SYMPLECTIC GEOMETRY AND HAMILTONIAN SYSTEMS

E. LERMAN

Contents

1. Lecture 1. Introduction and basic definitions

A **symplectic manifold** is a pair (M, ω) where M is a manifold and ω is a nondegenerate closed 2-form on M (nondegeneracy is defined below).

What are the reasons for studying symplectic geometry? It is an outgrowth of classical mechanics. It has connections with

- classical mechanics,
- quantum mechanics (via geometric quantizations, deformation quantization),
- representations of Lie groups – "the orbit method",
- PDEs (via microlocal analysis),
- gauge theory,
- geometric invariant theory in algebraic geometry...

We first define nondegeneracy of a skew-symmetric bilinear form on a vector space.

Definition 1. Let V be a vector space. Let $\omega : V \times V \longrightarrow \mathbb{R}$ be a skew-symmetric, bilinear 2-form, $\omega \in \bigwedge^2 V^*$. The form ω is **nondegenerate** if for every $v \in V$,

$$\omega(v, u) = 0 \quad \forall u \in V \qquad \Longrightarrow \qquad v = 0$$

Note that since ω is skew-symmetric $\omega(v, v) = -\omega(v, v)$, hence $\omega(v, v) = 0$.

Example 1. Let $V = \mathbb{R}^2$ with coordinates x and y. It is not hard to check that the bilinear form $\omega = dx \wedge dy$ is nondegenerate (also see below).

Here is a criterion for nondegeneracy. Given as skew-symmetric bilinear form ω on a (finite dimensional) vector space V, define the linear map $\omega^\sharp : V \longrightarrow V^*$ by $v \mapsto \omega(v, \cdot)$. Because ω is bilinear, $\omega^\sharp$ is linear, so $\omega^\sharp$ is well-defined. We claim that ω is non-degenerate if and only if $w^\sharp$ is 1-1. Indeed, suppose $\omega^\sharp(v) = 0$. Then $\omega^\sharp(v)(u) = 0 = \omega(v, u)$ for any $u \in V$.

Note that $\omega^\sharp$ is 1-1 if and only if $\omega^\sharp$ is an isomorphism, since for a finite dimensional vector space $\dim V$ we have $\dim V = \dim V^*$.

Definition 2. A **symplectic vector space** is a pair (V, ω) where V is a vector space and $\omega \in \bigwedge^2(V^*)$ is a nondegenerate bilinear skew-symmetric form.

We can now define a nondegenerate differential two-form. The condition is pointwise.

Definition 3. A two-form $\omega \in \Omega^2(M)$ is **nondegenerate** if and only if for any point $m \in M$, the bilinear form ω_m on the tangent space T_mM is nondegenerate.

Example 2. Again consider $V = \mathbb{R}^2$ with coordinates x and y. Consider $\omega = dx \wedge dy$ as a (constant coefficient) differential form. It is easy to see that $\omega^\sharp(\frac{\partial}{\partial x}) = \iota(\frac{\partial}{\partial x})(dx \wedge dy) = dy$ and that similarly $\omega^\sharp(\frac{\partial}{\partial y}) = -dx$. So $\omega^\sharp$ is bijective. Therefore $(\mathbb{R}^2, dx \wedge dy)$ is a symplectic manifold.

Let (M, ω) be a symplectic manifold. Then for every point $m \in M$, $(\omega_m)^\sharp : T_mM \longrightarrow T_m^*M$ is an isomorphism so there is a correspondence between 1-forms and vector fields. In particular, given a function $f \in C^\infty(M)$, the differential df of f is a one-form $m \mapsto df_m = \sum \frac{\partial f}{\partial x_i}(dx_i)_m$. This gives us a vector field $X_f(m) := ((\omega_m)^\sharp)^{-1}(df_m)$. Equivalently X_f is defined by $\iota(X_f)\omega = df$.

Definition 4. The vector field X_f defined above is called the **Hamiltonian vector field** of the function f on a symplectic manifold (M, ω).

Example 3. Consider the symplectic manifold $(M = \mathbb{R}^2, \omega = dx \wedge dy)$. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a smooth function. Let's compute its Hamiltonian vector field X_f:
$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$.
$(\omega^\sharp)(\frac{\partial}{\partial x}) = dy$, $(\omega^\sharp)(\frac{\partial}{\partial y}) = -dx$. Hence

$$X_f := (\omega^\sharp)^{-1}(df) = \frac{\partial f}{\partial x}(\omega^\sharp)^{-1}(dx) + \frac{\partial f}{\partial y}(\omega^\sharp)^{-1}(dy) = -\frac{\partial f}{\partial x}\frac{\partial}{\partial y} + \frac{\partial f}{\partial y}\frac{\partial}{\partial x}.$$

Exercise 1. Check that $(M = \mathbb{R}^2, \omega = (x^2+y^2+1)\, dx \wedge dy)$ is a symplectic manifold. Compute the Hamiltonian vector field of a function of f on this symplectic manifold.

Example 4. The manifold $M = \mathbb{R}^{2n}$ with coordinates $q_1, \dots, q_n, p_1, \dots, p_n$ and the form . $\omega = \sum dq_i \wedge dp_i$ is a symplectic manifold.
The Hamiltonian vector field of a function f is $X_f = \sum_{i=1}^n \left(-\frac{\partial f}{\partial q_i}\frac{\partial}{\partial p_i} + \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q_i}\right)$.

So far we haven't talked about the condition $d\omega = 0$. One consequence of the condition is:

Theorem 5 (Darboux theorem)**.** *Let M be a manifold, ω a symplectic form on M. Then* $\dim M$ *is even, say* $2n$. *Moreover, for every point* $m \in M$, *there are coordinates* $q_1, \dots, q_n, p_1, \dots, p_n$ *defined near* m *such that in these coordinates* $\omega = \sum dq_i \wedge dp_i$.

The proof will take the next few lectures. The first step of the proof is to consider the linear case.

Theorem 6. *Let (V, ω) be a finite dimensional symplectic vector space. Then*

1. *V is even dimensional, and*
2. *there is a basis $e_1, \dots, e_n, f_1, \dots, f_n$, where $n = \frac{1}{2}\dim V$ such that $\omega(e_i, f_j) = \delta_{ij}$ and $\omega(e_i, e_j) = 0 = \omega(f_i, f_j)$ where δ_{ij} is the Croniker delta function.*

Consequently, $\omega = \sum e_i^* \wedge f_i^*$ *where* $e_1^*, \dots, e_n^*, f_1^*, \dots, f_n^* \in V^*$ *is the dual basis.*

Remark 7. The basis $e_1, \dots, e_n, f_1, \dots, f_n$ in the above theorem is called a **symplectic** basis.

Proof. (induction on the dimension of V). Let e_1 be any nonzero vector in V. Since the form ω is nondegenerate there exists a vector $f \in V$ such that $\omega(e_1, f) \neq 0$. Let $f_1 = \frac{f}{\omega(e_1,f)}$. Then $\omega(e_1, f_1) = 1$. Let $W = Span\{e_1, f_1\} = \{ae_1 + bf_1 \,:\, a, b \in \mathbb{R}\}$.

We now briefly digress.

Definition 8. Let (V, ω) be a symplectic vector space and let $U \subset V$ a linear subspace. The **symplectic perpendicular** to U in V is the space $U^\omega := \{\, v \in V \,:\, \omega(v, u) = 0, \forall u \in U \,\}$.

Proposition 9. *Let* (V, ω) *be a symplectic vector space,* $U \subset V$ *a linear subspace and* U^ω *be the symplectic perpendicular. Then*

$$\dim U + \dim U^\omega = \dim V.$$

In fact the isomorphism $\omega^\sharp : V \to V^*$ *maps* U^ω *onto the annihilator* U° *of* U *in* V^*.

Recall that $U^\circ := \{\ell \in V^* \mid \ell|_U = 0\}$.

Proof. Since $\dim U^\circ = \dim V - \dim U$, it is enought to prove the last claim. If $v \in U^\omega$ than for any $u \in U$,

$$0 = \omega(v, u) = \omega^\sharp(v)(u).$$

Hence $\omega^\sharp(v) \in U^\circ$.

Conversely, since $\omega^\sharp$ is onto, any $\ell \in V^*$ is of the form $\ell = \omega^\sharp(v)$ for some v. If ℓ is in U°, then for any $u \in U$

$$0 = \ell(u) = \omega^\sharp(v)(u) = \omega(v, u).$$

Hence $v \in U^\circ$. □

Note $U^\omega \cap U$ need not be zero. For exaple if U is a line, i.e., $U = \mathbb{R}u$ for some $0 \neq u$ then since $\omega(u, u) = 0$, we have $\mathbb{R}u \subset (\mathbb{R}u)^\omega$. On the other hand we need not have $U \subset U^\omega$ either.

Lemma 10. *Let* (V, ω) *be a symplectic vector space,* $U \subset V$ *a linear subspace and* U^ω *be the symplectic perpendicular. Then*

$$(U^\omega)^\omega = U.$$

Proof. By Proposition 9, $\dim(U^\omega)^\omega = \dim V - \dim U^\omega = \dim V - (\dim V - \dim U) = \dim U$. Also $U \subset (U^\omega)^\omega$. Therefore $U = (U^\omega)^\omega$. □

Proof of Theorem 6 continued. Consider W^ω. We know $\dim W^\omega = \dim V - \dim W = \dim V - 2$. We claim that $W^\omega \cap W = \{0\}$.

Indeed, suppose $w \in W \cap W^\omega$. Then $w = ae_1 + bf_1$ for some $a, b \in \mathbb{R}$. $0 = \omega(w, e_1) = \omega(ae_1 + bf_1, e_1) = a\omega(e_1, e_1) + b\omega(f_1, e_1) = a \cdot 0 + b \cdot (-1) = -b$. So $b = 0$. Similarly, $\omega(w, f_1) = 0 = a$. This proves the claim and implies, since $\dim W + \dim W^\omega = \dim V$ that $V = W \oplus W^\omega$.

