to the dynamics generated by the Hamiltonian of a system. Note in particular that in canonical coordinates the differential equations (130) for the integral curves of any smooth $f : T^*M \to \mathbb{R}$ take the **form** of the classical Hamilton equations (9) with the function f playing the role of the Hamiltonian. The point I wish to emphasize here is that it is the 1^{st} structure equation (127) together with the assumption of canonical coordinates that lead to the specific differential equations (130).

DEFINITION #2: If $\mathcal{L} : TM \to \mathbb{R}$ is hyper-regular then the associated **Hamiltonian function** $H_{\mathcal{L}} : T^*M \to \mathbb{R}$ of the system is

$$H_{\mathcal{L}} = h_{\mathcal{L}} \circ \Lambda_{\mathcal{L}}^{-1} \ . \tag{131}$$

If the Lagrangian is regular then we replace $\Lambda_{\mathcal{L}}^{-1}$ in the above equation with $(\Lambda_{\mathcal{L}}|_U)^{-1}$ where U is an appropriate subset to ensure an inverse.

Usually one denotes the Hamiltonian function by the symbol H rather than by $H_{\mathcal{L}}$, and we will adopt this standard convention.

As a corollary to the above discussion we have that in canonical coordinates on T^*M the differential equations for the integral curves of the Hamiltonian vector field X_H defined by the $dH = -X_H \lrcorner\, \omega$ take the form

$$\begin{aligned} \dot{q}^i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q^i} \end{aligned} \tag{132}$$

which are the standard Hamilton's equations (9). We now show that if $\mathcal{L}$ is regular then $X_{\mathcal{L}}$ maps to X_H under the differential $d\Lambda_{\mathcal{L}}$ of the Legendre transformation. Let $h_{\mathcal{L}}$ be regular and define $X_{\mathcal{L}}$ by

$$dh_{\mathcal{L}} = -X_{\mathcal{L}} \lrcorner\, \omega_{\mathcal{L}} \quad . \tag{133}$$

(See equation (2.26)).

LEMMA #1: If $\mathcal{L} : TM \to \mathbb{R}$ is regular then $(d\Lambda_{\mathcal{L}})(X_{\mathcal{L}}) = X_H$.

Proof: Let Y be an arbitrary vector field on T^*M. Then

$$\begin{aligned} (d\Lambda_{\mathcal{L}}(X_{\mathcal{L}})) \lrcorner\, \omega)(Y) &= 2\omega(d\Lambda_{\mathcal{L}}(X_{\mathcal{L}}), Y) \\ &= 2\omega(d\Lambda_{\mathcal{L}}(X_{\mathcal{L}}), d\Lambda_{\mathcal{L}} \circ d\Lambda_{\mathcal{L}}^{-1}(Y)) \\ &= 2(\Lambda_{\mathcal{L}}^*(\omega))(X_{\mathcal{L}}, d\Lambda_{\mathcal{L}}^{-1}(Y)) \\ &= 2\omega_{\mathcal{L}}(X_{\mathcal{L}}, d\Lambda_{\mathcal{L}}^{-1}(Y)) \\ &= (X_{\mathcal{L}} \lrcorner\, \omega_{\mathcal{L}})(d\Lambda_{\mathcal{L}}^{-1}(Y)) \\ &= -dh_{\mathcal{L}}(d\Lambda_{\mathcal{L}}^{-1}(Y)) \\ &= -d(h_{\mathcal{L}} \circ \Lambda_{\mathcal{L}}^{-1})(Y) \\ &= -(dH)(Y) \end{aligned}$$

Since Y is arbitrary we obtain

$$dH = -(d\Lambda_{\mathcal{L}}(X_{\mathcal{L}})) \lrcorner\, \omega \ ,$$

which implies $d\Lambda_{\mathcal{L}}(X_{\mathcal{L}}) = X_H$. ▮

REMARK: This lemma shows that the dynamics encoded in $X_{\mathcal{L}}$ on TM is carried over to the the dynamics encoded in X_H on T^*M provided the Lagrangian is regular. Put simply, the Lagrange equations describe the same dynamics as do the Hamilton equations when the Lagrangian is regular.

