A very short Introduction to Classical Mechanics

Sylvain Rossi

November 12, 2018

Contents

1 Introduction

Classical mechanics is used in physics to describe dynamical systems. By that, one means to study the motion of a particle subject to different types of forces, e.g. a point particle attached to a string in a uniform gravitational field. A physicist has two approaches to derive the equations of motion of the particle: either setting up the Lagrangian and solve the Euler-Lagrange equation or setting up the Hamiltonian and solve Hamilton's equations. In physics, these two approaches are often assumed to be equivalent and used interchangeably. In fact, they are not interchangeable, only in case the Lagrangian and Hamiltonian do satisfy some specific conditions. We will discuss what these conditions are and how to switch from one formalism to the other. Finally, we will to show how the Noether theorem translates from one formalism into the other.

2 The two formalisms

2.1 Lagrangian formalism

This section is mostly based on the book [1] and the first two chapters of [5]. The first reference is more of a rough overview of classical mechanics which gives the most important ideas, the second reference is more rigorous, but skips a lot of steps.

Definition 2.1.1. A (smooth) *Lagrangian system* is a pair consisting of an n-dimensional smooth manifold M and a smooth function: $l : TM \times \mathbb{R} \to \mathbb{R}$ (called *Lagrangian*), let $l_t : TM \to \mathbb{R}$ denote the map $l_t(q, v) := l(q, v, t)$.

Definition 2.1.2. Let $I := [a, b] \subset \mathbb{R}$ be a bounded interval and (M, l) a Lagrangian system. The *action functional* $\mathcal{S}_l(\gamma)$ associated to l is defined by

$$\mathcal{S}_l(\gamma) := \int_a^b l_t(\gamma(t), \dot{\gamma}(t))dt.$$

with γ a smooth path in M.

Definition 2.1.3. Let M be an n-dimensional smooth manifold and $\gamma : [a, b] \to M$ a smooth path in M. A *variantion (with fixed endpoints) of* γ is a smooth function $\sigma : I \times I \to M$ with the following properties:

$$\forall (s, t) \in I \times I : \ \sigma(0, t) = \gamma(t), \ \sigma(s, a) = \gamma(a), \ \sigma(s, b) = \gamma(b).$$

Definition 2.1.4. A path is called *extremal* if $\frac{\partial}{\partial s}|_{s=0}\mathcal{S}_l(\sigma(s, \cdot)) = 0$.

Lemma 2.1.5. *(Fundamental Lemma of the Calculus of Variation) Let* $(a, b) \subset \mathbb{R}$ *be an open subset and* $f : (a, b) \to \mathbb{R}$ *a continuous function. If for all compactly* *supported smooth function* $h : (a,b) \to \mathbb{R}$,

$$\int_a^b f(x)h(x)dx = 0,$$

then f *is identically zero.*

Theorem 2.1.6. *(Euler-Lagrange) Let* (M,l) *be a Lagrangian system and* γ *a path in* M*. The path* γ *is extremal if and only if the following equation holds:* $\frac{d}{dt}\frac{\partial}{\partial v}l_t(\gamma,\dot\gamma) = \frac{\partial}{\partial q}l_t(\gamma,\dot\gamma)$ *(The Euler-Lagrange equations).*

Proof. The idea behind the proof is to variate the action along a path. To do this, we will split the path in many fragment such that each one fits into a chart and only variate within that chart. So let $(q,v) \in TM$ denote a point in local coordinates on the tangent bundle TM of M and $\xi_i(t)$ be equal to $\frac{\partial}{\partial s}\big|_{s=0}\sigma_i(s,t)$.

$$\begin{aligned}
0 &= \frac{\partial}{\partial s}\Big|_{s=0}\int_a^b l_t(\sigma(s,t),\frac{d}{dt}\sigma(s,t))dt = \int_a^b \frac{\partial}{\partial s}\Big|_{s=0} l_t(\sigma(s,t),\frac{d}{dt}\sigma(s,t))dt \\
&= \sum_{i=1}^n \int_a^b \Big(\frac{\partial}{\partial q_i}l_t(\sigma(s,t),\dot\sigma(s,t))\cdot\frac{\partial}{\partial s}\Big|_{s=0}\sigma_i(s,t) \\
&\qquad + \frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\cdot\frac{\partial}{\partial s}\Big|_{s=0}(\frac{d}{dt}\sigma_i(s,t))\Big)dt \\
&= \sum_{i=1}^n \int_a^b \Big(\frac{\partial}{\partial q_i}l_t(\sigma(s,t),\dot\sigma(s,t))\cdot\xi_i(t) + \frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\cdot\frac{d}{dt}\xi_i(t)\Big)dt \\
&= \sum_{i=1}^n \int_a^b \Big[\frac{\partial}{\partial q_i}l_t(\sigma(s,t),\dot\sigma(s,t)) - \frac{d}{dt}\big(\frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\big)\Big]\cdot\xi_i(t)dt \\
&\quad + \sum_{i=1}^n \int_a^b \frac{d}{dt}\Big(\frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\cdot\xi_i(t)\Big)dt \\
&= \sum_{i=1}^n \int_a^b \Big[\frac{\partial}{\partial q_i}l_t(\sigma(s,t),\dot\sigma(s,t)) - \frac{d}{dt}\big(\frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\big)\Big]\cdot\xi_i(t)dt \\
&\quad + \frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\cdot\underbrace{\xi_i(b)}_{=0} - \frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\cdot\underbrace{\xi_i(a)}_{=0} \\
&= \sum_{i=1}^n \int_a^b \Big[\frac{\partial}{\partial q_i}l_t(\sigma(s,t),\dot\sigma(s,t)) - \frac{d}{dt}\big(\frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\big)\Big]\cdot\xi_i(t)dt
\end{aligned}$$

Thus by the fundamental lemma of the calculus of variation, $\frac{\partial}{\partial q_i}l_t(\sigma(s,t),\dot\sigma(s,t)) - \frac{d}{dt}\big(\frac{\partial}{\partial v_i}l_t(\sigma(s,t),\dot\sigma(s,t))\big)$ has to vanish. □

2.1.1 Noether's theorem

Definition 2.1.7. A Lagrangian system (M, l) *admits* a one-parameter group of diffeomorphism θ_s if the following holds:

$$\forall s \in I, \forall (q, v) \in TM: \; l_t(\theta_s(q), \theta_s^*(v)) = l_t(q, v)$$

Theorem 2.1.8. *(Noether's Theorem) Let (M, l) be a Lagrangian system which admits a one parameter group of diffeomorphism θ_s.*
The Euler-Lagrange equation to l has a first integral:

$$\mathcal{I} : TM \to \mathbb{R}, \; \textit{in local coordinates } (q, v) \in TM : \mathcal{I}(q, v) = \frac{\partial}{\partial v} l_t \left. \frac{\partial}{\partial s} \right|_{s=0} \theta_s(q).$$

Proof. Let $\gamma : I \to M$ be an extremal path of l.
We define a variation of that path through $\kappa(s, t) := \theta_s(\gamma(t))$ and denote the pullback of the velocity vectorfield by $\dot{\kappa}(s, t) := \theta_s^* \dot{\gamma}(t)$. The Lagrangian function is invariant under this variation, therefore we get the following equation:

$$\begin{aligned} 0 &= \left. \frac{\partial}{\partial s} \right|_{s=0} l_t(\kappa(s,t), \dot{\kappa}(s,t)) \\ &= \sum_{i=1}^{n} \Big(\frac{\partial}{\partial q_i} l_t(\kappa(s,t), \dot{\kappa}(s,t)) \left. \frac{\partial}{\partial s} \right|_{s=0} \kappa_i(s,t) + \frac{\partial}{\partial v_i} l_t(\kappa(s,t), \dot{\kappa}(s,t)) \left. \frac{\partial}{\partial s} \right|_{s=0} \dot{\kappa}(s,t) \Big) \end{aligned}$$

For the same reason, the paths $\kappa(s, t)$ are extremal for any fixed $s \in I$ and satisfy the Euler-Lagrange equation. Thus, by using the relation above, one can see:

$$\begin{aligned} &\forall i \in \{1, \ldots, n\} : \big(\frac{\partial}{\partial q_i} l_t(\kappa(s,t), \dot{\kappa}(s,t)) - \frac{d}{dt} \frac{\partial}{\partial v_i} l_t(\kappa(s,t), \dot{\kappa}(s,t)) \big) \cdot \left. \frac{\partial}{\partial s} \right|_{s=0} \kappa_i(s,t) = 0 \\ &\Rightarrow \frac{\partial}{\partial v_i} l_t(\kappa(s,t), \dot{\kappa}(s,t)) \cdot \left. \frac{\partial}{\partial s} \right|_{s=0} (\frac{d}{dt} \kappa_i(s,t)) + \frac{d}{dt} \frac{\partial}{\partial v_i} l_t(\kappa(s,t), \dot{\kappa}(s,t)) \cdot \left. \frac{\partial}{\partial s} \right|_{s=0} \kappa_i(s,t) = 0 \\ &\Rightarrow \frac{d}{dt} \, \mathcal{I}(\gamma(t), \dot{\gamma}(t)) = 0. \end{aligned}$$

□

2.2 Hamiltonian formalism

Here again, as for the Lagrangian formalism, we will use [5] as our reference.

Definition 2.2.1. A (smooth) *Hamiltonian system* is a pair consisting of an n-dimensional manifold M and a smooth function $h : T^*M \times \mathbb{R} \to \mathbb{R}$ (called *Hamiltonian*). Let $h_t(q, p) := h(q, p, t)$.

Definition 2.2.2. Let (M, h) be a Hamiltonian system and ω a canonical symplectic form making T^*M into a symplectic manifold. Then one can define a *Hamiltonian vector field* X_h of h on T^*M by

$$i_{X_h}\omega = dh.$$

In local coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ on T^*M we have

$$X_h = \sum_{j=1}^{n} \frac{\partial h}{\partial p_j}\frac{\partial}{\partial q^j} - \frac{\partial h}{\partial q^j}\frac{\partial}{\partial p_j}.$$

Definition 2.2.3. Let (M, h) be a Hamiltonian system. Let X_h be the Hamiltonian vector field of h. The local flow $\varphi_h : \mathcal{D} \to T^*M$ generated by X_h is called *Hamiltonian flow* of h. Let $\Gamma : (t_0, t_1) \to T^*M$ be an integral curve of X_h, $\Gamma = (\gamma, \rho)$, where $\gamma : (t_0, t_1) \to M$ and ρ is a section of $\gamma^\star T^*M$. Then we say that Γ satisfies the *Hamilton equations* if in local coordinates (γ, ρ) satisfy:

$$\dot{\gamma}^j(t) = \frac{\partial h}{\partial p_j}(t, \gamma(t), \rho(t)),\ \dot{\rho}_j(t) = -\frac{\partial h}{\partial q^j}(t, \gamma(t), \rho(t)),\ \forall j = 1, \ldots, n.$$

Definition 2.2.4. Let (M, h) be a Hamiltonian system, λ the canonical one-form on T^*M and $I := [a, b] \subset \mathbb{R}$ a bounded interval. The *Hamiltonian action* $\mathcal{A}_h$ associated to h of a smooth curve $\Gamma : I \to T^*M$ is defined in the following way:

$$\mathcal{A}_h(\Gamma) := \int_a^b \Gamma^*\lambda - h_t(\Gamma(t))dt.$$

Proposition 2.2.5. *The extremal curves of $\mathcal{A}_h$ are precisely the ones satisfying Hamilton's equations.*

Proof. As for the Lagrangian setting, the idea behind the proof is to variate the Hamiltonian action along a path in the cotangent space. To do this, we will split the path in many fragment such that each one fits into a chart and only variate within that chart. So let $I = [a, b]$ and let $\Gamma \in I \times T^*M$, we denote the variation of Γ, by $\Gamma^s := (\gamma^s, \rho^s)$ with $\gamma^s \in M$ and $\rho^s \in T^*M$ for all $s \in I$.