To get to the inductive step, it is enough to show the restriction of ω to W^ω is nondegenerate. But we have already proved this since if there is $v \in W^\omega$ such that $\omega(v, w) = 0$ for all $w \in W^\omega$ then $v \in (W^\omega)^\omega = W$ and $W \cap W^\omega = \{0\}$, so $v = 0$.

By induction, $\dim W^\omega$ is even, say $2n-2$, and there is a basis of W^ω $e_2, \ldots, e_n,\ f_2, \ldots, f_n$ such that $\omega(e_i, f_j) = \delta_{ij}$, $\omega(e_i, e_j) = 0 = \omega(f_i, f_j)$ for all $i, j \geq 2$. Therefore, $\dim V$ is even, and $e_1, \ldots, e_n, f_1, \ldots, f_n$ is the desired basis. □

Homework Problem 1. Prove that $\omega \in \bigwedge^2 V^*$ is non-degenerate if and only if $\omega^n = \omega \wedge \cdots \wedge \omega \neq 0 \quad n = \frac{1}{2}\dim V$.

Homework Problem 2. Suppose V is a finite dimensional vector space (as usual over the reals). Let $\omega : V \times V \longrightarrow \mathbb{R}$ be any skew-symmetric, bilinear 2-form. Show that there is a basis $e_1, \ldots, e_k, f_1, \ldots, f_k, z_1, \ldots, z_\ell$ such that $\omega(z_i, v) = 0$ for all $v \in V$, $\omega(e_i, f_j) = \delta_{ij}$, $\omega(e_i, e_j) = 0 = \omega(f_i, f_j)$. Hence $\omega = \sum e_i^* \wedge f_i^*$ where $e_i^*, \ldots, e_k^*, f_1^*, \ldots, f_k^*, z_1^*, \ldots, z_\ell^*$ is the dual basis.

This provides another proof for the fact that $\omega \in \bigwedge^2 V^*$ is non-degenerate if and only if $\omega^n \neq 0$ where $n = \frac{1}{2}\dim V$.

2. Lecture 2. Symplectic linear algebra

There are several consequence of Theorem 6. The first one demonstrates the difference between skew-symmetric and symmetric nondegenerate bilinear forms. The second one is a through-away remark, though it is amusing.

1. **Corollary 11.** *Any two symplectic forms on a given vector space are the same, i.e. if V is a vector space and ω_1, ω_2 are two symplectic forms on V, then there is an invertible linear map $A : V \longrightarrow V$ such that $\omega_1(Av, Aw) = \omega_2(v, w)$ for all $v, w \in V$. That is, $A^*\omega_1 = \omega_2$ where $(A^*\omega_1)(v, w) = \omega_1(Av, Aw)$*

   Proof. There exists a basis $e_1, \ldots, e_n, f_1, \ldots, f_n$ such that $\omega_1 = \sum e_i^* \wedge f_i^*$, and a basis $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ such that $\omega_2 = \sum \alpha_i^* \wedge \beta_i^*$. Define $A : V \longrightarrow V$ by $A(\alpha_i) = e_i$, $A(\beta_j) = f_j$. Then $A^*(e_i^*) = \alpha_i^*$, $A^*(f_j^*) = \beta_j^*$. Consequently, $A^*\omega_1 = A^*(\sum e_i^* \wedge f_i^*) = \sum (A^* e_i^*) \wedge (A^* f_i^*) = \sum \alpha_i^* \wedge \beta_i^* = \omega_2$. □

2. **Corollary 12.** *Any even dimensional vector space posesses a symplectic form.*

   Proof. Choose a basis $e_1, \ldots, e_n, f_1, \ldots, f_n$, $n = \frac{1}{2}\dim V$. Let $e_1^*, \ldots, e_n^*, f_1^*, \ldots, f_n^*$ denote the dual basis. Let $\omega = \sum e_i^* \wedge f_i^*$, $e_1^*, \ldots, e_n^*, f_1^*, \ldots, f_n^*$ dual basis. The form ω is nondegenerate. □

Definition 13. Let (V, ω) be a symplectic vector space. A subspace $U \subset V$ is **symplectic** if the restriction of the symplectic form ω to U is non-degenerate. Equivalently U is symplectic if and only if $U \cap U^\omega = \{0\}$.

Exercise 2. Check that the two definitions of a symplectic subspace are are indeed equivalent.

Example 5. Let (V, ω) be a symplectic vector space of dimension $2n$ and let $e_1, \dots, e_n$, $f_1, \dots, f_n$ be a symplectic basis so that $\omega = \sum e_i^* \wedge f_i^*$. Let $U = \text{Span}\{e_1, f_1, e_2, f_2, \dots, e_k, f_k\}$ for some $k \leq n$. Then U is a symplectic subspace.

Exercise 3. Show that Example 5 is, in some sense, the only example. That is, given a symplectic vector space (V, ω) of dimension $2n$ and a symplectic subspace U show that there exists a symplectic basis $e_1, \dots, e_n$, $f_1, \dots, f_n$ of $((V, \omega)$ such that $U = \text{Span}\{e_1, f_1, e_2, f_2, \dots, e_k, f_k\}$ for some $k \leq n$

Note that since $(U^\omega)^\omega = U$, we have $U \cap U^\omega = \{0\}$ if and only if $U^\omega \cap (U^\omega)^\omega = \{0\}$. Hence a subspace U is symplectic if and only if its symplectic perpendicular U^ω is symplectic. Also since $\dim U^\omega = \dim V - \dim U$, if U is a symplectic subspace of (V, ω) then $V = U \oplus U^\omega$.

Definition 14. A subspace E of (V, ω) is **isotropic** if $\omega(e, e') = 0$ for all $e, e' \in E$. Equivalently, E is an isotropic subspace if and only if $E \subset E^\omega$.

Example 6. Again let (V, ω) be a symplectic vector space of dimension $2n$ with a symplectic basis $e_1, \dots, e_n$, $f_1, \dots, f_n$. Consider the subspaces $E := \mathit{span}\{e_1, \dots, e_k\}$, for some $k \leq n$, and $F = \{f_1, \dots, f_\ell\}$, for some $\ell \leq n$. Both E and F are isotropic. (Check this!)

Example 7. A line in a symplectic vector space is always isotropic.

Note that since $E \subset V$ is isotropic if and only if $E \subset E^\omega$, we have $\dim E \leq \dim E^\omega = \dim V - \dim E \Longrightarrow \dim E \leq \frac{1}{2} \dim V$.

Definition 15. A subspace L of a symplectic vector space (V, ω) is **Lagrangian** if and only if L is maximally isotropic, i.e. if E is any isotropic subspace of V with $E \supset L$, then $E = L$.

If $\dim L = \frac{1}{2} \dim V$, and L is isotropic, then it is Lagrangian. The converse is also true.

Lemma 16. *If $L \subseteq V$ is a Lagrangian subspace of a symplectic vector space (V, ω), the $\dim L = \frac{1}{2} \dim V$.*

Proof (by contradiction). Suppose $L \subset V$ is isotropic and $\dim L < \frac{1}{2} \dim V$. Then $\dim L^\omega = \dim V - \dim L > \frac{1}{2} \dim V > \dim L$. Therefore there is a vector $v \in L^\omega, v \neq 0, v \notin L$. Consider $E = L \oplus \mathbb{R}v$. We claim that E is again isotropic.

Indeed if $e, e' \in E$, there are $\ell, \ell' \in L$ and $a, a' \in \mathbb{R}$ such that $e = \ell + av$ and $e' = \ell' + a'v$. Now $\omega(e, e') = \omega(\ell + av, \ell' + a'v) = \omega(\ell, \ell') + aa'\omega(v, v) + a\omega(v, \ell') + a'\omega(\ell, v) = 0$. Therefore E is isotropic. Hence if $L \subset V$ is isotropic and $\dim L < \frac{1}{2} \dim V$, L is not maximal. □

Note that the above argument gives another proof that a symplectic vector space has to be even dimensional.

Note also that the existence of a symplectic basis $e_1, \dots, e_n$, $f_1, \dots, f_n$ of (V, ω) gives us a splitting of V as a direct sum of two Lagrangian subspaces: $V = E \oplus F$ with $E := \text{Span}\{e_1, \dots e_n\}$ and $F := \text{Span}\{f_1, \dots f_n\}$.

Example 8 (The one and only example of a symplectic vector space)**.** Let W be any vector space and let $W^* = \{\ell : W \longrightarrow \mathbb{R}\}$ be its dual. We have a natural bilinear pairing $W^* \times W \longrightarrow \mathbb{R}$ given by $(\ell, w) \longmapsto \ell(w) =: \langle \ell, w\rangle$. Consider $V = W^* \oplus$ and define a skew-symmetric bilinear form $\omega_0 : (W \oplus W^*) \times (W \oplus W^*) \longrightarrow \mathbb{R}$ by $\omega_0((\ell, w), (\ell', w')) = \langle \ell', w\rangle - \langle \ell, w'\rangle$. The form ω_0 is non-degenerate. Indeed, suppose $(\ell, w) \neq (0, 0)$. Then either $w \neq 0$ or $\ell \neq 0$. Say $w \neq 0$. Then there is ℓ' such that $\ell'(w) \neq 0$. Therefore $\omega_0((\ell, w), (0, \ell')) = \langle \ell', w\rangle - 0 \neq 0$. Similar argument for $\ell \neq 0$.

We will refere to the form ω_0 as the canonical symplectic form on $W \times W^*$.

Definition 17. An **isomorphism** of two **symplectic vector** spaces (V, ω) and (V', ω') is an invertible linear map $A : V \longrightarrow V'$ such that $\omega'(Av, Aw) = \omega(v, w)$. i.e. $A^*\omega' = \omega$.