EXERCISE: Let $M = \mathbb{R}^3$ and $\mathcal{L} = (m/2)\delta_{ij}v^i v^j$ be the Lagrangian for a free Newtonian particle of mass m. Show that $H = (m/2)\delta^{ij}p_i p_j$.

EXERCISE: Let $M = \mathbb{R}^3$ and $\mathcal{L} = (m/2)\delta_{ij}v^i v^j - k(1/2)\delta_{ij}q^i q^j$ be the Lagrangian of the non-relativistic isotropic harmonic oscillator. Find H.

6 The Poisson Algebra of C^∞ Functions on T^*M.

Two fundamental aspects of Hamiltonian dynamics that are central to both the canonical quantization scheme and geometric quantization are the Poisson bracket and the related Lie algebraic structure of C^∞ functions on T^*M. These ideas are introduced in this section along with some general facts relating to canonical transformations.

Let μ denote a differential form and X a smooth vector field on T^*M. From differential geometry we have the following general formula that relates the exterior derivative operator d, the Lie derivative operator L_X, and the (left) hook product $\lrcorner$:

$$L_X(\mu) = X \lrcorner\ d(\mu) + d(X \lrcorner\ \mu)\ . \tag{134}$$

Consider a 2-form $\mu = \mu_{ij} dx^i \wedge dx^j$ and a vector field $X = X^i \frac{\partial}{\partial x^i}$ in local coordinates (x^i) on a manifold M. Then equation (134), for μ a 2-form, can be derived from the familiar local coordinate formula

$$L_X(\mu_{ij}) = X^k \frac{\partial \mu_{ij}}{\partial x^k} + \mu_{kj} \frac{\partial X^k}{\partial x^i} + \mu_{ik} \frac{\partial X^k}{\partial x^j} \tag{135}$$

for the Lie derivative $L_X(\mu)$. Rewriting the second and third terms in this equations using the identity

$$\mu_{kj} \frac{\partial X^k}{\partial x^i} = \frac{\partial}{\partial x^i}(X^k \mu_{kj}) - X^k \frac{\partial \mu_{kj}}{\partial x^i}\ ,$$

and grouping terms, we have

$$\begin{aligned} L_X(\mu_{ij}) &= X^k \left(\frac{\partial \mu_{ij}}{\partial x^k} + \frac{\partial \mu_{jk}}{\partial x^i} + \frac{\partial \mu_{ki}}{\partial x^j}\right) + \frac{\partial}{\partial x^i}(X^k \mu_{kj}) - \frac{\partial}{\partial x^j}(X^k \mu_{ki}) \\ &= (X \lrcorner\ d\mu)_{ij} + (d(X \lrcorner\ \mu))_{ij} \end{aligned}$$

which is equation (134) expressed in local coordinates.

We will also need the following general formula from differential geometry which expresses the commutation relation between L_X and $Y \lrcorner$. If μ is a differential form and X and Y are smooth vector fields on a manifold M, then

$$L_X(Y \lrcorner\ \mu) - Y \lrcorner\ (L_X \mu) = [X, Y] \lrcorner\ \mu\ , \tag{136}$$

where $[X, Y]$ denotes the Lie bracket of the vector fields X and Y (i.e. the Lie derivative of Y with respect to X.)

EXERCISE: Derive equation (136) in local coordinates using the local coordinate formula

$$L_Z(\mu_i) = Z^j \partial_j(\mu_i) + \mu_j \partial_i(Z^j)$$

for the Lie derivative of a 1-form. Here I am using the notation $\partial_i = \frac{\partial}{\partial x^i}$.

Now recall (see equation (105)) that a **locally Hamiltonian vector field** X on T^*M is a vector field X such that

$$L_X \omega = 0 \ . \tag{137}$$

Combining equations (134) and (137) and using the fact that ω is closed we have that if X is locally Hamiltonian then

$$d(X \lrcorner\ \omega) = 0 \ . \tag{138}$$

This equation implies that $X \lrcorner\ \omega$ is a closed 1-form, and thus by the Poincaré lemma **locally** there exists a function f such that

$$df = -X \lrcorner\ \omega \ . \tag{139}$$

REMARK: If the first (De Rham) cohomology group $H^1(T^*M)$ of the manifold T^*M is trivial then all closed 1-forms are exact, and when this is true the function f will be globally defined. However, $H^1(T^*M)$ will generally not be trivial for an arbitrary manifold.