$$\begin{aligned}
0 &= \frac{\partial}{\partial s}\Big|_{s=0} \int_a^b \Gamma^{s*}\lambda - h_t(\gamma^s(t), \rho^s(t))dt \\
&= \int_a^b \frac{\partial}{\partial s}\Big|_{s=0} \Big(\rho^s(t)\big(\dot\gamma^s(t)\big) - h_t(\gamma(t), \rho(t))\Big)dt \\
&= \int_a^b \Big(\frac{\partial}{\partial s}\Big|_{s=0} \rho^s(t)\big(\dot\gamma^s(t)\big) - \frac{\partial}{\partial s}\Big|_{s=0} \gamma^s(t)\ \frac{\partial}{\partial q}\Big|_q h_t(\gamma^s(t), \rho^s(t)) \\
&\quad - \frac{\partial}{\partial s}\Big|_{s=0} \rho^s(t)\ \frac{\partial}{\partial p}\Big|_p h_t(\gamma^s(t), \rho^s(t))\Big)dt \\
&= \int_a^b (\frac{\partial}{\partial s}\Big|_{s=0} \rho^s(t))\big(\dot\gamma^s(t) - \frac{\partial}{\partial p}\Big|_p h_t(\gamma^s(t), \rho^s(t))\big)dt \\
&\quad - \int_a^b (\frac{\partial}{\partial s}\Big|_{s=0} \gamma^s(t))\big(\dot\rho^s(t)(\dot\gamma^s(t)) - \frac{\partial}{\partial q}\Big|_q h_t(\gamma^s(t), \rho^s(t))\big)dt \\
&\quad -\Big[\frac{\partial}{\partial s}\Big|_{s=0} \gamma^s(t)\ \frac{\partial}{\partial q}\Big|_q \rho^s(t)\big(\dot\gamma^s(t)\big)\Big]\Big|_a^b .
\end{aligned}$$

The last term of the equation equals zero as a variation with fixed endpoints has vanishing derivative at the endpoints. The two other equations have to be zero, therefore the extremal path of the Hamiltonian action satisfies the Hamilton equation of motion. □

2.2.1 Noether's theorem

Definition 2.2.6. A Hamiltonan h is said to be invariant under a one-parameter group of symplectormorphisms θ_s (w.r.t. the canonical symplectic form ω on T^*M) if the following holds:

$$\forall s \in I,\ \forall (q,p) \in T^*M :\ h_t(\theta_s(q), \theta_s^*(p)) = h_t(q,p).$$

Definition 2.2.7. Suppose (M, ω) is a connected symplectic manifold and $\Phi : G \times M \to M$ is a *symplectic action* of the Lie group G on M, i.e. for all $g \in G$, the map $\Phi_g(x) := \Phi(g, x)$ is a symplectomorphism. Let $J : M \to \mathfrak{g}^* = T_e^*G$ be a smooth map and $\xi \in \mathfrak{g} = T_eG$. Define $\hat J : M \to \mathbb{R}$ by $\hat J(\xi)(x) := J(x) \cdot \xi$ and $\xi_M(x) := \frac{d}{dt}\big|_{t=0} \Phi(\exp(t\xi), x)$, called the *infinitesimal generator of the action corresponding to* ξ.
The map J is called a *momentum mapping* for the action provided that for every $\xi \in \mathfrak{g}$:

$$d\hat J(\xi) = i_{\xi_M}\omega,$$

or equivalently, if $X_{\hat{J}(\xi)}$ is the Hamiltonian vector field of $\hat{J}(\xi)$, then

$$X_{\hat{J}(\xi)} = \xi_M.$$

Definition 2.2.8. Let $f, g : M \to \mathbb{R}$ be two smooth maps and $X_f,\ X_g$ be their respective Hamiltonian flows, then we define the *Poisson bracket* as

$$\{f, g\} = -i_{X_f} i_{X_g} \omega.$$

Remark. Observe that $\{f, g\} = -i_{X_f} i_{X_g} \omega = -i_{X_f} dg = -dg \cdot X_f = -\mathcal{L}_{X_f} g$.

Lemma 2.2.9. *With f, g as above, the following are equivalent:*

(i) *f is constant on flows of X_g,*

(ii) *g is constant on flows of X_f,*

(iii) *$\{f, g\} = 0$.*

Proof. suppose φ_t is the flow of X_f, then:

$$\frac{d}{dt}(g \circ \varphi_t) = \frac{d}{dt}(\varphi_t^* g) = \varphi_t^* \mathcal{L}_{X_f} g = -\varphi_t^* \{f, g\}.$$

The result follows from the fact that $\varphi_t^* \{f, g\} = 0$ for all $t \in I$ if and only if $\{f, g\} = 0$. □

Theorem 2.2.10. *(Noether's Theorem). Let θ_s be a one-parameter group of symplectomorphisms and suppose that for all $s \in \mathbb{R}$, $h : T^*M \to \mathbb{R}$ is invariant under θ_s. Then J is an integral for X_h, i.e. if φ_h is the flow of X_h, then:*

$$J(\varphi_h(x)) = J(x).$$

A complete proof of this theorem can be found in [4].

3 Connecting the two formalism

3.1 The Legendre transformation

This section is mainly copied out of [3] and the missing steps in the proofs were completed.

We will show how to transition from a Lagrangian to a Hamiltonian in three steps. First, we will define the Legendre transform on a vector space. Secondly, use step 1. on a tangent space and extend it to the tangent bundle. Finally, we will check that everything that we defined is smooth when needed.

Definition 3.1.1. Let V be a real vector space, a subset C is said to be *convex* if $\alpha x + (1-\alpha)x'$ is in C whenever $x,\ x'$ are in C, and $\alpha \in]0,1[$.

Proposition 3.1.2. *(Properties of convex sets) Let V be a real vector space, then we have the following properties:*

(i) Let $(C_j)_{j\in J}$ be a family of convex sets, then the intersection $\bigcap_{j\in J} C_j$ is also convex.

(ii) For $i = 1, \ldots, k$, let C_i be a convex subset of the vector space V_i. Then $C_1 \times \cdots \times C_k$ is a convex subset of $V_1 \times \cdots \times V_n$.

(iii) The opposite $-C$ is a convex set.

(iv) Let α_1 and α_2 be two real numbers. The sum

$$\alpha_1 C_1 + \alpha_2 C_2 := \{x = \alpha_1 x_1 + \alpha_2 x_2 | x_1 \in C_1,\ x_2 \in C_2\}$$

of two convex sets C_1 and C_2 is convex.

Proof. The properties (i) and (ii) are straightforward. To show properties (ii), (iii) and (iv), we will prove the following claim:
Let V, W be real vector spaces, $A : V \to W$ be an affine mapping and $C \subset V$ and $D \subset W$ convex sets. Then the image $A(C)$ of C under A is convex in W and the inverse image $A^{-1}(D)$ is convex in V.
Indeed, for any x, x' in V, the image under A of the segment $[x, x']$ is the segment $[A(x), A(x')]$. In the same way, if x and x' are such that $A(x)$ and $A(x')$ lie in D, then every point of the segment $[x, x']$ has its image in $[A(x), A(x')] \subset D$. □

Definition 3.1.3. Let V be an n-dimensional vector space over $\mathbb{R}$ and $F : V \to \mathbb{R}$ a smooth function. Then F is *strictly convex* if, for all $x \in V$ and $v \in V \setminus \{0\}$, the *Hessian*

$$d^2 F_x(v) := \left.\frac{d^2}{dt^2}\right|_{x=0} F(x + vt) > 0.$$

Lemma 3.1.4. *If F is strictly convex, the following conditions are equivalent:*

(i) $\exists x \in V :\ dF_x = 0$.

(ii) F has a local minimum at some point p_0

(iii) F has a unique local minimum (hence a unique global minimum).

(iv) $F(x)$ tends to $+\infty$ as x tends to infinity in V.

Proof. ((i) $\Rightarrow$ (ii)) If $dF_x = 0$ at $x_0 \in V$, then for all $v \in V \setminus \{0\}$, then $f(t) = F(x_0 + tv)$ has a critical point at $t = 0$ and since F is strictly convex, we have that $f''(0) > 0$. Thus F restricted to $t \mapsto x_0 + tv$ has a local minimum at $t = 0$, i.e. at x_0 and hence F has a local minimum at x_0.
((ii) $\Rightarrow$ (iii)) suppose X_1, x_2 are two local minima of F. Take $v := x_2 - x_1$. Then $f(t) := F(x_1 + tv)$ would have two local minima at $t = 0$ ant $t = 1$, but this would contradict the convexity of f (one can easily prove this using the mean value theorem).
((iii) $\Rightarrow$ (iv)) Let us give V an inner product. Since F has a unique global minimum at x_0, $f(t) := F(x_0 + tv)$ (for any $v \in V$ with $||v|| = 1$) has a global minimum at $t = 0$. Hence $f'(t) \neq 0$ for all $t \neq 0$ and $f'(t) > 0$ for all $t > 0$, thus $f|_{t>0}$ is strictly increasing. If $\lim_{t\to\infty} f(t) < \infty$, then $f|_{t>0}$ is bounded from above, thus $\liminf_{t\to\infty} f'(t) = 0$, but this contradicts $f''(t) > 0$, so $\lim_{t\to\infty} f(t) = \infty$. Since the sphere $\{v \in V|\ ||v|| = 1\}$ is compact, F takes on its maximum, thus $F(x) \to \infty$ for $||x|| = ||x_0 + tv|| \to +\infty$.
((iv) $\Rightarrow$ (i)) For $r > 0$ large enough, $B := \{x \in V | F(x) \leq r\}$ is compact and non-empty. Since F is continuous, it attains its minimum on B, if the minimum is on the interior of B, then it must be a local minimum and hence $dF = 0$ at that point. If the maximum occurs on the boundary of B, then $F(x_0) = r$ and $F = r$ on some neighbourhood of $x_0 \in B$. Since $F > r$ on the complement of B, x_0 also has to be a local minimum with $dF(x_0) = 0$. □

From now assume F is strictly convex. For $x \in V$, one has a canonical identification: $T_x^*V \cong V^*$. Meaning, we can regard the section dF of the cotangent bundle as map

$$dF : V \to V \times V^*.$$

Definition 3.1.5. Let $\pi_2 : V \times V^* \to V^*$ such that $\pi_2((v, l)) = l$. The *Legendre transform* of F is the map

$$L_F : V \to V^*,\ L_F := \pi_2 \circ dF.$$

Remark. If $(x^1, \ldots, x^n)$ are a linear system of coordinates in V and $(l_1, \ldots, l_n)$ the dual coordinates on V^*, this is just the map:

$$L_F(x) = l \Leftrightarrow l_i = \frac{\partial F}{\partial x^i},\ i = 1, \ldots, n.$$

Under these coordinates, strict convexity is equivalent to

$$\det\Big(\frac{\partial^2 F}{\partial x^i \partial x^j}\Big) > 0.$$

Thus, this also means that for every point x in V, there is a neighbourhood around x which gets mapped diffeomorphically onto a neighbourhood of its image.

Definition 3.1.6. Let $F : V \to \mathbb{R}$ be strictly convex. Then F is stable if it satisfies any one of the conditions of the previous lemma.

Given an element $l \in V^*$, we will denote by $F_l : V \to \mathbb{R}$ the function

$$F_l(x) := F(x) - l(x), \ \forall x \in V.$$

Remark. It is clear that since F is strictly convex, so is F_l as they posses the same Hessian.

Definition 3.1.7. Let F be strictly convex. The *stability set of* F is the set $V_s^* \subset V^*$ of all $l \in V^*$ such that F_l is stable.

Remark. The image of the Legendre transform is V_s^* as F_l is stable if and only if, for some $x \in V$, $(dF_l)_x = 0 \ \Leftrightarrow \ dF_x = l$.