Exercise 4. Suppose $A : V \longrightarrow V'$ is a linear map between two vector spaces . Let ω, ω' be symplectic forms on V, V' respectively. Show that if $A^*\omega' = \omega$, then A is injective.

Proposition 18. *Given a symplectic vector space (V, ω) there exists a Lagrangian subspace E of V and an isomorphism $A : (V, \omega) \longrightarrow (E^* \times E, \omega_0)$ of symplectic vector spaces, where ω_0 is the canonical symplectic form on $E^* \times E$ (see Example 8 above).*

Note that this gives another proof that symplectic vector spaces are classified by their dimension.

Proof. Let $e_1, \ldots, e_n, f_1, \ldots, f_n$ be a sympletic basis of (V, ω). Let $E = \mathit{span}\{e_1, \ldots, e_n\}$ and let $F = \{f_1, \ldots, f_n\}$. Then $V = E \oplus F$. Define $A : E \oplus F \longrightarrow E^* \times E$ by $A(e, f) = (\omega(\cdot, f), e)$.

Then

$$\omega_0(A(e, f), A(e', f')) = \omega_0((\omega(\cdot, f), e), (\omega(\cdot, f'), e')) =$$
$$\langle \omega(\cdot, f'), e\rangle - \langle \omega(\cdot, f), e'\rangle = \omega(e, f') + \omega(e, e') + \omega(f, f') + \omega(f, e') =$$
$$\omega(e + f, e' + f').$$

□

We finish with an example of a symplectic basis of $(W^* \oplus W, \omega_0)$: Let $e_1, \ldots, e_n$ be any basis of W. Let $e_1^*, \ldots, e_n^*$ be the dual basis. The collection $\{e_1, \ldots, e_n, e_1^*, \ldots, e_n^*\}$ is a symplectic basis.

3. Lecture 3. Cotangent bundles and the Liouville form

Recall that a symplectic manifold is a pair (M, ω), where M is a manifold and ω is a closed non-degenerate 2-form.

Example 9. Consider $M = S^2$. Let ω be any 2-form on S^2 such that $\omega_x \neq 0, \forall x \in S^2$. Then (S^2, ω) is a symplectic manifold. For example if we identify the two sphere with the set $\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ we may take $\omega = (x_1 dx_2 \wedge dx_3 + x_2 dx_3 \wedge dx_1 + x_3 dx_1 \wedge dx_2)|_{S^2}$.

More generally, any surface Σ with a nowhere vanishing two-form ω is symplectic.

Remark 19. If σ is a form, we will write interchangeably σ_x and $\sigma(x)$ for its value at a point x.

Theorem 20. *The cotangent bundle $M = T^*X$ of a manifold X is naturally a symplectic manifold.*

Proof. We start by constructing the so called tautological one-form α (also known as the Liouville form).

Let $\pi : T^*X \longrightarrow X$ denote the projection.

Consider a covector $\eta \in T^*X$. Let $x = \pi(\eta)$, so that $\eta \in T^*_x X$.[1] Let $v \in T_\eta(T^*X)$ so that $d\pi_\eta(v) \in T_x X$, where $d\pi_\eta : T_\eta(T^*X) \longrightarrow T_{\pi(\eta)}X = T_x X$. We define

$$\alpha_\eta(v) = \eta(d\pi_\eta(v)).$$

We need to show that α is smooth. Fix a covector η_0 and let $x_0 = \pi(\eta_0)$. Choose coordinates $x_1, \dots, x_n$ on X near x_0. Denote the corresponding coordinates on T^*X by $x_1 \circ \pi, \dots, x_n \circ \pi$, $\xi_1, \dots, \xi_n$. Recall that ξ_i's are defined by $\xi_i(\eta) = \eta(\frac{\partial}{\partial x_i}|_x)$ for any covector η in the coordinate patch on T^*X. Then $\alpha(x_1 \circ \pi, \dots, x_n \circ \pi, \xi_1, \dots, \xi_n) = \alpha_{(\sum \xi_i dx_i)} = (\sum \xi_i dx_i) \circ d\pi = \sum \xi_i d(x_i \circ \pi)$. If we abuse notation by writing x_i for $x_i \circ \pi$ then $\alpha = \sum \xi_i dx_i$. So $\alpha_{\sum \xi_i dx_i} = \sum \xi_i dx_i$. Therefore the tautological one-form α is smooth.

Next observe that $d\alpha$ is a closed two-form and that in coordinates $d\alpha = \sum d\xi_i \wedge dx_i$. Hence $d\alpha$ is non-degenerate. So $\omega := \pm d\alpha$ is a symplectic form on the cotangent bundle T^*X. $\square$

We next address the naturality of the symplectic form $d\alpha$ on T^*X.

Homework Problem 3. Suppose X and Y are two manifolds. If $f : X \longrightarrow Y$ a diffeomorphism, then it naturally lifts to a map $\tilde{f} : T^*X \longrightarrow T^*Y$ of cotangent bundles:

For every point $x \in X$ we have $df_x : T_x X \longrightarrow T_{f(x)}Y$. Since f is invertible we get $(df_x)^{-1} : T_{f(x)}Y \longrightarrow T_x X$. By taking the transpose we get $(df_x^{-1})^T : T^*_x X \longrightarrow T^*_{f(x)}Y$. We now define $\tilde{f}$ by $\tilde{f}(x, \eta) = (f(x), (df_x^{-1})^T \eta)$, where, of course, $x = \pi(\eta)$.

Let α_Y be the tautologiacal 1-form on T^*Y, and α_X be the tautologiacal 1-form on T^*X. Show that $(\tilde{f})^*\alpha_Y = \alpha_X$. Conclude that $(\tilde{f})^* d\alpha_Y = d\alpha_X$.

Remark 21. In literature both $d\alpha$ and $-d\alpha$ are used as canonical/natural symplectic forms on T^*X. Recall that if f is a function on a symplectic manifold (M, ω), it defines a vector field X_f by $\iota(X_f)\omega = \pm df$. We always want in natural coordinates $x_1, \dots, x_n, \xi_1, \dots, \xi_n$ on T^*X, $X_f = \sum_{i=1}^n \frac{\partial f}{\partial \xi_i}\frac{\partial}{\partial x_i} - \frac{\partial f}{\partial x_i}\frac{\partial}{\partial x i_i}$ (this is traditional). So we can either choose $\omega = -d\alpha$ and $\iota(X_f)\omega = df$ or $\omega = d\alpha$ and $\iota(X_f)\omega = -df$.

Homework Problem 4. Let X be a vector space. Then $T^*X \cong X \oplus X^*$. Show that $d\alpha_X = \pm\omega_0$ where ω_0 is the natural symplectic form on $X \oplus X^*$ defined earlier (in Example 8).

Exercise 5. If μ is any closed 2-form on a manifold X, then $d\alpha_X + \pi^*\mu$ is also a symplectic form on T^*X.

[1]We will oftent abuse notation and write redundantly a pair (x, η) for a covector $\eta \in T^*_x X$

We now come to the first serious theorem.

Theorem 22 (Darboux Theorem)**.** *Let (M, ω) be a symplectic manifold and let $m \in M$ be a point. Then there exist coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ defined on a neighborhood $U \subseteq M$ of m such that $\omega = \sum dp_i \wedge dq_i$ on U.*

In other words all symplectic manifolds of a given dimension are locally isomorphic.

Let us now schetch the main ideas of the proof before doing the details.

1. Since the claim is local, we may assume M is a disk in $\mathbb{R}^{2n}$, i.e., $M = \{(x_1^2, \ldots, x_{2n}^2) \mid x_1^2 + \cdots + x_{2n}^2 < R^2\}, \omega = \sum \omega_{ij}(x) dx_i \wedge dx_j$. Let $\omega_0 = \sum \omega_{ij}(0) dx_i \wedge dx_j$.

2. Next we observe that it is enough to find $\varphi : M \longrightarrow M$ (or on a smaller disk) such that $\varphi^* \omega = \omega_0$. This is because, as we saw in all constant coefficient symplectic forms are the same. More precisely there is an $A : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ such that

$$A^* \omega_0 = \sum_{i=1}^{n} dx_i \wedge dx_{i+n}$$

(see Theorem 6 and its corollaries).

3. (Moser's deformation argument) First make problem harder. Let $\omega_t = t\omega + (1-t)\omega_0$. Note that $\omega_t|_{t=0} = \omega_0$, $\omega_t|_{t=1} = \omega$, and that $\omega_t(0) = t\omega(0) + (1-t)\omega_0 = \omega_0$. So $\omega_t(0)$ is nondegenerate for all $0 \leq t \leq 1$. Hence ω_t is symplectic for all t in some neighborhood of 0.

Now, look for a family of diffeomorphisms $\{\varphi_t : M \longrightarrow M\}_{0 \leq t \leq 1}$ such that $\varphi_t^* \omega_t = \omega_0$ and $\varphi_0 = id$. The vector field $X_t(x) := \frac{d}{ds}\big|_{s=1} \varphi_s(x)$ is a time-dependent vector field. We will see that knowing X_t is equivalent to knowing the isotopy $\{\varphi_t\}_{0 \leq t \leq 1}$.

We next observe that if such an isotopy exists then $\frac{d}{dt}(\varphi_t^* \omega_t) = 0$. On the other hand, we'll see that $\frac{d}{dt}(\varphi_t^* \omega_t) = \varphi_t^*(L_{X_t} \omega_t) + \varphi_t^* \dot{\omega}_t = 0$ where L denotes the Lie derivative.