Consider next a Hamiltonian vector field X_f on T^*M determined by a function $f : T^*M \to \mathbb{R}$. By definition #5.1 X_f satisfies the equation

$$df = -X_f \lrcorner\ \omega \ . \tag{140}$$

Combining this equation with equation (134) we see that

$$L_{X_f} \omega = 0 \ . \tag{141}$$

We have the result that **each globally Hamiltonian vector field X_f is locally Hamiltonian , but generally a locally Hamiltonian vector field will not be globally Hamiltonian.**

DEFINITION #1: The set of all locally Hamiltonian vector fields on T^*M is denoted by

$$LHV \equiv LHV(T^*M) \ ,$$

and the set of (globally) Hamiltonian vector fields on T^*M is denoted by

$$HV \equiv HV(T^*M) \ .$$

The following proposition is due to S. Sternberg (*Lectures on Differential Geometry*, 1964.)

Proposition #1: If $X, Y \in LHV(T^*M)$, then $[X,Y] \in HV(T^*M)$. In particular, $[X,Y] = X_f$ where $f = 2\omega(X,Y)$.

Proof: We need to show that if X, Y are locally Hamiltonian vector fields, then there is a function $f : T^*M \to \mathbb{R}$ such that $[X,Y] \lrcorner\ \omega = -df$. Using the identity (136) and the fact that $L_X \omega = 0$ we have

$$\begin{aligned} [X,Y]\lrcorner\ \omega &= L_X(Y\lrcorner\ \omega) - Y\lrcorner\ (L_X\omega) \\ &= L_X(Y\lrcorner\ \omega) \end{aligned}$$

Now use the identity (134) and $d(Y\lrcorner\ \omega) = 0$ to obtain

$$\begin{aligned} [X,Y]\lrcorner\ \omega &= X\lrcorner\ d(Y\lrcorner\ \omega) + d(X\lrcorner\ (Y\lrcorner\ \omega)) \\ &= d(X\lrcorner\ (Y\lrcorner\ \omega)) \\ &= d(-2\omega(X,Y)) \end{aligned}$$

Thus $[X,Y]$ is the Hamiltonian vector field on T^*M determined by the function $f = 2\omega(X,Y)$. ▮

The importance of this proposition is that it reveals an underlying algebraic structure associated with Hamiltonian vector fields on T^*M. Recall that the set $\mathcal{X}(M)$ of all smooth vector fields on a manifold is an infinite dimensional Lie algebra under the Lie bracket. That is to say, $\mathcal{X}(M)$ is an infinite dimensional vector space, and the Lie bracket provides $\mathcal{X}(M)$ with the following multiplication rule:

$$\forall\ X, Y \in \mathcal{X}(M) \quad (X,Y) \longrightarrow [X,Y] \in \mathcal{X}(M) \ .$$

The Lie bracket satisfies

$$(a) \qquad [X,Y] = -[Y,X] \tag{142}$$

$$(b) \qquad [X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0 \tag{143}$$

A (non-associative) algebra with a multiplication rule that satisfies these two properties is called a **Lie algebra**. Property (b) is referred to as the Jacobi identity.

EXERCISE: Prove that the sets LHV and HV of locally and globally Hamiltonian vector fields, respectively, on T^*M are infinite-dimensional vector spaces. Thus you need to show that if $X, Y \in LHV$ and $\alpha \in \mathbb{R}$, then $X + Y \in LHV$ and $\alpha X \in LHV$, and similarly for HV.