Theorem 3.1.8. *V_s^* is an open convex subset of V^*, and the Legendre transform maps V diffeomorphically onto this set.*

Proof. Let $l_1, l_2 \in V^*$. Then $F_{l_1}(x)$ and $F_{l_2}(x)$ both tend to $+\infty$ as x tends to infinity in V, so for $0 < t < 1$,

$$(tF_{l_1} + (1-t)F_{l_2})(x) \to +\infty, \text{ for } x \to \infty.$$

Thus $tl_1 + (1-t)l_2$ is in V_s^*, which implies that V_s^* is convex. As discussed above, V_s^* is the image of a map which is locally a diffeomorphismat every point, so it is open. Finally, it follows that $L_F(x)$ is bijective, because $L_F(x) = l$ if and only if x is the unique minimum point of F_l. □

Corollary. *Given $l \in V_s^*$, the pre-image, $p = L_l^{-1}(l)$, is the unique point in V where F_l achieves its minimum value.*

We now ask ourselves when he stability set V_s^* is equal to the total dual space V^* (i.e. that every point $l \in V^*$) is in the stability set. One of these criteria is the following one, which is often satisfied in physical Lagrangian systems:

Theorem 3.1.9. *Let $||\cdot||$ be a metric on V. If $F : V \to \mathbb{R}$ is strictly convex and superlinear i.e*

$$\lim_{||v|| \to \infty} \frac{F(x)}{||x||} = +\infty.$$

Then $V_s^* = V^*$.

Proof. Let $l \in V^*$, then $l \in V_s^*$ if and only if F_l is superlinear too, but for $v \in V \setminus \{0\}$

$$\lim_{a \to \infty} \frac{F_l(av)}{||av||} = \lim_{a \to \infty} \frac{F(av) - al(v)}{a||v||} = +\infty.$$

Hence $l \in V_s^*$. □

Theorem 3.1.10. *Let $F_1, F_2 : V \to \mathbb{R}$ be strictly convex functions, with stability sets V_1^* and V_2^* respectively. Then $V_1^* + V_2^*$ is the stability set of $F_1 + F_2$.*

Proof. The sum of two convex sets is convex, so $V_1^* + V_2^*$ is convex. If $l_1 \in V_1^*$ and $l_2 \in V_2^*$, then $(F_1)_{l_1}$ and $(F_2)_{l_2}$ are stable, hence their sum is stable, so there exists a point $x \in V$ at which

$$d(F_1 + F_2)_x = l_1 + l_2.$$

□

Definition 3.1.11. Let $F : V \to \mathbb{R}$ be strictly convex and $l \in V_s^*$. The *dual function* F^* of F is

$$F^*(l) := -\min_{x \in V} F_l(x),$$

let $p = L_F^{-1}(l)$, then F^* takes the concrete form

$$F^*(l) = l(p) - F(p).$$

Theorem 3.1.12. *The inverse of the Legendre transform*

$$L_F : V \to V_s^*$$

is the Legendre transform $L_{F^*} : V^* \to (V^*)^* \cong V$.

Proof. As $V \times V^*$ corresponds to the cotangent bundle of V, by reversing the product, we get another equality:

$$V \times V^* \underbrace{\cong}_{switch} V^* \times V \cong V^* \times (V^*)^*,$$

but $V^* \times (V^*)^*$ is the cotangent bundle of V^*. So there is a natural identification of the cotangent bundle of V with the cotangent bundle of V^*. Both can be identified with $V \times V^*$. Let α_I and α_{II} be the canonical "tautology" one-forms on T^*V and T^*V^*. Given the identification, we can think of α_I and α_{II} as living in $V \times V^*$.

We claim that $\alpha_I = dB - \alpha_{II}$, where $B : V \times V^* \to \mathbb{R}$ is the function $B(p, l) = l(p)$.

Indeed, let $v_0 \in V$ and $(p_0, v_0^{**}), (v_p, v_{q^{**}}) \in V^* \times (V^*)^*$, we have that locally α_{II} (on $v^* \times (V^*)^*$) is

$$\alpha_{II}|_{(p_0, v_0^{**})} \big((v_p, v_{q^{**}})\big) = \big(v^{**} dp\big)\big|_{(p_0, v_0^{**})} \big((v_p, v_{q^{**}})\big) = v_0^{**}(v_p) = \mathrm{ev}_{v_0}(v_p) = v_p(v_0).$$

Now let $(p_0, v_0^{**}) := f(v_0, p_0) \in V^* \times (V^*)^*$ and $(v_q, v_p) \in V \times V^*$. We have

$$\begin{aligned} f^*\alpha_{II}|_{(v_0,p_0)}(v_q, v_p) &= \alpha_{II}|_{f(v_0,p_0)}\left(df|_{(v_0,p_0)}(v_q, v_p)\right) \\ &= \alpha_{II}|_{(p_0,v_0^{**})}\left((v_p, v_{q^{**}})\right) \\ &= v_p(v_0). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} dB|_{(p_0,v_0)}(v_p, v_q) &= \left.\frac{dB(p_0 + tv_p, v_0 + tv_q)}{dt}\right|_{t=0} \\ &= \left.\frac{d(p_0(x_0) + tv_p(v_0) + tp_0(v_q) + t^2 v_p(v_q))}{dt}\right|_{t=0} \\ &= v_p(v_0) + p_0(v_q). \end{aligned}$$

(This can be done as B is defined on a vector space.)
But,

$$\begin{aligned} \alpha_I|_{(v_0,p_0)}\left((v_q, v_p)\right) &= pdv|_{(v_0,p_0)}\left((v_q, v_p)\right) \\ &= p_0(v_q). \end{aligned}$$

Hence, $\alpha_I + f^*\alpha_{II} = dB$ as $(v_0, p_0), (v_q, v_p) \in V \times V^*$ were random.

Thus the canonical symplectic forms $\omega_I = d\alpha_I$ and $\omega_{II} = d\alpha_{II}$ satisfy

$$\omega_I = -\omega_{II}$$

Now let $F : V \to \mathbb{R}$ be a strictly convex function on V, and let $\Lambda_F \subset V \times V^*$ be the graph of its Legendre transform. By identifying $V \times V^* \cong T^*V$, the graph Λ_F becomes the image of the map:

$$dF : V \to T^*V,$$

by giving $V \times V^*$ a symplectic structure using ω_I, Λ_F becomes a Lagrangian submanifold of $V \times V^*$, i.e. $d\omega_I = 0$ on Λ_F. It continues to be Lagrangian even if on replaces ω_I by ω_{II}.
Let $pr_1 : \Lambda_F \to V$ and $pr_2 : \Lambda_F \to V_s^*$ be the restriction of the projection $V \times V^* \to V$ and $V \times V^* \to V$ to Λ_F respectively. Since Λ_F is a graph, pr_1 and pr_2 are diffeomorphisms.

Let $i : \Lambda_F \to V \times V^*$ be the inclusion. We claim that

$$i^*\alpha_I = d(pr_1)^*F,$$

which determines F up to an additive constant since pr_1 is a diffeomorphism.

Let $(x,l),(x_0,l_0) \in V \times V_s^*$, we have

$$i^*\alpha_I|_{(x_0,l_0)}(x,l) = l_0 d\tilde{x}(x,l) = l_0(x),$$

but l_0 is the image of L_F at the point x_0. Hence $l_0 = dF|_{x_0} = L_F(x_0)$. On the other hand, we have

$$d((pr_1)^*F)|_{(x_0,l_0)}(x,l) =, dF|_{pr_1(x_0,l_0)}\big(Dpr_1\big)(x,l) = dF|_{(x_0)}(x) = l_0(x).$$

Thus, the claim is proven.

To prove that the Legendre transformation associated with F^* is the inverse of the Legendre transformation associated to F, it suffices to show that F^* satisfies the transpose identity

$$i^*\alpha_{II} = d(pr_2)^*F^*.$$

However, we have that

$$i^*\alpha_{II} = di^*B - i^*\alpha_I = d(i^* - (pr_1)^*F)$$

and F^* satisfies

$$(pr_2)^*F^* = i^*B - (pr_1)^*F$$

by definition. □

We have now a way to pass from a vector space to its dual space. What we need to do now is to work fiberwise:
Let M be an n-dimensional smooth manifold and $F : TM \to \mathbb{R}$ smooth function such that for all $x \in M$:

$$F_x : T_xM \to \mathbb{R},$$

is a strictly convex superlinear function.
Let us now fix $x \in M$ and $u \in T_xM$. As seem above there is a natural identification of the cotangent bundle of a vector space and the cotangent bundle of its dual, hence $T_u^*T_xM \cong T_x^*M$. We define the *Legendre transform at* $x \in M$:

$$\begin{array}{rcl} L_{F_x} : T_xM & \to & T_u^*T_xM \cong T_x^*M \\ u & \mapsto & \big(d(F_x)\big)_u \end{array}$$

Let us now go to local coordinates $(x^1,\dots,x^n,u^1,\dots,u^n) \in T_xM$, the Legendre transform of F looks like:

$$d(F_x)_u = \sum_{j=1}^{n} \frac{\partial F}{\partial u^j} dx^j.$$

Definition 3.1.13. Let M be a smooth manifoldand $F : TM \to \mathbb{R}$ a smooth function. The *fiberwise derivative* of F at $(x, v) \in T_xM$ is the element of T^*_xM given by

$$d(F(x,v))_v = \partial_v F(x,v) = \sum_{j=1}^{n} \frac{\partial F}{\partial v^j}(x,v)dx^j$$

Now, we can finally define the Legendre-dual.

Definition 3.1.14. The *Legendre-dual* of a smooth map $F : TM \to \mathbb{R}$ is the map $F^* : T^*M \to \mathbb{R}$

$$F^* \circ dF(v) := dF(v)v - F(v), \ \forall v \in TM.$$

One now has to check that the Legendre transform and hence the Legendre-dual depends smoothly on the basepoint x chosen in M, but this is a consequence of the Picard-Lindelöf theorem which states that the solution of an ODE depend smoothly of the initial conditions.

3.2 From Lagrangian to Hamiltonian

Definition 3.2.1. Let M be a smooth manifold, a *Riemannian metric* $\langle \cdot, \cdot \rangle$: $TM \times TM \to \mathbb{R}$ is a 2-form on M such that for all $q \in M$, $\langle \cdot, \cdot \rangle_q : T_qM \times T_qM\mathbb{R}$ is a metric on T_qM.

Let M be a closed (i.e. compact) manifold, let us fix a Riemannian metric $\langle \cdot, \cdot \rangle$ on that manifold.

Definition 3.2.2. Let M be a smooth manifold, $F : M \to \mathbb{R}$ a smooth function and $x \in M$. The fiberwise Hessian of F is defined fiberwise in the following way: Let $v, w \in T_xM$, the Hessian at x is

$$\frac{d^2}{dt^2}F(tv)(x)[w] \underbrace{=}_{\text{loc. coord.}} \sum_{j,h=1}^{n} \frac{\partial^2 F}{\partial v^i \partial v^j}(x,v)w^i w^j.$$

Definition 3.2.3. We say that a smooth Lagrangian $l : \mathbb{R} \times TM \to \mathbb{R}$ is *Tonelli* when it satisfies the following conditions:

(T1) the fiberwise Hessian of l is positive-definite, i.e.,

$$\sum_{j,h=1}^{n} \frac{\partial^2 l}{\partial v^j \partial v^h}(t,q,v)w^i w^j > 0,$$

for all $(t, q, v) \in \mathbb{R} \times TM$ and $w = \sum_{j=1}^{n} w^j \frac{\partial}{\partial w^j} \in T_qM$ with $w \neq 0$.

(T2) l is fiberwise superlinear, i.e.,

$$\lim_{|v|_q \to \infty} \frac{l(t,q,v)}{|v|_q} = \infty,$$

for all $(t,q) \in \mathbb{R} \times M$.