Recall that for any form σ and any vector field Y we have

$$L_Y(\sigma) = d\iota(Y)(\sigma) + \iota(Y) d(\sigma).$$

Consequently

$$\varphi_t^*(d\iota(X_t)\omega_t + \iota(X_t) d\omega_t + \dot{\omega}_t) = 0$$

$$d\iota(X_t)\omega_t + \iota(X_t) d\omega_t + \dot{\omega}_t) = 0 \quad \text{(since pullback by a diffeomorphism is injective)}$$

$$d\iota(X_t)\omega_t = -\dot{\omega}_t \quad \text{(since } \omega_t \text{ is closed)}$$

Since $\dot{\omega}_t = \frac{d}{dt}(t\omega + (1-t)\omega_0) = \omega - \omega_0$, we have

$$d\iota(X_t)\omega_t = \omega_0 - \omega.$$

The form ω is closed, so $\omega_0 - \omega$ is closed as well. Hence, by Poincaré Lemma, $\omega_0 - \omega = d\theta$ for some one-form θ. We may assume $\theta(0) = 0$.

We therefore are looking for a time-dependent vector field X_t such that

$$d(\iota(X_t)\omega_t) = d\theta.$$

Enough to find X_t such that

$$\iota(X_t)\omega_t = \theta. \tag{1}$$

The form ω_t is non-degenerate near 0 in $\mathbb{R}^{2n}$ for all $t \in [0,1]$. So equation (1) has a unique solution $X_t = (\omega_t^\sharp)^{-1}\theta$. Also since $\theta(0) = 0$, we have $X_t(0) = 0$. Therefore

- the isotopy φ_t is defined for all $t \in [0,1]$ on a neighborhood 0 in $\mathbb{R}^{2n}$, and
- $\varphi_t(0) = 0$.

Also by construction, $\iota(X_t)\omega_t = \theta$, hence $d\iota(X_t)\omega_t = d\theta$, hence .... $\frac{d}{dt}(\varphi_t^*\omega_t) = 0$ for all t, hence $\varphi_t^*\omega_t = \varphi_0^*\omega_0 = \omega_0$.

4. Lecture 4. Isotopies and time-dependent vector fields

Definition 23. Let N be a manifold. A family of maps $f_t : N \longrightarrow N, t \in [0,1]$ is an **isotopy** if

1. each f_t is a diffeomorphism,
2. $f_0 = id_N$, and
3. f_t depend smoothly on t, i.e., the map $[0,1] \times N \longrightarrow N$, $(t,x) \longmapsto f_t(x)$ is smooth.

A time-dependent vector field $\{X_t\}_{t\in[0,1]}$ is a smooth collection of vector fields, one for each t. Again "smooth" means the map $(t,x) \longmapsto X_t(x)$, $\mathbb{R} \times N \longrightarrow TN$ is smooth.

We will see shortly that there is a correspondence between isotopies and time-dependent vector fields.

Detour: vector fields and flows.

Recall that a manifold N is Hausdorff, if for all $x, x' \in N$, $x \neq x'$, there are neighborhoods $U \ni x, U' \ni x'$ with $U \cap U' = \emptyset$.

Recall that if N is a Hausdorff manifold and X a vector field on N, then for every point $x \in N$, there is a unique curve $\gamma_x(t)$ such that $\gamma_x(0) = x$ and $\frac{d}{dt}\gamma_x(t) = X(\gamma_x(t))$ for all t. The curve γ_x is called an integral curve of X. It is defined on some interval containing 0. The interval itself depends on the point x.

Example 10. Let $N = (0,1)$ and $X = \frac{d}{dt}$. Then $\gamma_x(t) = x + t$, hence $\gamma_x(t)$ is defined whenever $0 < x + t < 1$, that is, $-x < t < 1 - x$.

Example 11 (of a non-Hausdorff manifold)**.** Let $X = \mathbb{R} \times \{0,1\}$, topologized as the disjoint union of two copies of $\mathbb{R}$. Define an equivalence relation on X by setting $(x,\alpha) \sim (y,\beta)$ if and only if either $\alpha = \beta$ and $x = y$, or $\alpha \neq \beta$ and $x = y \neq 0$. Thus $\{(0,0)\}$ and $\{(0,1)\}$ each constitute a distinct equivalence class, but all other classes are pairs $\{x,0),(x,1)\}$ where $x \neq 0$. The quotient space $N = X/\sim$ is a non-Hausdorff manifold. It is easy to see that integral curves of $\frac{d}{dx}$ on N are not unique.

From now on all the manifolds will be assumed to be Hausdorff (unless noted otherwise).

Let X be a vector field on a manifold N. Recall that the **flow** of X on N is a map ϕ from a subset of $N \times \mathbb{R}$ into N given by

$$\phi : (x,t) \mapsto \gamma_x(t).$$

where as above $\gamma_x(t)$ is the integral curve of X passing through a point x at $t = 0$.

Example 12. Let $N = (0,1)$ and let $X = \frac{d}{dt}$. The flow of X on N is $\phi(x,t) = x + t$.

We next recall a few facts about flows.

1. The domain of the flow ϕ of a vector field X on a manifold N is an open subset of $N \times \mathbb{R}$.
2. Write $\phi_t(x) = \phi(x,t)$. Then $\phi_t(\phi_s(t)) = \phi_{t+s}(x)$ whenever both sides are defined. This is the "group property" of the flow.
3. If the set $\overline{\{x \in N, X(x) \neq 0\}}$ is compact, then the flow ϕ_t is defined for all t.

Note: it follows from (1) that if for some $x_0 \in N, \phi_t(x_0)$ exists for $t \in [0,T]$, then for some neighborhood U of x_0 in N such that for all $x \in U, \phi_t(x)$ exists for all $t \in [0,T]$. (This is because $\{x_0\} \times [0,T]$ is compact.)

In particular, if $X(x_0) = 0$, then $\phi_t(x_0) = x_0$ for all t. And then there is a neighborhood U of x_0 such that for all points $x \in U$, the flow $\phi_t(x)$ is defined for arbitrarily large values of t.

For a vector field X and its flow $\{\phi_t\}$, we have

- $\phi_0 = id$
- $\frac{d}{dt}\phi_t(x) = X(\phi_t(x))$. Hence in particular $\frac{d}{dt}\big|_{t=0}\phi_t(x) = X(x)$

This ends our brief summary of facts about vector fields and flows. We now turn to isotopies and time-dependent vector fields.

Given an isotopy, $\{f_t : N \to N\}$, we define the corresponding time-dependent vector field X_t by

$$X_t(f_t(x)) := \frac{d}{ds}\big|_{s=t} f_s(x).$$

An **integral curve of a time-dependent vector field** X_t is a curve γ such that $\frac{d}{ds}\gamma(s)|_{s=t} = X_t(\gamma(t))$.

Proposition 24. *Let N be a compact manifold and let $\{X_t\}$ a time-dependent vector field on N. Then there is an isotopy $\{f_t\}$ defined for all t such that*

$$\frac{d}{ds} f_s(x)|_{s=t} = X_t(f_t(x)),$$

that is, for each $x \in N$ the curve $\gamma_x(t) := f_t(x)$ is an integral curve of X_t.

Proof. Consider the vector field $\overline{X}(x,t) := (X_t(x), \frac{d}{dt})$ on the manifold $N \times \mathbb{R}$. Let F_t denote the flow of $\overline{X}$. Assume for a moment that $\{F_t\}$ is defined for all t. The flow F_t has the form $F_t(x,s) = (G_t(x,s), s+t)$ for some map $G : N \times \mathbb{R} \to N$.

Since F_t is a flow, we have $F_{-t}(F_t(x,s)) = (x,s) = (G_{-t}(G_t(x,s), s+t), s)$. So for t, s fixed, $x \mapsto G_t(x,s)$ is a diffeomorphism with inverse $y \mapsto G_{-t}(y, s+t)$. We define $f_t(x) = G_t(x,0)$. This is our isotopy. Finally observe that $\frac{d}{dt}G_t(x,0) = X_t(G_t(x,0))$, i.e. $\frac{d}{dt}f_t(x) = X_t(f_t(x))$.

It remains to check that the the flow $\{F_t\}$ is defined for all t. In fact it is enough to show that for any $(x_0, q) \in N \times \mathbb{R}$ and any $T > 0$, the integral curve of $\overline{X}(x,t) = (X_t(x), \frac{d}{dt})$ through (x_0, q) is defined for all $t \in [0,T]$. This is not immediate since $N \times \mathbb{R}$ is not compact.

Choose a smooth function ρ on $\mathbb{R}$ such that ρ has compact support. i.e. $\overline{\{t : \rho(t) \neq 0\}}$ is compact, and such that $\rho = 1$ on $[q, q+T]$. Consider $\overline{X_\rho}(x,t) = (X_t(x), \rho(t)\frac{d}{dt})$. Since N is compact, the vector field $\overline{X_\rho}$ is compactly supported on $N \times \mathbb{R}$ and $\overline{X_\rho} = \overline{X}$ for $t \in [q, q+T]$. Therefore the integral curves through (x_0, q) of the vector fields $\overline{X_\rho}$ and $\overline{X}$ are the same for the time $t \in [0,T]$. In particular the integral curve of $\overline{X}(x,t)$ through (x_0, q) is defined for all $t \in [0,T]$. ∎

Proposition 25. *Let $\{X_t\}$ be a time dependent vector field on a manifold N. Suppose for some $x_0 \in N, X_t(x_0) = 0$ for all t. Then there is a neighborhood U of x_0 in N, and a family of maps $f_t : U \longrightarrow N$, $t \in [0,1]$ such that $f_0 = id$, and for all t we have: $\frac{d}{dt}(f_t(x)) = X_t(f_t(x))$, $f_t(x_0) = x_0$, f_t is 1-1 and $(df_t)_x$ is invertible. i.e. f_t is a diffeomorphism from U to $f_t(U)$.*

Proof. Consider $\overline{X}(x,t) = (X_t(x), \frac{d}{dt})$. Since $\overline{X}(x_0,t) = (0, \frac{d}{dt})$, we have that $\gamma(t) = (x_0, t)$ is an integral curve of $\overline{X}$ through $(x_0, 0)$. This curve exists for all t, so there is a neighborhood V of $(x_0, 0)$ in $N \times \mathbb{R}$ such that the flow of $\overline{X}$ exists on V for $t \in [0,1]$. Let $U = V \cap (N \times \{0\})$. ∎

Homework Problem 5. Consider the time-dependent vector field $X_t = t\frac{d}{d\theta}$ on S^1 where θdenotes the angle. Find the corresponding isotopy $\{\varphi_t\}$.