Since LHV is an infinite dimensional vector space, the proposition shows that the Lie bracket provides LHV with the algebraic structure of an infinite dimensional Lie algebra, which is a proper subalgebra of $\mathcal{X}(T^*M)$. Note that since each Hamiltonian vector field is itself locally Hamiltonian, the set HV of Hamiltonian vector fields is also a Lie algebra under the Lie bracket, and HV is a subalgebra of LHV. This Lie algebraic structure of the set HV leads naturally to the concept of the Poisson bracket of functions on T^*M.

We can obtain some understanding of the geometrical significance of the proper sub-algebras LHV and HV of the Lie algebra $\mathcal{X}(T^*M)$ by using the symplectic 2-form ω to relate each to subsets of the space $\Lambda^1 \equiv \Lambda^1(T^*M)$ of differential 1-forms on T^*M . Each vector field $X \in \mathcal{X}(T^*M)$ corresponds to a unique 1-form λ_X via

$$\lambda_X = -X \lrcorner\, \omega \ .$$

From the definitions we see that

$$\begin{array}{ll} (a) & X \in LHV \Longrightarrow \lambda_X \text{ is closed} \\ (b) & X \in HV \Longrightarrow \lambda_X \text{ is exact} \\ (c) & X \notin LHV \Longrightarrow \lambda_X \text{ is not closed} \end{array}$$

Thus locally and globally Hamiltonian vector fields on T^*M correspond, respectively, to closed and exact differential 1-forms.

Denote by $C^\infty(T^*M)$ the set of all smooth (C^∞) functions $f : T^*M \to \mathbb{R}$. This set forms an infinite dimensional vector space, since the sum of two smooth functions and a scalar multiple of a smooth function are again smooth functions on T^*M. Moreover, we may think of $C^\infty(T^*M)$ as a **commutative algebra** under the pointwise multiplication

$$f \cdot g(x) = f(x)g(x) \ .$$

This view of $C^\infty(T^*M)$ as a commutative algebra will be used shortly.

Now consider two classical observable $f, g \in C^\infty(T^*M)$. The associated Hamiltonian vector fields determined by f and g are, from equation (129),

$$\begin{aligned} X_f &= \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial}{\partial p_i} \\ X_g &= \frac{\partial g}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial g}{\partial q^i}\frac{\partial}{\partial p_i} \end{aligned} \qquad (144)$$

If X is a smooth vector field and f is a smooth function on T^*M , then $L_X(f) \equiv X(f)$ is another smooth function on T^*M . In particular, applying the Hamiltonian vector field X_f to the classical observable g we obtain

$$X_f(g) = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} \tag{145}$$

which we recognize as the classical **Poisson bracket** of the observables g and f, which we denote by $\{g, f\}$. We have the result that if $X_f \in HV$ and $g \in C^\infty(T^*M)$ then

$$\boxed{X_f(g) = \{g, f\}} \tag{146}$$

REMARK: We take equation (146) as the **definition** of the Poisson bracket of functions on T^*M . As remarked above this definition agrees with the classical definition, and with Śniatycki's definition (page 39). Note, however, that the definition (146) has the disagreeable feature of reversing the order of appearance of the functions f and g on the two sides of the equation. Other authors (eg. Woodhouse, *Geometric Quantization*, 1980, pg. 11) use the definition $\{f, g\} := X_f(g)$ instead of (146), but this results in the introduction of a minus sign in the coordinate expression for the Poisson bracket.

The Poisson bracket $\{g, f\}$ can also be expressed as $2\omega(X_f, X_g)$ since

$$\begin{aligned} \{g, f\} &= X_f(g) \\ &= dg(X_f) \\ &= -(X_g \lrcorner\, \omega)(X_f) \\ &= -2\omega(X_g, X_f) \\ &= 2\omega(X_f, X_g) \end{aligned} \tag{147}$$

What we want to do now is to use Proposition #1 to transfer the Lie algebraic structure of HV over to $C^\infty(T^*M)$. Thinking of $C^\infty(T^*M)$ as an infinite dimensional vector space we define a multiplication on elements of $C^\infty(T^*M)$ by

$$\forall\ f, g \in C^\infty(T^*M) \quad f \cdot g \overset{def}{=} \{f, g\} \in C^\infty(T^*M)\ . \tag{148}$$