Definition 3.2.4. The *Legendre transform* associated to l is the map $\text{Leg}_l : \mathbb{R} \times TM \to \mathbb{R} \times T^*M$ given by

$$\text{Leg}_l(t,q,v) = (t,q,\partial_v l(t,q,v)), \ \forall (t,q,v) \in \mathbb{R} \times TM.$$

Definition 3.2.5. We can now define the *Legendre-dual* Hamiltonian $h : \mathbb{R} \times T^*M \to \mathbb{R}$ associated to l as

$$h \circ \text{Leg}_l(t,q,v) := \partial_v l(t,q,v)v - l(t,q,v), \ \forall (t,q,v) \in \mathbb{R} \times TM.$$

Let $\mathcal{A}_h$ be the Hamilton action associated to h and $\mathcal{S}_l$ be the action associated to l, with l strictly convex and superlinear. Then if l and h are Legendre-dual of each other, the Hamiltonian action and the Action for the Lagrangian are equal. Let $\Gamma = (\gamma, \rho) : I \to T^*M$ be a solution to that Hamilton equations of the Hamilton system (M,h) (so that $\gamma : I \to M$ is a solution to the Euler-Lagrange equation of l), then:

$$\begin{aligned}\mathcal{A}_h(\Gamma) &= \int_a^b \big(\Gamma^*\lambda - h_t(\Gamma(t))\big)dt = \int_a^b \big(\rho(t)[\dot\gamma(t)] - h_t(\gamma(t),\rho(t))\big)dt \\ &= \int_a^b l_t(\gamma(t),\dot\gamma(t))dt = \mathcal{S}_l(\gamma).\end{aligned}$$

3.3 Noether's Theorem - From the Lagrangian to the Hamiltonian formalism

Let M be a smooth n-dimensional manifold equipped with a Riemmanian metric and let $l : TM \times \mathbb{R} \to \mathbb{R}$ be a Tonelli Lagrangian. On top of that, suppose l is admits the one-parameter groups of diffeomorphism generated by the action μ of a Lie group G and that $h : T^*M \to \mathbb{R}$ is the Legendre dual Hamiltonian associated to l.
We will show that under these conditions, the conserved quantity $\mathcal{I}$ of l in theorem 2.1.8 and J of h in theorem 2.2.10 are related to each other.

Definition 3.3.1. Let $\mu_g : M \to M$ denote the diffeomorphism $\mu_g(x) := \mu(g,x)$. We call $\mu_g^\sharp$ the *lift* of μ_g on the tangent space, it is defined by

$$\begin{aligned}\mu_g^\sharp &: TM \to TM \\ &(q,v) \mapsto (\mu_g(q), (d\mu_g)_q(v)).\end{aligned}$$

As μ_g is a diffeomorphism, we can invert its differential. This enables us to define a dual lift $\mu_g^\flat$ of μ_g on the cotangent space in the following way,

$$\begin{aligned}\mu_g^\flat : T^*M &\to T^*M \\ (q,\xi) &\mapsto (\mu_g(q), \xi \circ (d\mu_g)^{-1}_{\mu_g(q)}).\end{aligned}$$

Theorem 3.3.2. *The lift $\mu^\flat$ of the action μ of the Lie group G on M preserves the tautological 1-form λ.*

Proof. Let λ denote the tautological 1-form on T^*M. In a chart about a point $(x,\xi) \in T^*M$, we get the following

$$\begin{aligned}(\mu_g^\flat)^*\lambda &= \sum_{i,j=1}^n \xi_i \left[(d\mu_g)^{-1}_{\mu_g(x)}\right]^i_j d\mu_g^j(x) \\ &= \sum_{i,j=1}^n \xi_i \underbrace{\left(\frac{\partial \mu_g^i(y)}{\partial y^j}\right)^{-1}_{\mu_g(x)} \left(\frac{\partial \mu_g^j(x)}{\partial x^k}\right)}_{\delta^j{}_k} dx^k \\ &= \sum_{i=1}^n \xi_i dx^i = \lambda.\end{aligned}$$

□

Starting now, we will do a slight abuse of notation and call both $\mu^\sharp$ and $\mu^\flat$ simply μ, and keep in mind whenever we are working on TM or T^*M.

Corollary. *The canonical symplectic form ω on T^*M is invariant under the action μ of the Lie group G. Hence for every $g \in G$, the diffeomorphism μ_g : $\mu(g,\cdot)$ is a symplectomorphism.*

Proof. We have $\omega = d\lambda$, but as $\mu^*\lambda = \lambda$ and knowing that the pullback commutes with the exterior derivative, we get

$$\mu^*\omega = \mu^* d\lambda = d\mu^*\lambda = d\lambda = \omega.$$

□

Lemma 3.3.3. *If l is invariant under the action $\mu : G \times M \to M$, then the fiberwise derivative of l is invariant under the pullback of that action.*

Proof. For $(x,v) \in TM$, we have

$$\mu^* d(l(x,v))_v = d\mu^* l(x,v)_v = d(l(x,v))_v.$$

Thus the fiberwise derivative is invariant under the Lie group action. □

Corollary. *The Legendre-dual h is invariant under the action of μ of the Lie group G.*

Proof. Let us recall the definition of the Legendre-dual h associated to l:

$$h \circ \mathrm{Leg}_l(t,q,v) := \partial_v l(t,q,v)v - l(t,q,v),\ \forall (t,q,v) \in \mathbb{R} \times TM.$$

Every term on the right hand side is invariant under the action μ, hence the left hand side has to be too. □

Now all that is left to do is to relate $\mathcal{I}$ and J, let us recall the two theorems:

Definition. A Lagrangian system (M,l) *admits* a one-parameter group of diffeomorphism θ_s if the following holds:

$$\forall s \in I, \forall (q,v) \in TM:\ l_t(\theta_s(q), \theta_s^*(v)) = l_t(q,v)$$

Theorem. *(Noether's Theorem) Let (M,l) be a Lagrangian system which admits a one parameter group of diffeomorphism θ_s.*
The Euler-Lagrange equation to l has a first integral:

$$\mathcal{I} : TM \to \mathbb{R},\ \textit{in local coordinates } (q,v) \in TM : \mathcal{I}(q,v) = \frac{\partial}{\partial v} l_t \left.\frac{\partial}{\partial s}\right|_{s=0} \theta_s(q).$$

Theorem. *(Noether's Theorem). Let θ_s be a one-parameter group of symplectomorphisms and suppose that for all $s \in \mathbb{R}$, $h : T^*M \to \mathbb{R}$ is invariant under θ_s. Then the momentum map $J : M \to \mathfrak{g}^*$ is an integral for X_h, i.e. if φ_h is the flow of X_h, then:*

$$J(\varphi_h(q)) = J(q),\ \forall q \in M.$$

Let $\xi \in \mathfrak{g}$ and let the Lagrangian l be invariant under the action μ of a Lie group G. Then infinitesimal generator of the action is defined as $\xi_M(q) = \left.\frac{\partial}{\partial t}\right|_{t=0} \mu(\exp(t\xi), q)$. By setting $\theta_s(q) := \mu(\exp(t\xi), q)$, we get

$$\mathcal{I}(q,v) = \frac{\partial}{\partial v} l_t \left.\frac{\partial}{\partial s}\right|_{s=0} \theta_s(q) = \frac{\partial}{\partial v} l_t(\xi_M(q)).$$

Now let us look at the second Noether theorem. Let $\hat{J}(\xi)(q) = J(q)\xi$, we get the following property for $\hat{J}$:

$$d\hat{J}(\xi) = i_{\xi_M}\omega.$$

But we only work with the tautological 1-form λ, hence $\omega = d\lambda$. Using the Cartan formula, we get

$$i_{\xi_M}\omega = i_{\xi_M} d\lambda = \mathcal{L}_{\xi_M}\omega - d i_{\xi_M}\lambda = -d\lambda(\xi_M).$$

The Lie derivative of ω vanishes as the flow induced by ξ_M is a symplectomorphism.
Hence $\hat{J}(\xi) = -\lambda(\xi_M) + df$, for some $f \in C^\infty(M)$.

Since l is Tonelli, the Legendre transform of l induces a fiberwise bijection between TM and T^*M. Thus for all $p \in T_q^*M$ there exists a $v \in T_qM$ such that $L_l(v) = p$.

What does this implies? Let us look at the map $\hat{J}$ in more details.

Let's pick a basis of T_qM such that each $p \in T_q^*M$ is the image of the Legendre transform applied to a basis vector. Hence we get

$$\hat{J}(\xi) = -\lambda(\xi_M) + df = -\sum_i p_i dq^i(\xi_M) + df = -\sum_i \frac{\partial l}{\partial v^i} dq^i(\xi_M) + df = -\mathcal{I}(q, v) + df$$

$$\Rightarrow \hat{J}(\xi) + \mathcal{I}(q, v) = df.$$

Up to an exterior differential of a $C^\infty(M)$ function, the two conserved quantities agree with each other.

What does this mean? This seems to be a really fundamental result. On one hand, we have a completely natural construction, the tautological one form and the momentum map, on the other hand, this very "artificial" construction of picking a specific Lagrangian. Observed under a certain angle, this means that the symmetries of the system are more fundamental that the functional setting we put ourselves in by picking a Lagrangian. Indeed, the equation above means that I can fully "delete" the functional setting by doing just regular symplectic geometry. Under another angle this means that even with a very specific choice of Lagrangian (namely, it needs to be Tonelli and invariant under the Lie group action) one can already extract a lot of information out of a manifold. The results one gets are independent of the choice of the Lagrangian as we have just shown that they have a natural equivalent.

A Differential Geometry

In order study Lagrangian and Hamiltonian systems, we need to start with a manifold. Why a manifold? Usually in physics, systems have constraints, therefore the path of particles are restraint to surfaces within $\mathbb{R}^{3n}$, the so called *configuration space* for n point particles. Hence we will use manifold which enable use to encode these constraints not within the functional setting, but in the geometry of our space.

A.1 Manifolds, smooth maps and tangent spaces

This first part of the thesis is a collection of tools needed to make a rigorous analysis of the Lagrange Hamilton dualism later on. Most topics in this section are taken over from [6].

Definition A.1.1. A *smooth manifold of dimension n* is a pair (M, Σ) where M is a topological manifold of dimension n and Σ is a smooth structure on M.

Definition A.1.2. Let $\varphi : M \to N$ be a continuous map between the two manifolds. We say φ is *smooth* if for every point $x \in M$, if $\sigma : U \to O$ is any chart on M with $x, \in U$ and $\tau : V \to Q$ any chart on N, with $\varphi(x) \in V$, the composition

$$\tau \circ \varphi \circ \sigma^{-1} : \underbrace{\sigma(U \cap \varphi^{-1}(V))}_{\subset \mathbb{R}^n} \to \underbrace{\tau(\varphi(U) \cap V)}_{\subset \mathbb{R}^k}$$

is smooth (in the normal limit sense on euclidean space).

Definition A.1.3. A *smooth function* on a manifold is a smooth map $f : M \to \mathbb{R}$ in the sense that for any chart $\sigma : U \to O$ on M, the composition $f \circ \sigma^{-1} : O \to \mathbb{R}$ is a smooth function in the euclidean sense.

Remark. We denote the set of smooth function on M by $C^\infty(M)$. This space with pointwise addition, scalar multiplication and multiplication forms an *algebra*.

Now suppose $f : O \subset \mathbb{R}^n \to \mathbb{R}^m$ is smooth. If $x \in O$ and $v \in \mathbb{R}^n$, then the vector $Df(x)[v]$ can be thought of as the partial derivative of f in the direction v:

$$Df(x)[v] = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t}.$$

Denote by $u^i : \mathbb{R}^n \to \mathbb{R}$ the function and $e_i \in \mathbb{R}^n$ the i'th basis vector, we have:

$$u^i(x^1, \ldots, x^n) = x^i,$$
$$u^i(e_j) = \delta^i_j.$$

We therefore get the more familiar expression:

$$Df(u^i \circ f)(x)[e_j] = \frac{\partial f^i}{\partial x^j}(x).$$

Definition A.1.4. If $x \in M$ and σ is a chart defined on a neighbourhood of x then we will denote the function $u^i \circ \sigma$ by x^i. We call the x^i the *coordinate* of the chart σ, and we say that the x^i are *local coordinates about* x.