Homework Problem 6. Consider two 2-forms $\omega_0 = dx \wedge dy, \omega_1 = (1 + x^2 + y^2)dx \wedge dy$ on $\mathbb{R}^2$. Find a time-dependent vector field X_t such that its isotopy $\{\varphi_t\}$ satisfies $\varphi_1^*\omega_1 = \omega_0$. Hint: convert everything to polar coordinates.

5. Lecture 5. Poincaré lemma

Having gotten the preliminaries on time-dependent vector fields and isotopies out of the way, we now start a proof of the Darboux theorem. Recall the statement:

Theorem 26 (Darboux)**.** *Let ω_0, ω_1 be two symplectic forms on a manifold M. Suppose for some $x_0 \in M, \omega_0(x_0) = \omega_1(x_0)$. Then there are neighborhoods U_0, U_1 of x_0 in M and a diffeomorphism $\varphi : U_0 \longrightarrow U_1$ a such that $\varphi^*\omega_1 = \omega_0$.*

Proof. (Due to Moser and Weinstein). We define a family of forms ω_t by $\omega_t := t\omega_1 + (1-t)\omega_0$, $t \in [0,1]$. We'd like to construct an isotopy $\{\varphi_t\} : U \hookrightarrow M$ on a neighborhood U of x_0[2] such that for all $t \in [0,1]$ we have $\varphi_t^*\omega_t = \omega_0$ and $\varphi_t(x_0) = x_0$, and such that $\varphi_0(x) = x$ for all $x \in U$.

If such an isotopy exists, then $\frac{d}{dt}(\varphi_t^*\omega_t) = 0$. This equation should impose a condition on the time-dependent vector field X_t generated by the isotopy φ_t. Let us find out what this condition is.

[2]Strictly speaking $\{\varphi_t\}$ is a family of open embeddings and not a family of diffeomorpishm.

By the chain rule, for a (perhaps vector valued) function $G(\alpha, \beta)$ of two real variables we have

$$\frac{d}{dt}(G(t,t)) = \frac{\partial G}{\partial \alpha}(t,t) + \frac{\partial G}{\partial \beta}(t,t).$$

Therefore

$$\frac{d}{dt}(\varphi_t^* \omega_t) = \frac{d}{ds}(\varphi_s^* \omega_t)|_{s=t} + \frac{d}{ds}(\varphi_t^* \omega_s)|_{s=t}.$$

Exercise 6. Show that $\frac{d}{ds}(\varphi_t^* \omega_s)|_{s=t} = \varphi_t^*(\frac{d}{ds}\omega_s|_{s=t})$.

Let us now examine the term $\frac{d}{ds}(\varphi_s^* \omega_t)|_{s=t}$.

Recall that by the definition of a Lie derivative, if $\{\psi_t\}$ is the flow of a vector field Y and if ν is form then

$$\frac{d}{dt}(\psi_t^* \nu) = \psi_t^*(L_Y \nu).$$

Recall also Cartan's formula:

$$L_Y \nu = d(\iota(Y)\nu) + \iota(Y) d\nu$$

An analogous statement holds for isotopies and the corresponding time-dependent vector fields.

Proposition 27. *Let $\{\varphi_t\}$ be an isotopy on a manifold N and let X_t be the time-dependent vector field generated by $\{\varphi_t\}$. Then*

$$\frac{d}{dt}(\varphi_t^* \nu) = \varphi_t^*(L_{X_t} \nu) \tag{2}$$

for any form ν on N.

Proof. Fix $t \in [0,1]$. Define two operators $Q_1$and Q_2 on differential forms $\Omega^*(N)$ on N by

$$Q_1(\nu) = \frac{d}{dt}(\varphi_t^* \nu),$$
$$Q_2(\nu) = \varphi_t^*(L_{X_t} \nu).$$

We want to show that $Q_1 = Q_2$.

Let us show first that $Q_i \circ d = d \circ Q_i$ for $i = 1, 2$. Recall that the exterior differentiation d commutes with pullbacks: $d \circ \varphi_t^* = \varphi_t^* \circ d$. Applying $\frac{d}{dt}$ to both sides gives us $Q_1 \circ d = d \circ Q_1$. Similarly since $d \circ \varphi_t^* = \varphi_t^* \circ d$ and since $L_{X_t} \circ d = d \circ L_{X_t}$ we get $Q_2 \circ d = d \circ Q_2$.

We next show that $Q_i(\nu \wedge \mu) = Q_i(\nu) \wedge \varphi_t^* \mu + \varphi_t^* \nu \wedge Q_i(\mu)$ for $i = 1, 2$. Note that $\varphi_t^*(\nu \wedge \mu) = (\varphi_t^* \nu) \wedge (\varphi_t^* \mu)$. Now let's differentiate both sides with respect to t. We get

$$\frac{d}{dt}\varphi_t^*(\nu \wedge \mu) = (\frac{d}{dt}(\varphi_t^* \nu)) \wedge (\varphi_t^* \mu) + (\varphi_t^* \nu) \wedge \frac{d}{dt}(\varphi_t^* \mu).$$

Similarly, since L_{X_t} is a derivation,

$$L_{X_t}(\nu \wedge \mu) = (L_{X_t} \nu) \wedge \mu + \nu \wedge (L_{X_t} \mu).$$

The result for Q_2 now follows by applying φ_t^* to both sides of the above equation.

We now check that Q_1 and Q_2 agree on zero forms, i.e., on functions. Let f be a function on N. Then at a point $x \in N$

$$\frac{d}{dt}(\varphi_t^* f)(x) = \frac{d}{dt}(f(\varphi_t(x)) =$$
$$\langle df(\varphi_t(x), X_t(\varphi_t(x))\rangle = (L_{X_t} f)(\varphi_t(x)) =$$
$$\varphi_t^*(L_{X_t} f)(x).$$

Therefore $Q_1(f) = Q_2(f)$.

Since locally any differential form is a sum of expressions of the form $f\, dx_1 \wedge \cdots \wedge x_m$, the result follows. $\square$

We conclude that there is an isotopy φ_t such that $\varphi_t^* \omega_t = \omega_0$ if and only if

$$0 = \varphi_t^* L_{X_t} \omega_t + \varphi_t^*(\frac{d}{dt}\omega_t).$$

Since φ_t is an open embedding, φ^* is injective. Hence the desired isotopy exists if and only if

$$0 = L_{X_t}\omega_t + \frac{d}{dt}\omega_t. \tag{3}$$

Next note that since ω_1, ω_0 are closed, $d\omega_t = d(t\omega_1 + (1-t)\omega_0) = td\omega_1 + (1-t)d\omega_0 = 0$. Hence $L_{X_t}\omega_t = (d\iota(X_t) + \iota(X_t)d)\omega_t = d\iota(X_t)\omega_t$. Therefore equation (3) is equivalent to

$$d(\iota(X_t)\omega_t) = \omega_0 - \omega_1. \tag{4}$$

Lemma 28 (Poincaré)**.** *Let ν be a closed k-form ($k > 0$) on a subset U of $\mathbb{R}^n$ such that $0 \in U$ and such that for any $v \in U$ we have $tv \in U$ for all $0 \le t \le 1$.*[3] *Then there is a $(k-1)$-form μ such that $\nu = d\mu$.*

Let us assume the lemma for a moment. Then there is a one-form μ on some neighborhood of x_0 in M such that $\omega_0 - \omega_1 = d\mu$ on this neighborhood. Since μ is defined up to an exact form we may also assume that $\mu(x_0) = 0$.

Since $\omega_0 - \omega_1 = d\mu$, equation (4) is implied by the equation

$$\iota(X_t)\omega_t = \mu. \tag{5}$$

Equation (5) has a unique solution X_t if the form ω_t is nondegenerate. Now $\omega_t(x_0) = t\omega_1(x_0) + (1-t)\omega_0(x_0) = \omega_0(x_0)$ is nondegenerate for all $t \in [0,1]$. Therefore for all $t \in [0,1]$ the form $\omega_t(x)$ is nondegenerate for all x in a small neighborhood of x_0. On this neighborhood we can define X_t by

$$X_t = (\omega_t)^{-1}(\mu).$$

Then $\iota(X_t)\omega_t = \mu$, hence $d\iota(X_t)\omega_t = d\mu = \omega_0 - \omega_1 = -\frac{d}{dt}\omega_t$. Therefore $L_{X_t}\omega_t + \frac{d}{dt}\omega_t = 0$. Note that since by construction $\mu(x_0) = 0$, we have $X_t(x_0) = 0$ as well. Therefore on a sufficiently

[3]Such subsets are called **star-shaped**.

small neighborhood of x_0 the isotopy φ_t of X_t is defined for all $t \in [0,1]$. This implies that $\varphi_t^* L_{X_t}\omega_t + \varphi_t^* \frac{d}{dt}\omega_t = 0$, which, in turn, implies that

$$\frac{d}{dt}(\varphi_t^*\omega_t) = 0, \quad \text{for all } t \in [0,1].$$

Since φ_0 is the identity, we conclude that

$$\varphi_1^*\omega_1 = \varphi_0^*\omega_0 = \omega_0.$$

This finishes our proof of the Darboux theorem modulo the Poincaré lemma. □

Corollary 29. *Suppose ω_0, ω_1 are two symplectic forms on M. Let $x_0 \in M$ be a point. Then there exist open neighborhoods U_0, U_1 of x_0 and a diffeomorphism $\varphi : U_1 \longrightarrow U_0$ with $\varphi(x_0) = x_0$ such that $\varphi^*\omega_0 = \omega_1$.*