From the definition (146) one may show that this multiplication rule satisfies, $\forall\ f, g, h \in C^\infty(T^*M)$,

$$\begin{aligned} &(a) \qquad \{f, g\} = -\{g, f\} \\ &(b) \qquad \{f, \{g, h\}\} + \{h, \{f, g\}\} + \{g, \{h, f\}\} = 0 \end{aligned} \tag{149}$$

The result is that the vector space $C^\infty(T^*M)$, provided with the multiplication rule (148), is an **infinite dimensional Lie algebra**. Moreover, above we have indicated that $C^\infty(T^*M)$ also has the structure of an infinite dimensional **commutative algebra** under pointwise multiplication. For $f, g, h \in C^\infty(T^*M)$ we have $gh \in C^\infty(T^*M)$, and we can ask how the Poisson bracket in the Lie algebra is related to the pointwise multiplication in the commutative algebra. (Both algebras have the same underlying vector space $C^\infty(T^*M)$.) From the definition (146) we have

$$\begin{aligned}\{f, gh\} &= -X_f(gh) \\ &= -X_f(g)h - gX_f(h) \\ &= \{f,g\}h + g\{f,h\}\end{aligned} \tag{150}$$

This equation shows that the Poisson bracket acts as **a derivation** on the commutative algebra. Guillemin and Sternberg (*The moment map and collective motion*, Ann. Physics, **127**(1980), pp. 220-253) have given the name **Poisson algebra** to the algebraic structure of a commutative algebra with a "bracket" multiplication that is anti-symmetric, satisfies a Jacobi identity, and which also acts as a derivation on the commutative algebra. The set $C^\infty(T^*M)$ equipped with the Poisson bracket defined in (146) is thus a Poisson algebra.

DEFINITION #2: The Poisson algebra of C^∞ real-valued functions on T^*M with multiplication rule (146) will be denoted by

$$HF \equiv HF(T^*M, \mathbb{R}) \; . \tag{151}$$

We now have two Lie algebras on T^*M defined by the symplectic 2-form ω, namely HV with the Lie bracket of vector fields as the product, and HF with the Poisson bracket of functions as the product. How are these two Lie algebras related? We have a natural map $\phi : HF \to HV$ given by

$$f \longrightarrow \phi(f) = X_f \; . \tag{152}$$

Moreover, from equation (147) we have

$$\{f, g\} = -2\omega(X_f, X_g) \; . \tag{153}$$

This last equation together with proposition #1 implies

PROPOSITION #2: $[X_f, X_g] = -X_{\{f,g\}}$.

From this proposition we infer the following fact:

$$\phi(\{f, g\}) = -[X_f, X_g] \ , \tag{154}$$

This shows that the map $\phi : HF \to HV$ is an **anti-homomorphism** of the Poisson algebra HF into the Lie algebra HV.

EXERCISE: Show that the **center** of the Poisson alegbra $(C^\infty(T^*M), \{\ ,\ \})$, that is the set of all $f \in C^\infty(T^*M)$ such that $\{f, g\} = 0$ for all $g \in C^\infty(T^*M)$, consists of the constant functions on T^*M .

EXERCISE: Show that the kernel of the map $f \longrightarrow X_f$ is the center of the Poisson algebra.

If we partition HF into equivalence classes, two functions f and g being equivalent if they differ by a constant, then we may define the quotient space $HF/\mathbb{R}$, where we are identifying the constant functions (i.e. the center of the Poisson algebra) with the real numbers $\mathbb{R}$. Thus as Lie algebras we have

$$HF/\mathbb{R} = HV \ . \tag{155}$$

The sequence of maps

$$0 \longrightarrow \mathbb{R} \longrightarrow HF \longrightarrow HV \longrightarrow 0 \tag{156}$$

is exact in the sense that the image of each map is the kernel of the following map. Because HV misses being isomorphic to HF only by elements of the center of HF, the Lie algebra HF is referred to as **a central extension** of HV (cf. Guillemin and Sternberg, *Symplectic techniques in physics*,p. 91, 1984, Cambridge).