Definition A.1.5. Let M be a smooth manifold and let $x \in M$. Let U and V be two neighbourhoods of x, and suppose $f \in C^\infty(U)$ and $g \in C^\infty(V)$. We say that f and g have the same *germ* at x if there exists a smaller neighbourhood $W \subset U \cap V$ such that

$$f|_W \equiv g|_W .$$

Remark. We denote a germ $\underline{f}$ and let $\mathcal{F}_x(M)$ denote the set of germs at x. This set forms another algebra with pointwise addition, scalar multiplication and multiplication.
A germ at x has a well-defined value at x (although nowhere else), and this gives us the map

$$\text{eval}_x : \mathcal{F}_x(M) \to \mathbb{R},\ \text{eval}_x(\underline{f}) := f(x),$$

where (U, f) is any representative of $\underline{f}$.
Remark. Let $f : O \subset \mathbb{R}^n \to \mathbb{R}$ be smooth, with $O \subset \mathbb{R}^n$ open. Let $x \in O$ and $v \in \mathbb{R}^n$. If one considers (x, v) fixed and let f vary:

$$(x, v) : C^\infty(O) \to \mathbb{R},\ (x, v)(f) := Df(x)[v].$$

As the value of $Df(x)[v]$ only depends on the germ of f at x, we can think of the vector v as a linear map

$$v : \mathcal{F}_x(O) \to \mathbb{R},\ v(\underline{f}) := Df(x)[v].$$

In fact, the map v is not any linear map, it is also a *derivation* in the sense that

$$v(\underline{f} \cdot \underline{g}) = \text{eval}_x(\underline{f}) \cdot v(\underline{g}) + \text{eval}_x(\underline{g}) \cdot v(\underline{g}),$$

which is just a fancy way of expressing the *Leibniz rule.*

Definition A.1.6. Let M be a smooth manifold of dimension n and let $x \in M$. A *tangent vector at* x is a linear map

$$v : \mathcal{F}_x(M) \to \mathbb{R},$$

which satisfies the *derivation property*:

$$v(\underline{f} \cdot \underline{g}) = \text{eval}_x(\underline{f}) \cdot v(\underline{g}) + \text{eval}_x(\underline{g}) \cdot v(\underline{g}).$$

The set of tangent vectors at x in M will be denoted by: T_xM.

Definition A.1.7. Let M be a smooth manifold of dimension n, let $x \in M$ and let U be any neighbourhood of x. A *derivation of* $C^\infty(U)$ *at* x is a linear map $w : C^\infty(U) \to \mathbb{R}$ which satisfies the *derivation property*

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

Definition A.1.8. Let M and N be smooth manifolds and let $\varphi : M \to N$ be a smooth map. Fix $x \in M$ and $v \in T_xM$. We define a tangent vector $w \in T_{\varphi(x)}N$ by setting

$$w(f) := v(f \circ \varphi), \ \forall f \in C^\infty(N).$$

It is pretty simple to check that this indeed is a linear derivation, hence an element of $T_{\varphi(x)}N$. Moreover if we denote w by $D\varphi(x)[v]$, it is immediate that the map $v \mapsto D\varphi(x)[v]$ is a linear map. We call $D\varphi(x)$ the *derivative of* φ *at* x.

Proposition A.1.9. *(The chain rule on manifolds). Let M, N and P be smooth manifolds, and suppose $\varphi : M \to N$ and $\psi : N \to P$ are smooth maps. Then*

$$D(\psi \circ \varphi)(x) = D\psi(\varphi(x)) \circ D\varphi(x).$$

Proof. Take $v \in T_xM$ and $f \in C^\infty(P)$ and unfold definitions. □

Lemma A.1.10. *Let $\varphi : M \to N$ be a smooth manifold between two smooth manifolds, where M has dimension n and N has dimension m. Let $x \in M$, and let $\sigma : U \to O$ be a chart on M about x, and let $\tau : V \to Q$ ba a chart on N around $\varphi(x)$. Denote the local coordinates of σ by (x^i) and the local coordinates of τ by (y^i). Then the matrix of $D\varphi(x)$ with respect to the basis $\{\frac{\partial}{\partial x^i}\big|_x | i = 1, \ldots, n\}$ of T_xM and $\{\frac{\partial}{\partial y^j}\big|_{\varphi(x)} | j = 1, \ldots, m\}$ at $T_{\varphi(x)}N$ is given by the matrix $D(\tau \circ \varphi \circ \sigma^{-1})(\sigma(x))$.*

Proof.

$$\begin{aligned} D\varphi(x)[\frac{\partial}{\partial x^j}\Big|_x] &= \sum_{i=1}^n D\varphi(x)[\frac{\partial}{\partial x^j}\Big|_x]y^i \, \frac{\partial}{\partial y^i}\Big|_{\varphi(x)} \\ &= \sum_{i=1}^n \frac{\partial}{\partial x^j}\Big|_x (y^i \circ \varphi) \, \frac{\partial}{\partial y^i}\Big|_{\varphi(x)} \\ &= \sum_{i=1}^n D(u^i \circ \tau\varphi\sigma^{-1})(\sigma(x))[e_i] \, \frac{\partial}{\partial y^i}\Big|_{\varphi(x)}. \end{aligned}$$

□

Definition A.1.11. A *curve (or path)* in a smooth manifold M is a smooth map $\gamma : (a, b) \to M$, where we think of (a, b) as a one dimensional smooth manifold. Fix $t \in (a, b)$. There are, a priori, two different way of defining an element $\gamma'(t)$ of $T_{\gamma(t)}M$, which we will call the *velocity vector of γ at time t*.

1. Firstly, we can define a derivation on $C^\infty(M)$ at $\gamma(t)$ by setting

$$\gamma'(t)(f) := (f \circ \gamma)'(t), \; f \in C^\infty(M).$$

2. Think of γ as a smooth map between manifolds then we can define a tangent vector $\gamma'(t)$ at $\gamma(t)$ via the derivative $D\gamma(t)$:

$$\gamma'(t) := D\gamma(t)[\left.\frac{\partial}{\partial t}\right|_t] \in T_{\gamma(t)}M.$$

Remark. By going to a chart one sees that the two definitions are equivalent.

Proposition A.1.12. *Let M be a smooth manifold and let $\delta, \gamma : (-\epsilon, \epsilon) \to M$ be two smooth curves such that $\gamma(0) = \delta(0)$. Then $\gamma'(0) = \delta'(0)$ in $T_{\gamma(0)}M$ if and only if for some chart σ defined on a neighbourhood of $\gamma(0)$, we have*

$$(\sigma \circ \gamma)'(0) = (\sigma \circ \delta)'(0).$$

Proof. Go to local coordinate basis of $T_{\gamma(0)}M$ □

Remark. Hence we get yet another way of defining the tangent space T_xM of a manifold: a tangent vector at $x \in M$ is an equivalence class of smooth curves a $\gamma : (-\epsilon, \epsilon) \to M$ such that $\gamma(0) = x$, where $\gamma \sim \delta$ if and only if for some chart the previous condition holds.

Proposition A.1.13. *Let $\varphi : M \to N$ be a smooth map between two smooth manifolds and let $\gamma : (a, b) \to M$ be a curve in M. Then*

$$D\varphi(x)[\gamma'(t)] = (\varphi \circ \gamma)'(t).$$

Proof. Unfold definitions. □

Definition A.1.14. Let M be a smooth manifold of dimension n and let $x \in M$. We denote the dual vector space $\mathrm{Hom}(T_xM, \mathbb{R})$ by T_x^*M and call it the *cotangent* space of M at x.

Remark. Let U be a neighbourhood of x and $f \in C^\infty(U)$. Then f defines an element $\left.df\right|_x \in T_x^*M$ by

$$\left.df\right|_x(v) := v(f), \; v \in T_xM.$$

Thus $df|_x$ is a linear function $T_xM \to \mathbb{R}$. In contrast, the differential $Df(x)$ is a linear function $T_xM \to T_{f(x)}\mathbb{R}$. Under the identification $T_{f(x)}\mathbb{R} \equiv \mathbb{R}$ these become the same map:

$$Df(x)[v] = df|_x(v) \cdot \left.\frac{\partial}{\partial t}\right|_{f(x)}.$$

Proposition A.1.15. *Let M be a smooth manifold of dimension n and let $x \in M$. Let $\sigma : U \to O$ be a chart about x with local coordinates $x^i = u^i \circ \sigma \in C^\infty(U)$. Then $\{ dx^i|_x \}$ is a basis of T_x^*M.*

Proof. $dx^j|_x (\frac{\partial}{\partial x^i}|_x) = \delta_i^j$. □

Definition A.1.16. Let M be a smooth manifold. The *tangent bundle* of M is the disjoint union ot the tangent spaces:

$$TM = \bigsqcup_{x \in M} T_xM.$$

We denote an element in TM as a pair (x, v) to indicate that $v \in T_xM$. There is a map $\pi : TM \to M$ given by $\pi(x, v) = x$. We call π the *footpoint map.*

Definition A.1.17. Let M be a smooth manifoldand TM its tangent bundle. Let $x \in M$, the set $\pi^{-1}(x)$ is called the *fibre over the point x*. By construction $\pi^{-1}(x) \equiv T_xM$.

Theorem A.1.18. *Let M be a smooth manifold of dimension n. The smooth structure on M naturally induces a smooth structure on TM, making TM into a smooth manifold of dimension $2n$. Moreover the map $\pi : TM \to M$ is smooth.*

Proof. Let $\Sigma = \{\sigma_a : U_a \to O_a | a \in A\}$ be a smooth atlas on M. Write $x_a^i = u^i \circ \sigma_a$ for the local coordinates of σ_a. We build a chart on TM:

$$\tilde{\sigma}_a(x, v) = \Big(\sigma_a(x),\ \sum_{i=1}^{n} dx_a^i\big|_x(v) \cdot e_i\Big),\ x \in U_a, v \in T_xM.$$

This does the job. □

Definition A.1.19. Let $\varphi : M \to N$ be a smooth map between two smooth manifolds. Define the *derivative of φ* to be the map

$$D\varphi : TM \to TN,\ D\varphi(x, v) := D\varphi(x)[v].$$

Definition A.1.20. Let M be a smooth manifold. The *cotangent bundle* of M is the disjoint union of the cotangent spaces:

$$T^*M = \bigsqcup_{x \in M} T_x^*M.$$

We denote an element of T^*M as (x, p) to indicate that $p \in T_x^*M$.

Remark. In an analogous way, one can define a canonical smooth structure on T^*M, making it into a $2n$ dimensional manifold and a smooth footpoint map $\pi^* : T^*M \to M$.

Remark. One can also call $\pi^{-1}(x)$ for $x \in M$ the fibre over the point x. Hence one has to be a bit careful and see if the footpoint maps it defined on the tangent or cotangent bundle, although in most cases, this should not be particularly difficult.

A.2 Vector fields, their algebra and flows

Definition A.2.1. Let M be a smooth manifold and $W \subset M$ be a non-empty open set (possibly equal to all M). A *vector field* X on W is a smooth map $X : W \to TM$ (where we regard W as a smooth manifold) that satisfies the *section property*:

$$\pi(X(x)) = x, \ \forall x \in W,$$

where $\pi : TM \to M$ is the footpoint map.
We denote by $\mathfrak{X}(W)$ the set of all vector fields on W.

Remark. Let $\sigma : U \to O$ be a chart on M. Suppose $X : U \to TM$ is a vector field. Since $\{ \frac{\partial}{\partial x^i}\big|_x \mid i = 1, \dots, n\}$ is a basis of T_xM, we can write

$$X(x) = X^i(x) \left.\frac{\partial}{\partial x^i}\right|_x .$$

The X^i define smooth functions $X^i : U \to \mathbb{R}$.
There is another way to think about it. Suppose $f \in C^\infty(U)$ and think of $X(x)$ as a derivation of $C^\infty(U)$. This gives us a function $X(f) : U \to \mathbb{R}$:

$$X(f)(x) := X(x)(f), \ \forall x \in U.$$

As X is smooth, the function $X(f)$ is smooth too.

Proposition A.2.2. *Let M be a smooth manifoldand $W \subset M$ be a non-empty open set. Let $X : W \to TM$ be any function satisfying the section property.*
Then the following are equivalent:

(i) $X \in \mathfrak{X}(W)$.