Proof. Exercise. Hint: recall the linear case first. □

Proof of Poincaré lemma. Consider the radial vector field $R(v) := v$ (in coordinates $R(x_1, \ldots, x_n) = \sum x_i \frac{\partial}{\partial x_i}$.) Its flow ψ_t is given by $\psi_t(v) = tv$. Then $\psi_1(v) = v$ for all $v \in U$ and $\psi_0(v) = 0$ for all v. Hence $\psi_1^*\nu = \nu, \psi_0^*\nu = 0$. Therefore,

$$\nu = \psi_1^*\nu - \psi_0^*\nu = \int_0^1 \frac{d}{dt}(\psi_t^*\nu)\, dt = \int_0^1 \psi_t^*(L_R\nu)\, dt = \int_0^1 \psi_t^*(d\iota(R) + \iota(R)d)\nu\, dt.$$

Since the form ν is closed, we have $\nu = \int_0^1 (\psi_t^*(d\iota(R)\nu)\, dt = d(\int_0^1 (\psi_t^*(\iota(R)\nu))\, dt)$. So $\mu = \int_0^1 \psi_t^*(\iota(R)\nu)\, dt$ is the desired $(k-1)$-form. □

6. Lecture 6. Lagrangian embedding theorem

Recall that a subspace L of a symplectic vector space (V, ω) is Lagrangian if (and only if) $\dim L = \frac{1}{2}\dim V$ and for all $\ell, \ell' \in L$ we have $\omega(\ell, \ell') = 0$. Equivalently, L is a Lagrangian subspace if (and only if) $L^\omega = L$.

Definition 30. A submanifold L of a symplectic manifold (M, ω) is **Lagrangian** if and only if $\forall x \in L$, T_xL is Lagrangian subspace of $(T_xM, \omega(x))$, i.e., $\dim L = \frac{1}{2}\dim M$ and $\omega(x)(v, v') = 0$ for all $v, v' \in T_xL$.

Equivalently we say that an embedding j of a manifold L into a symplectic manifold (M, ω) is **Lagrangian** if (1) $\dim L = \frac{1}{2}\dim M$ and (2) $j^*\omega = 0$.

Example 13. Let X be a manifold. Then its cotangent bundle T^*X is naturally a symplectic manifold with the symplectic form $\omega = \omega_{T^*X}$. Recall that if $x_1, \ldots, x_n$ are coordinates on X and $x_1, \ldots, x_n, \xi_1, \ldots, \xi_n$ are the corresponding coordinates on T^*X, then in these coordinates $\omega_{T^*X} = \sum dx_i \wedge d\xi_i$. The manifold X embeds into T^*X as the set of zero covectors. This embedding X in (T^*X, ω_{T^*X}) is Lagrangian.

The Lagrangian embedding theorem asserts that the above example of a Lagrangian embedding is, in an appropriate sense, the only example.

Theorem 31 (Lagrangian embedding theorem)**.** *If $X \subset (M, \omega)$ is a Lagrangian submanifold, then there is a neighborhood U of X in M, a neighborhood U' of X in T^*X (where, as before, we think of X as zero covectors), and a diffeomorphism $\varphi : U \longrightarrow U'$ such that (1) $\varphi^*(\omega_{T^*X}) = \omega$ and (2) $\varphi(x) = x$ for all $x \in X$.*

The Lagrangian embedding theorem is due to A. Weinstein. The theorem has a number of applications. For example it is at the heart of the method of generating functions. Also it is prototypical for a number of other theorems that are quite useful. Finally it gives us an excuse to introduce a number of ideas such as vector bundles, tubular neighborhoods, smooth homotopy invariance of de Rham cohomology, almost complex structures, geodesic flows as Hamiltonian systems ...

Homework Problem 7. Let X a manifold, and let α denote the tautological (Liouville) one-form on T^*X. Every one-form $\mu \in \Omega^1(X)$ is a map $\mu : X \longrightarrow T^*X$. Show that $\mu^*\alpha = \mu$. Use this to show that $\mu(X)$ is a Lagrangian submanifold of $(T^*X, d\alpha)$ if and only if $d\mu = 0$.

We now begin an introduction of the tools that we will use to prove the Lagrangian embedding theorem.

Vector bundles. Informally, a vector bundle is a family of isomorphic vector spaces parameterized by points of a manifold. More precisely,

Definition 32. A smooth map $\pi : E \to M$ of one manifold onto another is a smooth **real vector bundle of rank** k if the following conditions are satisfied:

- For each $x \in M$ the set $\pi^{-1}(x) = E_x$, called the **fiber** above x, is a real vector space of dimension k.
- For every $x \in M$, there exists a neighborhood U of x in M and a diffeomorphism $\varphi : \pi^{-1}(U) \longrightarrow U \times \mathbb{R}^k$ such that $\varphi(E_y) \subseteq \{y\} \times \mathbb{R}^k$ for all $y \in U$, i.e., the diagram

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\phi"] \arrow[dr, "\pi"'] & & U \times \mathbf{R}^k \arrow[dl, "pr_1"] \\ & U & \end{tikzcd}$$

  commutes, where pr_1 is the projection on the first factor, and $\varphi : E_y \longrightarrow \{y\} \times \mathbb{R}^k$ is a vector space isomorphism for all $y \in U$.

The map $\varphi : \pi^{-1}(U) \longrightarrow U \times \mathbb{R}^k$ is called a **local trivialization** of E.

For a vector bundle $\pi : E \to M$, the manifold E is called the **total space** and M is called the **base**.

Given a vector bundle $\pi : E \to M$ we will often say that E is a vector bundle over M.

Example 14. Let M be a manifold.

The tangent bundle $TM \xrightarrow{\pi} M$ is a vector bundle of rank $k = \dim M$ over M.

The cotangent bundle $T^*M \longrightarrow M$ is a vector bundle over M of rank $k = \dim M$.

$M \times \mathbb{R}^n \longrightarrow M$ is a vector bundle of rank n over M called **the trivial vector bundle**.

Let $E \xrightarrow{\pi_E} M$ and $F \xrightarrow{\pi_F} M$ be two vector bundles over a manifold M. A **morphism (map)** of vector bundles is a smooth map $\psi : E \longrightarrow F$ such that for all $x \in M$ we have $\psi(E_x) \subset F_x$, and the restriction of ψ to the fibers $\psi : E_x \longrightarrow F_x$ is a linear map. A morphism of vector bundles which has an inverse is called an **isomorphism**.

Example 15. Let (M, ω) be a symplectic manifold. The map $\omega^\sharp : TM \longrightarrow T^*M$ defined by $v \mapsto \omega(\pi(v))(v, \cdot)$ is an isomorphism of vector bundles.

Example 16. If (M, g) is a Riemannian manifold (i.e. if g is a Riemannian metric on the manifold M), the map $g^\sharp : TM \longrightarrow T^*M$ defined by $v \mapsto g(\pi(v))(v, \cdot)$ is an isomorphism of vector bundles.

Definition 33. A vector bundle $F \longrightarrow M$ is a **subbundle** of a vector bundle $E \longrightarrow M$ if F is a subset of E and if the inclusion $\psi : F \hookrightarrow E$ is a map of vector bundles. Equivalently, a subbundle of a bundle E is a pair (F, j) where F is a vector bundle and $j : F \to E$ an **injective** map of vector bundles.

Proposition 34. *Suppose $\pi : E \longrightarrow M$ is a vector bundle and $N \hookrightarrow M$ is an embedded submanifold, then $E|_N := \pi^{-1}(N)$ is a vector bundle called the* **restriction** *of E to N.*

Proof. The map π is a submersion. □

The above construction can be generalized as follows. Let $\pi : E \longrightarrow M$ be a vector bundle and let $f : N \to M$ be a smooth map of manifolds. We define the **pull-back** bundle $f^*E \to N$ to be the set

$$f^*E = \{(x, e) \in N \times E \mid f(x) = \pi(e)\}$$

with the projection $f^*E \to N$ given by $(x, e) \mapsto x$. It is not hard to see that f^*E is a manifold and that it is, in fact, a vector bundle with a fiber $(f^*E)_x = E_{f(x)}$.

Example 17. Let $N \hookrightarrow M$ be an embedded submanifold. The restriction of the tangent bundle of M to N is the bundle $T_N M := TM|_N = \coprod_{x \in N} T_x M$. The tangent bundle TN is a subbundle of $T_N M$.

Example 18. Let us make the above example more concrete. Consider the round sphere $S^2 \subset \mathbb{R}^3$, $S^2 = \{x \in \mathbb{R}^3 || x|^2 = 1\}$. The restriction of the tangent bundle of $\mathbb{R}^3$ to the sphere is $T_{S^2} T\mathbb{R}^3 = \{ (x, v) \in S^2 \times \mathbb{R}^3 \}$, and tangent bundle to the sphere is $TS^2 = \{ (x, v) \in S^2 \times \mathbb{R}^3 \mid x \cdot v = 0 \}$.

Example 19 (conormal bundle). Let $N \subset M$ embedded submanifold. The conormal bundle $\nu^*(N)$ is the subbundle of $T^*M|_N$ defined by

$$\nu^*(N) = \{ \xi \in T^*M|_N \,:\, \xi|_{T_{\pi(\xi)}N} = 0 \}$$

It is not hard to see that the conormal bundle is a vector bundle by studying it in coordinates. In more details, let $n = \dim N$ and $m = \dim M$. Since N is an embedded submanifold of M, for every point p of M there are coordinates $x_1, \dots, x_m$ on M defined on some open neighborhood U of p such that $U \cap N = \{x_{n+1} = x_{n+2} = \cdots = 0\}$. Let $x_1, \dots, x_m, \xi_1, \dots, \xi_m$ be the corresponding coordinates on T^*M. Then

$$\nu^*(N)|_{N \cap U} = \{x_{n+1} = x_{n+2} = \cdots = 0, \xi_1 = \xi_2 = \cdots = \xi_n = 0\}.$$

Therefore $\nu^*(N)$ is a submanifold of the cotangent bundle T^*M. It is easy to check it is in fact a subbundle of $T^*M|_N$.

Note that the dimension of the conormal bundle is $\dim \nu^*(N) = 2m - (m-n) - n = m = \frac{1}{2} \dim T^*M$. In fact, the conormal bundle is a Lagrangian submanifold of T^*M. This is easy to see in coordinates: $\omega_{T^*M} = \sum_{i=1}^m dx_i \wedge d\xi_i$ and $T_{(x,\xi)}(\nu^*(N)) = span\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}, \frac{\partial}{\partial \xi_{n+1}}, \dots, \frac{\partial}{\partial \xi_m}\}$

Definition 35. Let $\pi : E \longrightarrow M$ be a vector bundle. A **section** s of the vector bundle E is a map $s : M \longrightarrow E$ such that $\pi(s(x)) = x$ for all $x \in M$.

Example 20. A vector field on M is a section of of the tangent bundle TM. A one-form is a section of the cotangent bundle T^*M.

For every vector bundle $E \longrightarrow M$, there is a section called the **zero section**:

$$\begin{aligned} 0 : M &\longrightarrow E \\ x &\longmapsto 0_x \quad \text{zero in} \quad E_x \end{aligned}$$

We will often refer to the **image** of the zero section as the zero section.

Theorem 36 (Tubular neighborhood)**.** *If $N \hookrightarrow M$ is an embedded submanifold, there exists a neighborhood U of N in M, a neighborhood U_0 of the zero section in the conormal bundle $\nu^*(N)$ and a diffeomorphism $\psi : U \longrightarrow U_0$ such that $\psi(x) = x$ for all $x \in N$.*

We will prove the theorem in the next section.

Normal bundles. The statement of the tubular neighborhood theorem above is a little non-standard. In this subsection we describe the more standard statement and the concepts associated with it.

We start by defining local sections. A **local section** of a vector bundle $E \longrightarrow M$ is a section of $E|_U$ for some open subset U of M. Local section always exist, for given a sufficiently small subset U of M the restriction $E|_U$ is isomorphic to $U \times \mathbb{R}^k$ (by definition of a vector bundle) and $U \times \mathbb{R}^k \longrightarrow U$ has lots of sections.

Next we define a Whitney sum of two vector bundles. There are two equivalent ways of defining the sum. Here is one. Suppose $E \xrightarrow{\pi_E} M$ and $F \xrightarrow{\pi_F} M$ are two vector bundles. Then we have a vector bundle $E \times F \xrightarrow{\pi_E \times \pi_F} M \times M$. The manifold M embeds into $M \times M$ as the diagonal:

$$\begin{aligned} \triangle : M &\longrightarrow M \times M, \\ x &\longmapsto (x, x) \end{aligned}$$

We define the **sum** $E \oplus F \longrightarrow M$ to be the pull-back $\triangle^*(E \times F)$. Note that the fiber above x of$E \oplus F$ is $(E \oplus F)_x = \{(e, f) \in E \times F \mid \pi_E(e) = \pi_F(f) = x\} = E_x \oplus F_x$.

Having defined direct sums, we can now try and define quotients. Let $E \to M$ be a vector bundle and let F be a subbundle of E. If we can find a subbundle F' of E such that $E = F \oplus F'$ then it would make sense to define the quotient E/F to be F'.

Now suppose that E has a **metric** g, that is a smoothly varying inner product on the fibers E_x of E. Here "smoothly varying" means : if s, s' are any two local sections, then $x \longmapsto g(x)(s(x), s'(x))$ is smooth function of x. Using partitions of unity one can show that metrics always exist.

Given a vector bundle E with a metric g and a subbundle F, we define the complementary bundle F^g by $F^g = \coprod_{x \in M} F_x^g$, where F_x^g is the orthogonal complement of F_x in E_x with respect to the inner product $g(x)$. One shows that F^g is a vector bundle and that $E = F \oplus F^g$. The proof that F^g is a vector bundle amounts to doing Gram-Schmidt.

Another way to define quotients is to define E/F to be

$$\coprod_{x \in M} E_x/F_x.$$

One then needs to prove that E/F is a vector bundle. This is a more invariant way of thinking about quotients.

We can now define normal bundles. Let M be a manifold and let $N \hookrightarrow M$ be an embedded submanifold. The **normal bundle** of N in M is, by definition

$$\nu(N) = T_N M/TN.$$

This is a definition in the category of differential manifolds. Note that the fiber of the conormal bundle $\nu^*(N)_x = T_x N^\circ$ is the annihilator of $T_x N$ in $T_x^* M$. Now given a vector space V and a subspace W, the annihilator W° is naturally isomorphic to $(V/W)^*$. Therefore the fibers of the normal and of the conormal bundle are dual vector spaces.

Given a vector bundle $E \to M$ one can define the dual bundle $E^* \to M$ whose fibers are dual to the fibers of E. The normal and the conormal bundles of an embedding are an example of this. Another example are the tangent and the cotangent bundles of a manifold.

If M is a Riemannian manifold with a metric g, then, tautologically, the tangent bundle TM has the metric g. The Riemannian normal bundle of N is the bundle

$$\nu_g(N) = TN^g, \quad \nu_g(N)_x = T_x N^g \quad \text{for all } x \in N.$$

This definition depends on the metric and defines $\nu_g(N)$ as a particular subbundle of $T_N M$. Recall that the metric g defines the bundle isomorphism $g^\sharp : T_N M \longrightarrow T_N^* M$; it is given by $(x, v) \longmapsto g(x)(v, \cdot)$. It is easy to see that $(g^\sharp)^{-1} \nu^*(N) = TN^g = \nu_g(N)$. Since the preimage of a subbundle under a bundle isomorphism is a subbundle this provides a proof that the Riemannian normal bundle is a vector bundle.

A more standard version of the tubular neighborhood theorem reads:

Theorem 37 (Riemannian tubular neighborhood theorem)**.** *If $N \hookrightarrow M$ is an embedded submanifold of a Riemannian manifold (M, g), there exists a neighborhood U of N in M, a neighborhood U_0 of the zero section in the normal bundle $\nu_g(N)$, and a diffeomorphism $\psi : U \longrightarrow U_0$ such that $\psi(x) = x$ for all $x \in N$.*

7. Lecture 7. Proof of the tubular neighborhood theorem

Homework Problem 8. Let $\pi : E \longrightarrow M$ be a vector bundle. Then M embedds into E as the zero section. Suppose σ_1 and σ_2 are two closed k-forms on E, $(k > 0)$, such that for every $x \in M$, $\sigma_1(x) = \sigma_2(x)$. Show there is a $(k-1)$-form ν on E such that

1. $\sigma_1 - \sigma_2 = d\nu$
2. $\nu(x) = 0$ for all $x \in M$

Hint: Let $\sigma = \sigma_1 - \sigma_2$. Consider $\rho_t : E \longrightarrow E$, $t \in [0, 1]$, given by $\rho_t(v) = tv$. Then $\rho_0(v) = \pi(v)$ and so $\rho_0^*\sigma = 0$ while $\rho_1^*\sigma = \sigma$. Next write $\sigma = \rho_1^*\sigma - \rho_0^*\sigma = \dots$

Our goal for this lecture is to prove Theorem 36, the tubular neighborhood theorem.

Proof of the tubular neighborhood theorem. Choose a Riemannian metric g on M. This produces a metric g^* on the cotangent bundle (the so called dual metric) as follows. Recall that $g^\sharp : TM \longrightarrow T^*M$ an isomorphism of vector bundles defined by $g^\sharp(x, v) = g(x)(v, \cdot), x = \pi(v)$. Define the inner product $g^*(x)$ on T_x^*M by $g^*(x)(\xi, \eta) = g(x)((g^\sharp)^{-1}(\xi), (g^\sharp)^{-1}(\eta))$ for $\xi, \eta \in T_x^*M$.

Let us see what this metric looks like in coordinates. Let $x_1, \dots, x_n$ be coordinates on M. Define $g_{ij}(x) = g(x)(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$. Then

$$g^\sharp(x)(\frac{\partial}{\partial x_i}) = \sum g_{ij}(x) dx_j.$$

The reader should check that the matrix $(g^{ij}(x))$ defined by

$$g^{ij}(x) = g^*(x)(dx_i, dx_j)$$

satisfies $\sum_j g^{ij}(x) g_{jk}(x) = \delta_{ik}$, i.e. the matrices $(g_{ij}(x))$ and $(g^{ij}(x))$ are inverses of each other.

Define a function $H : T^*M \longrightarrow \mathbb{R}$ by $H(x, \xi) = \frac{1}{2} g^*(x)(\xi, \xi)$ for $\xi \in T_x^*M$; it is the so called "kinetic energy" function.

Recall that the cotangent bundle T^*M has a natural symplectic form $\omega = \omega_{T^*M}$ and that $\omega = -d\alpha$, where α is the tautological 1-form. Recall also that in coordinates $\omega = \sum dx_i \wedge d\xi_i$.

The function H and the symplectic form on T^*M define a vector field X by the equation $\iota(X)\omega = dH$. The flow of X is called the **geodesic flow** of the metric g. We will see that projections onto M of integral curves of X "minimize" distances in M, i.e. they are geodesics.[4] For this reason we will refere to the flow ϕ_t of the vector field X as geodesic flow.

[4]The geodesics are extremals of the distance funcitonal and not necessarily the local minima hence the quotation marks.