(ii) *If $\sigma : U \to O$ any chart on M with $U \subset W$, then the functions X^i defined above belong to $C^\infty(U)$.*

(iii) *If $V \subset W$ any open set and $f \in C^\infty(V)$ then the function $X(f)$ belongs to $C^\infty(V)$.*

Proof. We first show $(i) \Leftrightarrow (ii)$. Let $x \in W$, and let $\sigma : U \to O$ ba a chart about x. X^i is smooth if and only if $X^i \circ \sigma^{-1}$ is smooth, but we have that $X^i \circ \sigma^{-1}$ is the function

$$z \mapsto dx^i|_{\sigma^{-1}(z)}\Big(X(\sigma^{-1}(z))\Big),\ z \in O.$$

Let us define the chart $\tilde{\sigma}$ on TM by

$$\tilde{\sigma}(x,v) = (\sigma(x), dx^i|_x(v)e_i),\ x \in U,\ v \in T_xM.$$

By definition, X is smooth at x if and only if the composition

$$\tilde{\sigma} \circ X \circ \sigma$$

is smooth in the normal sense. Explicitly, this is the map

$$O \to O \times \mathbb{R}^n,\ z \mapsto (z, dx^i|_{\sigma(z)}(X(\sigma(z))e_i).$$

Hence, we see that this map is smooth if and only if $X^i \circ \sigma^{-1}$ is smooth for each $i = 1, \dots, n$. Let us now prove $(ii) \Rightarrow (iii)$: Let $V \subset W$ and $f \in C^\infty(V)$. Choose a chart $\sigma : U \to O$ with $U \subset V$. Then for $x \in U$, we have that

$$X(f)(x) = X^i(x)\ \frac{\partial}{\partial x^i}\bigg|_x (f).$$

The function $x \mapsto \frac{\partial}{\partial x^i}\big|_x (f)$ is smooth. By (ii) we have that the X^i are smooth, hence the whole is smooth as the pointwise product of smooth functions.
Finally, $(iii) \Leftrightarrow (ii)$. Indeed, since if $\sigma : U \to O$ is a chart about x with local coordinates x^i, then the functions X^i are simply $X(x^i)$, where we think of x^i as elements of $C^\infty(U)$. □

Remark. Suppose $\sigma : U \to O$ is a chart on M with local coordinates $x^i = u^i \circ \sigma$, we can think of $\frac{\partial}{\partial x^i}$ as defining a vector field on U via:

$$\frac{\partial}{\partial x^i}(x) := \frac{\partial}{\partial x^i}\bigg|_x.$$

Definition A.2.3. If $f \in C^\infty(U)$ then we denote the function $\frac{\partial}{\partial x^i}(f)$ by

$$\frac{\partial f}{\partial x^i}(x) := \frac{\partial}{\partial x^i}(f) = D(f \circ \sigma^{-1})(\sigma(x)).$$

Remark. Let $W \subset M$ be an open subset, the space $\mathfrak{X}(W)$ forms a module on the ring $C^\infty(W)$.

Definition A.2.4. Let $X,\ Y \in \mathfrak{X}(W)$. Then the commutator $[X,Y] := X \circ Y - Y \circ X$ is another derivation. We call $[X,Y]$ the *Lie bracket* of X and Y.

Proposition A.2.5. *Let $\sigma : U \to O$ be a chart on M with local coordinates x^i, and let $X,\ Y \in \mathfrak{X}(U)$. Write $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^i \frac{\partial}{\partial x^i}$. Then*

$$[X,Y] = \sum_{i,j=1}^{n} \Big(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \Big) \frac{\partial}{\partial x^j}.$$

Definition A.2.6. A (real) *Lie algebra* is a vector space $\mathfrak{g}$ endowed with a bilinear operation called *Lie bracket*

$$\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g},\ (x,y) \mapsto [x,y]$$

which in addition is antisymmetric, $[x,y] = -[y,x]$ and satisfies the *Jacobi identity*

$$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0,\ \forall x,y,z \in \mathfrak{g}.$$

The *dimension* of the Lie algebra $\mathfrak{g}$ is simply its dimension as a vector space.

Theorem A.2.7. *Let M be a smooth manifoldand let $W \subset M$ be a non-empty open subset. Then $\mathfrak{X}(W)$ is a Lie algebra.*

Definition A.2.8. Suppose $\varphi : M \to N$ is a diffeomorphismand $X \in \mathfrak{X}(M)$. We define the *pushforward* vector field $\varphi_\star(X) \in \mathfrak{X}(N)$ by defining

$$\varphi_\star\big(X\big)(y) := D\varphi(\varphi^{-1})[X(\varphi^{-1}(y))].$$

Remark. $\varphi_\star(X) : N \to TN$ is smooth because the composition

$$N \xrightarrow{\varphi^{-1}} M \xrightarrow{X} TM \xrightarrow{D\varphi} TN$$

is a composition of smooth maps, hence smooth.
The map $\varphi_\star$ is a module homomorphism of $\mathfrak{X}(W)$ over $C^\infty(W)$. (One especially has $\varphi_\star(fX) = \varphi_\star(f)\varphi_\star(X),\ \forall X \in \mathfrak{X}(W),\ \forall f \in C^\infty(W)$)

Definition A.2.9. Let $\mathfrak{g}$ and $\mathfrak{h}$ be two Lie algebras. A *Lie algebra homomorphism* is a linear map $L : \mathfrak{g} \to \mathfrak{h}$ which respects the Lie bracket, i.e.

$$[Lx, Ly]_{\mathfrak{h}} = L[x,y]_{\mathfrak{g}},\ \forall x,y \in \mathfrak{g}.$$

A *Lie algebra isomorphism* is a bijective Lie algebra homomorphism whose inverse is also a Lie algebra homomorphism.

Proposition A.2.10. *Let $\varphi : M \to N$ be a diffeomorphism. Then $\varphi_\star : \mathfrak{X}(M) \to \mathfrak{X}(N)$ is a Lie algebra isomorphism.*

Definition A.2.11. Let M be a smooth manifoldand let $X \in \mathfrak{X}(M)$. Let $(a, b) \subset \mathbb{R}$ be an interval, and suppose $\gamma : (a, b) \to M$ is a smooth map. We say that γ is an *integral curve* of X if

$$\gamma'(t) = X(\gamma(t)), \ \forall t \in (a, b).$$

Definition A.2.12. Let $X \in \mathfrak{X}(M)$. Given a point $x \in M$, we denote by $(t^-(x), t^+(x))$ the maximal interval around 0 such that the integral curve $\gamma_x : (t^-(x), t^+(x)) \to M$ of X whose initial condition $\gamma_x(0) = x$ is defined. We call γ_x the *maximal integral curve through* x.

Theorem A.2.13. *(Maximal flow). Let M be a smooth manifoldand let $X \in \mathfrak{X}(M)$. There exists a unique open set $\mathcal{D} \subset \mathbb{R} \times M$ and a unique smooth map θ such that*

(i) For all $x \in M$ one has

$$\mathcal{D} \cup (\mathbb{R} \times \{x\}) = (t^-(x), t^+(x)) \times \{x\}.$$

(ii) $\theta(t, x) = \gamma_x(t), \ \forall (t, x) \in \mathcal{D}$.

We call θ the *flow* if X.

A.3 Lie groups and Lie algebras

Remark. We write Diff(M) to denote the set of diffeomorphisms of M. Diff(M) is actually a group under composition, where the identity is just the identity map.

Definition A.3.1. A *one-parameter group of diffeomorphisms* is a group homomorphism $\mathbb{R} \to \text{Diff}(M)$. Writing this as $t \mapsto \theta_t$, the group property tells us that

$$\theta_0 = \text{id}, \ \theta_s \circ \theta_t = \theta_{s+t}, \ \forall s, t \in \mathbb{R}.$$

If $\{\theta_t\}$ is a one-parameter group of diffeomorphisms, then we define its *infinitesimal generator* as the (necessarily complete) vector field

$$X(x) := D\theta(t, x)\left[\left.\frac{\partial}{\partial t}\right|_{(0,x)} \right].$$

Definition A.3.2. Let $X \in \mathfrak{X}(M)$ with flow θ_t. We define the *Lie derivative* of X to be the map

$$\mathcal{L}_X : C^\infty(M) \to C^\infty(M)$$

given by

$$(\mathcal{L}_X f)(x) := \lim_{t \to 0} \frac{f \circ \theta_t(x) - f(x)}{t}.$$

Lemma A.3.3. $\mathcal{L}_X f = X(f)$

Proof. From the definitions one has

$$X(f)(x) = X(x)(f) = \gamma_x'(0)(f) = (f \circ \gamma_x)'(0).$$

But then clearly

$$(f \circ \gamma_x)'(0) = \lim_{t \to 0} \frac{f \circ \gamma_x(t) - \gamma(x)}{t} = \lim_{t \to 0} \frac{f \circ \theta_t(x) - f(x)}{t}.$$

□

We will see how the Lie derivative is defined on vector fields and give an explicit formula, but first we need a preliminary result:

Lemma A.3.4. *Let $U \subset M$ be open and let $a < 0 < b$. Let $f : (a,b) \times U \to \mathbb{R}$ be a smooth function such that $f(0,x) = 0$ for all $x \in U$. Then there exists another smooth function $h : (a,b) \times U \to \mathbb{R}$ such that*

$$f(t,x) = th(t,x), \ \left.\frac{\partial}{\partial t}\right|_{(0,x)} (f) = h(0,x), \ \forall (t,x) \in (a,b) \times U.$$

Here $(t,x) \to \left.\frac{\partial}{\partial t}\right|_{(t,x)} := Di_x(t[\frac{\partial}{\partial t}])$ is the vector field on $(a,b) \times U$.

Proof. Define $h(t_0,x) := \int_0^1 \left.\frac{\partial}{\partial t}\right|_{(st_0,x)} (f)ds$.
To see that $f(t,x) = th(t,x)$, consider the curve $\gamma(s) := f \circ i_x(st)$. Then

$$f(t,x) = f(t,x) - f(0,x) = \gamma(1) - \gamma(0) = \int_0^1 \gamma'(s)ds.$$

But by definition $\gamma'(s) = \left.\frac{\partial}{\partial t}\right|_{(st_0,x)} (f)$. □

Definition A.3.5. Let $X \in \mathfrak{X}(M)$ with flow θ_t. We define the *Lie derivative* of X to be the map

$$\mathcal{L}_X : \mathfrak{X}(M) \to \mathfrak{X}(M)$$

given by

$$(\mathcal{L}_X Y)(x) := \lim_{t \to 0} \frac{D\theta_{-t}(\theta_t(x))[Y(\theta_t(x))] - Y(x)}{t}.$$

Theorem A.3.6. *For any two vector fields on M, one has $\mathcal{L}_X(Y) = [X,Y]$.*

Proof. Fix $x \in M$. By the local flow theorem, there exists $a < 0 < b$ and a neighbourhood U of x such that $(a,b) \times U \subset \mathcal{D}$, the domain of θ. Now fix $g \in C^\infty(M)$. We apply the previous lemma to the function $f(t,y) := g(\theta_t(y)) - g(y)$ to obtain a function h, writing $h_t(y) := h(t,y)$, we have:

$$g \circ \theta_t = g + th_t, \; h_0 = X(g),$$

Thus for another vector field Y, we have

$$\begin{aligned} D\theta_{-t}(\theta(x))[Y(\theta_t(x))](g) &= Y(\theta_t(x))(g \circ \theta_{-t}) \\ &= Y(\theta_t(x))(g - th_{-t}) \\ &= Y(g) \circ \theta_t(x) - tY(h_{-t}) \circ \theta_t(x). \end{aligned}$$

Hence,

$$\begin{aligned} (\mathcal{L}_X Y)(g)(x) &= \lim_{t \to 0} \frac{Y(g) \circ \theta_t(x) - (Y(g))(x)}{t} - \lim_{t \to 0} Y(h_{-t}) \circ \theta_t(x) \\ &= \mathcal{L}_X(Y(g))(x) - Y(h_0)(x) \\ &= X(Y(g))(x) - Y(X(g))(x) \\ &= [X,Y](g)(x). \end{aligned}$$

Since x and g are arbitrary this completes the proof. □

Proposition A.3.7. *Let $X, Y \in \mathfrak{X}(M)$ with flow θ_t^X and θ_t^Y respectively. Then $[X,Y] = 0$ if and only if the two flows commute, i.e. $\forall s, t$ small : $\theta_s^Y \circ \theta_t^X = \theta_t^X \circ \theta_s^Y$.*

A.3.1 Lie group

Definition A.3.8. A *Lie groups* G is a smooth manifold that is also a group in the algebraic sense, with the property that group multiplication

$$m : G \times G \to G, \; m(a,b) = ab,$$

and group inversion

$$i : G \to G, \; i(a) = a^{-1},$$

are both smooth maps.

Definition A.3.9. A *Lie group homomorphism* $\varphi : G \to H$ is a smooth map which is also a group homomorphism. A *Lie group isomorphism* is a Lie group homomorphism which is also a diffeomorphism, and whose inverse is also a Lie group homomorphism.

Definition A.3.10. Let G be a Lie group and let $a \in G$. Let $l_a : G \to G$ and $r_a : G \to G$ denote the *left translation* and *right translation* by a respectively

$$l_a(b) := ab, \; r_a(b) = ba.$$

These maps are both diffeomorphisms. For instance $l_a = m \circ i_a$, where $i_a : G \to G \times G$ it the map $b \mapsto (a, b)$, and hence is the composition of smooth maps. Moreover $l_{a^{-1}}$ is the inverse of l_a. Similar for r_a.

Proposition A.3.11. *Every Lie group homomorphism has constant rank.*

Proof. Let $\varphi : G \to H$ be a Lie group homomorphism. Fix $a \in G$. For all $b \in G$ one has

$$\varphi(l_a(b)) = \varphi(a)\varphi(b) = l_{\varphi(a)}(\varphi(b)),$$

that is,

$$\varphi \circ l_a = l_{\varphi(a)} \circ \varphi.$$

Differentiate at e and the job is done. □

Corollary. *A Lie group homomorphism is a Lie group isomorphism if and only if it is bijective.*

Definition A.3.12. Let G be a Lie group. We define the *Lie algebra* of G, which we will usually write as $\mathfrak{g}$, is simply the tangent space to G at the identity element e:

$$\mathfrak{g} := T_e G.$$

Definition A.3.13. Let G be a lie group. A vector field $X \in \mathfrak{X}(G)$ is said to be *left-invariant* if $(l_a)_\star(X) = X$ for all $a \in G$. Equivalently, this means that

$$Dl_a(b)[X(b)] = X(ab), \; \forall a, b \in G.$$

We denote by $\mathfrak{X}_l(G) \subset \mathfrak{X}(G)$ the set of left invariant vector fields.

Proposition A.3.14. *Let G be a Lie group and let $X, Y \in \mathfrak{X}_l(G)$. Then $[X, Y]$ also belongs to $\mathfrak{X}_l(G)$. Thus, $\mathfrak{X}_l(G)$ is a Lie subalgebra of $\mathfrak{X}(G)$.*

Proof. Use the properties of the pushforward $((l_a)_\star)$ and the Lie bracket. □

Theorem A.3.15. *Let G be a Lie group with Lie algebra $\mathfrak{g} = T_e G$. The evaluation map*

$$eval_e : \mathfrak{X}_l(G) \to \mathfrak{g}, \; eval_e(X) = X(e)$$

is a vector space isomorphism. Thus $\mathfrak{X}_l(G)$ is a vector space of the same dimension as $\mathfrak{g}$.

Proof. The map eval_e is clearly linear. If $\mathrm{eval}_e(X) = 0$, then X is identically zero, since for any $a \in G$ one has

$$X(a) = Dl_a(e)(X(e)) = 0.$$

Thus we need to show that eval_e is surjective. Fix $v \in \mathfrak{g}$. Define that map $X_v : G \to TG$ by

$$X_v(a) := Dl_a(e)[v].$$

Then X_v satisfies the section property. We still need to show that our map is smooth, we will show $X_v(f)$ is smooth for any $f \in C^\infty(G)$. Choose a curve $\gamma : (-\varepsilon, \varepsilon) \to G$ such that $\gamma(0) = e$ and $\gamma'(0) = v$. Then

$$\forall a \in G : \; X_v(f)(a) = X_v(a)(f) = Dl_a(e)[v](f) = v(f \circ l_a) = (f \circ l_a \circ \gamma)'(0).$$

The curve $f \circ l_a \circ \gamma$ is given by $t \mapsto f(m(a, \gamma(t)))$. Since f, γ and m are all smooth, this smooth. Finally, we need to show that X_v is left-invariant. Indeed if $a, b \in G$ then

$$Dl_a(b)[X_v(b)] = Dl_a(b) \circ Dl_b(e)[v] = D(l_a \circ l_b)(e)[v] = Dl_{ab}(e)[v] = X_v(ab).$$

Since $\mathrm{eval}_e(X_v) = X_v(e) = v$, this shows eval_e is surjective. □

Corollary. *Let G be a Lie group of dimension n. Then its Lie algebra is a Lie algebra of dimension n.*

Proof. Define a a Lie bracket on $\mathfrak{g}$, set

$$[v, w] := \mathrm{eval}_e[X_v, X_w], \; v, w \in \mathfrak{g}.$$

□

Proposition A.3.16. *Let G be a Lie group and let $X \in \mathfrak{X}_l(G)$. Then X is complete.*

Proof. By the local flow theorem, there exists some $\epsilon > 0$ such that the integral curve $\gamma_e(t)$ of X with initial conditions e exists on $(-\epsilon, \epsilon)$. Now observe that $l_a \circ \gamma_e$ is an integral curve of X starting at a. Hence equal to γ_a. Thus γ_a is also defined on $(-\epsilon, \epsilon)$. □

A.3.2 The exponential map

Definition A.3.17. Let G be a Lie group with Lie algebra $\mathfrak{g}$. A *one-parameter subgroup* of G is a homomorphism $\mathbb{R} \to G$.

Proposition A.3.18. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $v \in \mathfrak{g}$, and let $X_v \in \mathfrak{X}_l(G)$ denote the unique left-invariant field with $X_v(e) = v$ (defined as in the proof of theorem A.3.15). Let $\gamma^v = \gamma^v_e : \mathbb{R} \to G$ denote the integral curve of X_v with $\gamma^v(0) = e$ (by the previous proposition, this is defined on all $\mathbb{R}$). Then γ^v is a one-parameter subgroup. Moreover if $\gamma : \mathbb{R} \to G$ is any other one-parameter subgroup, then $\gamma = \gamma^v$ for some $v \in \mathfrak{g}$.*

Proof. We must show that γ^v is a one-parameter subgroup, hence show

$$\gamma^v(s+t) = \gamma^v(s)\gamma^v(t)$$

for all $s, t \in \mathbb{R}$, where on the right-hand side we use the multiplication in G. Consider $\delta(t) := \gamma^v(s)^{-1}\gamma^v(s+t)$, differentiate at 0 ($\delta(0) = e$) and do not forget to use the left invariance of the object defined. Finally, use the uniqueness of integral curves to prove that $\delta(t) = \gamma^v(t)$.
Conversely, suppose γ is a one-parameter subgroup. Let $v := \gamma'(0) \in \mathfrak{g}$. Prove that $\gamma'(t) = X_v(\gamma(t))$. (Use $\gamma(t+s) = \gamma(t)\gamma(s) = l_{\gamma(t)}\gamma(s)$) □

Lemma A.3.19. *For any $s, t \in \mathbb{R}$ one has*

$$\gamma^v(st) = \gamma^{sv}(t).$$

Proof. Differentiate the left-hand side and show that it satisfies the same ODE. □

Let use denote by $\theta^v_t : G \to G$ the flow of X^v. Thus by definition $\gamma^v(t) = \theta^v_t(e)$.

Proposition A.3.20. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $\gamma : \mathbb{R} \to G$ be a smooth curve such with $\gamma(0) = e$ and $\gamma'(0) = v \in \mathfrak{g}$. The following are equivalent:*

(i.) γ *is a one-parameter subgroup,*

(ii.) $\gamma(t) = \gamma^v(t)$,

(iii.) $\theta^v_t = r_{\gamma(t)}$.

Proof. We know (i) is equiv to (ii). If (iii) holds, then

$$\gamma^v(t) = \theta^v_t(e) = r_{\gamma(t)}e = \gamma(t). \ (i.)$$

Assume (ii) holds, fix $a \in G$. Then

$$\left.\frac{d}{dt}\right|_{t=0} a\gamma^v(t) = \left.\frac{d}{dt}\right|_{t=0} l_a(\gamma^v(t)) = Dl_a(e)[v] = X_v(a).$$

Thus $t \mapsto r_{\gamma(t)}(a) = \gamma(t)$ is an integral curve of X_v with initial condition a, so by uniqueness of integral curves, one has $r_{\gamma(t)}(a) = \theta^v_t(a)$. □

Definition A.3.21. Let G be a Lie group with Lie algebra $\mathfrak{g}$. The *exponential map* is the map

$$\exp : \mathfrak{g} \to G, \ v \mapsto \gamma^v(1).$$

Proposition A.3.22. *(Properties of the exponential map). The exponential map* $\exp : \mathfrak{g} \to G$ *satisfies:*

(i) $\exp((s+t)v) = \exp(sv)\exp(tv)$ *for all* $v \in \mathfrak{g}$ *and* $s, t \in \mathbb{R}$*,*

(ii) $\exp(-tv) = (\exp(tv))^{-1}, \ \forall v \in \mathfrak{g}, \ \forall t \in \mathbb{R}$*,*

(iii) The map $t \mapsto \exp(tv)$ *is precisely the one-parameter subgroup* $\gamma^v(t)$*,*

(iv) The flow θ_t^v *of* X_v *is given by* $\theta_t^v = r_{\exp(tv)}$*.*

Theorem A.3.23. *The exponential map* $\exp : \mathfrak{g} \to G$ *is smooth. Moreover up to canonical isomorphism* $T_0\mathfrak{g} \cong \mathfrak{g}$*, the derivative of the exponential map at* $0 \in \mathfrak{g}$ *is the identity.*

Proof. The proof takes two steps.
1. Consider $\hat{X}$ on the product manifold $G \times \mathfrak{g}$ given by

$$\hat{X}(a, v) := (X_v(a), 0) \in T_aG \times T_0\mathfrak{g} \cong T_{(a,v)}G \times \mathfrak{g}.$$

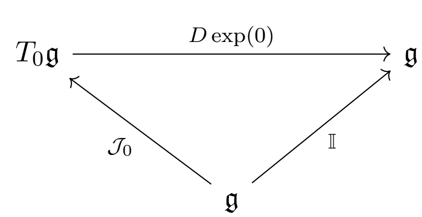
We claim that $\hat{X}$ is a vector field. It clearly satisfies the section property, thus we only need to check that it is smooth. For this, suppose $f \in C^\infty(G \times \mathfrak{g})$ is smooth. Given $v \in \mathfrak{g}$, let $f_v := f(\cdot, v) : G \to \mathbb{R}$ denote the smooth function given by regarding v as fixed. Then

$$\hat{X}(f)(a, v) = X_v(f_v)(a).$$

The vector field X_v depends linearly on v (and thus smoothly). The function f_v depends smoothly on v as f is a smooth function of a and v. Thus the expression $(a, v) \mapsto X_v(f_v)(a)$ depends smoothly on both a and v. Thus $\hat{X}$ is indeed a vector field, hence its flow $\hat{\theta}$ is also smooth and given by

$$\hat{\theta}(t, v, a) := (a\exp(tv), v), \ (t, v, a) \in \mathbb{R} \times \mathfrak{g} \times G.$$

In particular, $\hat{\theta}(1, e, \cdot) : \mathfrak{g} \to G$ is smooth. This is the map $v \mapsto (\exp(v), v)$. Thus exp is smooth.
2. We now compute the derivative of exp. Our claim is that the following diagram commutes:



So take $v \in \mathfrak{g}$. Then $\mathcal{J}_0(v) = \delta'(0)$ where $\delta(t) = tv$.

$$D\exp(0)[\mathcal{J}_0(v)] = (\exp \circ \delta)'(0) = \frac{d}{dt}\bigg|_{t=0} \exp(tv) = \frac{d}{dt}\bigg|_{t=0} \gamma^v(t) = v.$$

□

Corollary. *The exponential map is a diffeomorphismof some neighbourhood of the origin in $\mathfrak{g}$ onto its image in G.*

Proof. Since exp, by the theorem, has maximal rank at 0, this follows immediately from the inverse function theorem. □

Proposition A.3.24. *Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Let $\varphi : G \to H$ be a Lie group homomorphism. Then the following diagram commutes:*

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{D\varphi(e)} & \mathfrak{h} \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow[\varphi]{} & H \end{array}$$

Proof. If $\gamma : \mathbb{R} \to G$ is a homomorphism, then since φ is a homomorphism so is $\varphi \circ \gamma : \mathbb{R} \to H$. Applying this with $\gamma(t) = \exp(tv)$ shows that $t \mapsto \varphi(\exp(tv))$ is a one-parameter subgroup of H. Since

$$(\varphi \circ \gamma)'(0) = \frac{d}{dt}\bigg|_{t=0} \varphi(\exp(tv)) = D\varphi(e) \circ D\exp(0)[v] = D\exp(e)[v],$$

where we used the chain rule and the previous theorem. Thus by the uniqueness of integral curves, we get that $\varphi(\exp(tv)) = \exp(tD\varphi(e)[v])$. □

Corollary. *Let G be a Lie group with Lie algebra $\mathfrak{g}$ and $H \subset G$ a Lie subgroup with Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Then the exponential map* $\exp : \mathfrak{h} \to H$ *is simply given by the restriction of* $\exp : \mathfrak{g} \to G$ *to H.*

Proposition A.3.25. *Let $A \in \mathfrak{gl}(n) = Mat(n)$. Then the matrix exponential*

$$\exp(A) := \sum_{h=0}^{\infty} \frac{1}{h!} A^h$$

converges and defines an element of $GL(n)$. Moreover $A \mapsto \exp(A)$ is the exponential map of $GL(n)$.

Corollary. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then the exponential map* $\exp : \mathfrak{g} \to G$ *is given by matrix exponentiation:* $\exp(A) = e^A$.

Definition A.3.26. let G be a Lie group and let M be a manifold. A smooth map $\mu : G \times M \to M$ satisfying

$$\mu(ab, x) = \mu(a, \mu(b, x)),\ \ \mu(e, x) = x$$

for all $a, b \in G$ and $x \in M$ is called a *left action* of G *on* M. For a fixed $a \in G$, this implies that $x \mapsto \mu(a, x)$ is a diffeomorphismof M, which we denote by μ_a.

A.4 Differential forms and a bit of symplectic geometry

A.4.1 Differential forms

This section is taken over from [7], as the algebraic approach to differential forms seemed to be the fastest introducing the concepts needed further on.

Definition A.4.1. Let M be a smooth manifoldand T^*M is cotangent bundle. We define

$$\bigwedge^k (T^*M) = \bigsqcup_{p \in M} \bigwedge^k (T_p^*M),$$

to be the set of all alternating k-form at all points of M

Definition A.4.2. Let M be a smooth manifold. A *differential k-form* on M is a smooth map $\omega^k : \bigwedge^k (T^*M) \to \mathbb{R}$, such that for any $p \in M$, ω_p^k is an element of $\bigwedge^k (T_p^*M)$ and satisfies the section property:

$$\pi(\omega^k(x)) = x,$$

with $\pi : \bigwedge^k (T^*M) \to M$ the footpoint map.
We denote by $\Omega^k(M)$ the set of differential k-forms of M.

Definition A.4.3. Let $\varphi : M \to N$ be a smooth map between manifolds, $p \in M$ and $v^1, \ldots, v^k \in\in T_pM$. The *pullback* $\varphi^*\omega^k \in \Omega^k(M)$ of a k-form $\omega^k \in \Omega^k(N)$ is

$$(\varphi^*\omega^k)_p(v^1, \ldots, v^n) := \omega^k_{\varphi(p)}(\varphi_\star v^1, \ldots, \varphi_\star v^n).$$

Remark. Let use denote by $\Omega^*(M) = \bigoplus_{k \in \mathbb{N}} \Omega^k(M)$.

Definition A.4.4. An *exterior derivative* on a manifold M is a $\mathbb{R}$-linear map

$$D : \Omega^*(M) \to \Omega^*(M)$$

such that

1. D is an *antiderivation of degree 1*, i.e. $D(\omega \cdot \tau) = (D\omega) \cdot \tau + (-1)^k \omega(D\tau), \forall \omega \in \Omega^k(M),\ \tau \in \Omega^l(M);\ D(\omega \cdot \tau) \in \Omega^{k+l+1}(M)$.
2. $D \circ D = 0$.
3. if $f \in C^\infty(M)$ and $X \in \mathfrak{X}(M)$, then $(Df)X = Xf$.

This last property implies that on 0-forms an exterior derivative agrees with the differential *df* of a function $f \in C^\infty(M)$. Hence, in a coordinate chart (x^i) about a point $p \in M$,

$$Df = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i.$$

If $\omega^k \in \Omega^k(M)$ and $\sigma : U \subset M \to O$ is a coordinate chart about $p \in M$ with coordinates (x^i), then

$$\omega^k = \sum a_I dx^I,\ a_I \in C^\infty(O)$$

Then if d is a exterior derivation on O, then

$$\begin{aligned} d\omega^k &= \sum da_I \wedge dx^I + \sum a_I ddx^I \\ &= \sum (da_I) \wedge dx^I = \sum_I \sum_j \frac{\partial a_I}{\partial x^j} dx^j \wedge dx^I. \end{aligned}$$

Definition A.4.5. A k-form $\omega^k \in \Omega^k(M)$ is called *closed* if $d\omega^k = 0$.

Definition A.4.6. Let M be a smooth manifold, $\omega \in \Omega^{k+1}M)$ and $X \in \mathfrak{X}(M)$. Then define $i_X\omega \in \Omega^k(M)$ by:

$$i_X\omega(X_1, \ldots, X_k) = \omega(X, X_1, \ldots, X_k)$$

If $\omega \in \Omega^0(M)$, we define $i_X\omega = 0$. We call $i_X\omega$ the *inner product* of X and ω.

Proposition A.4.7. *For $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, we have:*

1. *i_X is an antiderivation.*
2. *$i_X \circ i_X = 0$.*
3. *$i_X df = L_X f$ (the Lie derivative of f).*

Proof. Use the definitions. □

Theorem A.4.8. *(Exterior differentiation under a Pullback) Let $\varphi : M \to N$ be a smooth map between smooth manifold. If $\omega \in \Omega^k(M)$, then $d\varphi^*\omega = \varphi^* d\omega$.*

Proof. We need to check the case $k = 0$: For $p \in N$ and $X_p \in T_pN$,

$$\begin{aligned}(d\varphi^*\omega)_p(X_p) &= X_p(\varphi^*\omega)\\ &= X_p(\omega \circ \varphi)\end{aligned}$$

and

$$\begin{aligned}(\varphi^* d\omega)_p(X_p) &= (d\omega)_{\varphi(p)}(\varphi_\star X_p)\\ &= (\varphi_\star X_p)\omega\\ &= X_p(\omega \circ \varphi).\end{aligned}$$

Now to check the general case one has to check it at a given point $p \in N$. Therefore, by going to local coordinates about p,

$$\begin{aligned}\varphi^*\omega &= \sum (\varphi^* a_I)\varphi^* dy^{i_1} \wedge \cdots \wedge \varphi^* dy^{i_k}\\ &= \sum (a_I \circ \varphi) d\varphi^{i_1} \wedge \cdots \wedge d\varphi^{i_k}.\end{aligned}$$

Hence

$$d\varphi^*\omega = \sum d(a_I \circ \varphi) \wedge d\varphi^{i_1} \wedge \cdots \wedge d\varphi^{i_k}.$$

On the other hand,

$$\begin{aligned}\varphi^* d\omega &= \sum \varphi^*(da_I \wedge dy^{i_1} \wedge \cdots \wedge dy^{i_k})\\ &= \sum \varphi^* da_I \wedge d\varphi^{i_1} \wedge \cdots \wedge d\varphi^{i_k}\\ &= \sum d(\varphi^* a_I) \wedge d\varphi^{i_1} \wedge \cdots \wedge d\varphi^{i_k}\\ &= \sum d(a_I \circ \varphi) \wedge d\varphi^{i_1} \wedge \cdots \wedge d\varphi^{i_k}\end{aligned}$$

Therefore,

$$d\varphi^*\omega = \varphi^* d\omega.$$

□

A.4.2 Symplectic geometry

Here we will orient ourselves on [2] to introduce some symplectic geometry.

Let V be an n-dimensional vector space and $\Omega : V \times V \to \mathbb{R}$ be a bilinear skew-symmetric maps. Let $\tilde{\Omega} : V \to V^*$ denote the linear map $\tilde{\Omega}(v)(u) = \Omega(v, u)$.

Definition A.4.9. The map Ω is called *symplectic*, if $\tilde{\omega}$ is bijective. Then (V, ω) is called a *symplectiv vector space.*

Definition A.4.10. A *symplectomorphism* $\varphi : (V, \Omega) \to (V', \Omega')$ between two symplectic vector spaces is a linear isomorphism $\varphi : V \tilde{\to} V'$ usch that $\varphi^*\Omega' = \Omega$. (with $(\varphi^*\Omega')(u, v) = \Omega'(\varphi(u), \varphi(v))$ for all $u, v \in V$). If a symplectomorphism, (V, ω) and (V', Ω') are called *symplectomorphic.*

Definition A.4.11. Let ω be a 2-form on a manifold M, it is called a *symplectic form* if it is closed and if for all $q \in M$, ω_q is a symplectic bilinear maps on $T_qM \times T_qM \to \mathbb{R}$.

Definition A.4.12. A *symplectic manifold* is a pair (M, ω) where M is a manifold and ω a symplectic form.

Example A.4.13. Let $M = \mathbb{R}^{2n}$ with linear coordinates $x^1, \dots, x^n, y^1, \dots, y^n$. The form

$$\omega_0 = \sum_{i=1}^{n} dx^i \wedge dy^i$$

is symplectic and the $\{(\frac{\partial}{\partial x^1}\big|_x)_p, \dots, (\frac{\partial}{\partial x^n}\big|_x)_p, (\frac{\partial}{\partial y^1}\Big|_y), \dots, (\frac{\partial}{\partial y^n}\Big|_y)\}$ is a symplectic basis of T_pM.

Definition A.4.14. Let M, ω and M', ω' be 2n-dimensional symplectic manifolds, and let $\varphi : M \to M'$ be a diffeomorphism. Then φ is a *symplectomorphism* if $\varphi^*\omega_2 = \omega_1$.

Definition A.4.15. Let M be an n-dimensional smooth manifold, then T^*M is a 2n-dimensional smooth manifold. Let $(q^1, \dots, q^n, p_1, \dots, p_n)$ be a set of local coordinate on T^*M. The *Liouville 1-form* $\lambda \in \Omega(T^*M)$ (also called *tautological 1-form*) is given in local coordinates by

$$\lambda = \sum_{i=1}^{n} p_i dq^i.$$

Alternatively, one can define λ without referring to coordinates. Denote by $\pi : T^*M \to M$ the footpoint map. Then

$$\lambda_{(q,p)}(\xi) = p(D\pi(q,p)(\xi)), \ \forall \xi \in T_{(q,p)}T^*M.$$

Theorem A.4.16. *Let M be an n-dimensional smooth manifold. Then its cotangent bundle T^*M can naturally be made into a symplectic manifold. Its canonical symplectic form ω on T^*M is given by*

$$\omega = d\lambda,$$

with λ being the Liouville 1-form defined above. In local coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ on T^*M, ω takes the form

$$\omega := \sum_{i=1}^{n} dp_i \wedge dq^i.$$

References

[1] V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Graduate Texts in Mathematics. Springer Verlag, second edition, 1989.

[2] Ana Cannas da Silva. *Lectures on Symplectic Geometry*. Lecture Notes in Mathematics. Springer Verlag, corrected 2nd printing edition, 2008.

[3] Victor Guillemin. *Moment Maps and Combinatorial Invariants of Hamiltonian T^n-spaces*. Progress in Mahtematics 122. Springer Verlag, 1994.

[4] Charles Hudgins. Understanding noether's theorem with symplectic geometry. math.uchicago.edu/~may/REU2017/REUPapers/Hudgins.pdf, 2017. Accessed: 06.11.2018.

[5] Marco Mazzucchelli. *Critical Point Theory for Lagrangian Systems*. Progress in Mahtematics 293. Springer Verlag, 2012.

[6] Will J. Merry. Differential geometry i (autumn 2018). https://www.merry.io/differential-geometry-i-lecture-notes, 2018. Accessed: 09.11.2018.

[7] Loring W.Tu. *An Introduction to Manifolds*. Universitext. Springer Verlag, 2008.