Let us compute the vector field X in coordinates. We have $dH = d(\frac{1}{2}\sum_{i,j} g^{ij}(x)\xi_i\xi_j) = \frac{1}{2}\sum_{i,j,k}\frac{\partial g^{ij}}{\partial x_k}\xi_i\xi_j dx_k + \frac{1}{2}\sum_{i,j} g^{ij}\xi_j d\xi_i + \frac{1}{2}\sum_{i,j} g^{ij}\xi_i d\xi_j = \frac{1}{2}\sum_{i,j,k}\frac{\partial g^{ij}}{\partial x_k}\xi_i\xi_j dx_k + \sum_{i,j} g^{ij}\xi_i d\xi_j$ since $g^{ij} = g^{ji}$. Consequently,

$$X(x,\xi) = \sum_{i,k} g^{ik}\xi_i \frac{\partial}{\partial x_k} - \frac{1}{2}\sum_{i,j,k}\frac{\partial g^{ij}}{\partial x_k}\xi_i\xi_j\frac{\partial}{\partial \xi_k}.$$

Proposition 38. *For any $w \in T_xM$ we have*

$$d\pi(X(x, g^\sharp(x)w)) = w.$$

Proof. One can give a coordinate-free proof of this fact, but a computation in coordinates is very simple. We write $w = \sum w_l \frac{\partial}{\partial x_l}$. Then $g^\sharp(x)w = g^\sharp(x)(\sum w_l \frac{\partial}{\partial x_l}) = \sum_{l,i} g_{l,i}(x) w_l dx_i$. Therefore

$$\begin{aligned} d\pi(X(x, g^\sharp(x)w)) &= \sum_{i,k} g^{ik}(x)(g^\sharp(x)w)_i \frac{\partial}{\partial x_k} = \sum_{i,k,l} g^{ik}(x) g_{li}(x) w_l \frac{\partial}{\partial x_k} \\ &= \sum_{k,l}\delta_{kl} w_l \frac{\partial}{\partial x_k} = \sum_k w_k \frac{\partial}{\partial x_k} = w. \end{aligned}$$

□

For every $s \neq 0$, we have a diffeomorphism $\rho_s : T^*M \longrightarrow T^*M$ defined by $\rho_s(x,\xi) = (x, s\xi)$.

Lemma 39. $X(\rho_s(x,\xi)) = d\rho_s(sX(x,\xi))$

Proof. We compute in coordinates. Since $\rho_s(x_1,\ldots,x_n,\xi_1,\ldots,\xi_n) = (x_1,\ldots,x_n,s\xi_1,\ldots,s\xi_n)$ we have

$$d\rho_s(\frac{\partial}{\partial x_k}) = \frac{\partial}{\partial x_k} \quad \text{and} \quad d\rho_s(\frac{\partial}{\partial \xi_k}) = s\frac{\partial}{\partial \xi_k}.$$

Therefore,

$$X(\rho_s(x,\xi)) = \sum g^{ik} s\xi_i \frac{\partial}{\partial x_k} - \frac{1}{2}\sum \frac{\partial g^{ij}}{\partial x_k} s^2 \xi_i\xi_j \frac{\partial}{\partial \xi_k}$$

while

$$\begin{aligned} d\rho_s(sX(x,\xi)) = s\,d\rho_s(X(x,\xi)) &= s\,d\rho_s(\sum g^{ik}\xi_i\frac{\partial}{\partial x_k} - \frac{1}{2}\sum\frac{\partial g^{ij}}{\partial x_k}\xi_i\xi_j\frac{\partial}{\partial \xi_k}) \\ &= s(\sum g^{ik}\xi_i\frac{\partial}{\partial x_k} - \frac{1}{2}\sum\frac{\partial g^{ij}}{\partial x_k}\xi_i\xi_j s\frac{\partial}{\partial \xi_k}). \end{aligned}$$

□

Corollary 40. *Let $\gamma_\xi(t)$ denote the integral curve of of the vector field X through a covector $\xi \in T^*M$ (i.e., suppose $\gamma_\xi(0) = \xi$ and $\frac{d}{dt}\gamma_\xi(t) = X(\gamma_\xi(t))$.) Then $\rho_s(\gamma_\xi(st)) = \gamma_{s\xi}(t)$. In terms of the geodesic flow ϕ_t we have*

$$\rho_s(\phi_{st}(\xi)) = \phi_t(\rho_s(\xi)).$$

Proof. Since $\rho_s(\gamma_\xi(s,0)) = \rho_s(\xi) = s\xi$ and since integral curves are unique, it is enough to check

$$\frac{d}{dt}\rho_s(\gamma_\xi(st)) = X(\rho_s(\gamma_\xi(st))).$$

We compute:

$$\frac{d}{dt}\rho_s(\gamma_\xi(st)) = d\rho_s(s\frac{d\gamma_\xi}{dt}(st)) = d\rho_s(sX(\gamma_\xi(st))) = X(\rho_s(\gamma_\xi(st)))$$

by Lemma 39. □

We can now define the exponential map $\exp^* : T^*M \longrightarrow M$ (determined by a metric g):

$$\exp^*(\xi) = \pi(\phi_1(\xi)),$$

where $\pi : T^*M \to M$ is the bundle projection and ϕ_1 is the time 1 map of the geodesic flow. Note that if we want to emphasize the base point of a covector ξ we would write

$$\exp^*(x,\xi) = \pi(\phi_1(x,\xi)),$$

where $x = \pi(\xi)$, or, equivalently $\xi \in T^*_x M$.

Similarly we define $\exp : TM \longrightarrow M$ by

$$\exp(x,v) = \pi\left(\phi_1(x, g^\sharp(x)v)\right)$$

for $v \in T_xM$, where $TM \xrightarrow{g^\sharp} T^*M$ is the usual metric induced isomorphism of vector bundles. Thus $\exp = \exp^* \circ g^\sharp$.

Since $X(x,0) = 0$ we have $\phi_t(x,0) = x$ for all $x \in M$. Hence the flow $\phi_t(x,\xi)$ is defined up to time 1 in some sufficiently small neighborhood of the zero section in T^*M. We conclude that the maps $\exp^*$ and $\exp$ are defined in some neighborhoods of the zero sections.

Next we compute the differential of the exponential map exp at the points of the zero seciton. We start with an observation that if $\pi : E \longrightarrow M$ is a vector bundle and $(x,0) \in E$ is a point on the zero section then the tangent space $T_{(x,0)}E$ is naturally isomorphic to the direct sum $T_xM \oplus (T_0E_x) = T_xM \oplus E_x$ since both M and the fiber E_x are submanifolds of E passing transversely through $(x,0)$. (There is no such natural splitting at points $e \in E$ off the zero section because there is no subspace of T_eE naturally isomorphic to $T_{\pi(e)}M$ unless $\pi(e) = e$.) In particular we have a natural identification

$$T_{(x,0)}(TM) \simeq T_xM \oplus T_xM.$$

Proposition 41. *The differential of the exponential map* $\exp : TM \to M$ *at the points of the zero section is given by*

$$d(\exp)(x,0)(v \oplus w) = v + w$$

for all $v \oplus w \in T_xM \oplus T_xM$.

In particular the proposition asserts that with respect to a basis of T_xM the differential $d(\exp)(x,0)$ is given by a matrix of the form

$$\begin{array}{cc} \overset{\text{zero section}}{T_xM} & \overset{\text{fiber}}{T_xM} \end{array}$$
$$\left(\begin{array}{ccc|ccc} 1 & & 0 & 1 & & 0 \\ & \ddots & & & \ddots & \\ 0 & & 1 & 0 & & 1 \end{array}\right).$$

Proof. Since $\exp(x,0) = x$ for all $x \in M$, $d\exp|_{\text{zero section}} = id_{T_xM}$. We now argue that $d\exp|_{\text{fiber}} = id_{T_xM}$ as well. Fix $w \in T_xM$. We want to show that $\frac{d}{dt}|_{t=0}\exp(x,tw) = w$. Now,

$$\begin{aligned}
\frac{d}{dt}|_{t=0}\exp(x,tw) &= \frac{d}{dt}|_{t=0}\pi\phi_1(x, g^\sharp(x)(tw)) \\
&= \frac{d}{dt}|_{t=0}\pi(\phi_1(x, tg^\sharp(x)w)) \\
&= \frac{d}{dt}|_{t=0}\pi(\rho_t(\phi_t(x, g^\sharp(x)w))) \quad (\text{ by Corollary 40}) \\
&= \frac{d}{dt}|_{t=0}\pi(\phi_t(x, g^\sharp(x)w))) \quad (\text{since } \pi\circ\rho_t = \pi) \\
&= d\pi(X(x, g^\sharp(x)w) = w \quad (\text{by Proposition 38}).
\end{aligned}$$

□

The following is the key point.

Corollary 42. $\exp|_{T_xM} : T_xM \longrightarrow M$ *is a local diffeomorphism on a neighborhood of* 0 *in* T_xM*.*

Proof. This follows from the inverse function theorem since $d(\exp|_{T_xM})(0) = id$. □

Proposition 43 (Tubular neighborhood theorem, version 3)**.** *Let N be an embedded submanifold of a Riemannian manifold (M,g) Then the restriction of the exponential map to the normal bundle* $\exp : TN^g \longrightarrow M$ *is a diffeomorphism on some neighborhood of the zero section. Moreover* $\exp(x) = x$ *for all $x \in N$.*

Example 21. Consider the round sphere $S^2 \subset \mathbb{R}^3$. Let g be the standard inner product on $\mathbb{R}^3$. It is easy to see that $(TS^2)^g = \{(x,v) \in S^2 \times \mathbb{R}^3 \,|\, v = \lambda x, \text{for some } \lambda \in \mathbb{R}\}$. Check that the exponential map $\exp : (TS^2)^g \longrightarrow \mathbb{R}^3$ is given by $(x,v) \mapsto x+v$. It is a diffeomorphism for v sufficiently small (less than the radius of the sphere).

Note that the result is false if N is injectively immersed but not embedded. For example take $M = \mathbb{R}^2$ and let $N = (0,1)$ be immersed as figure eight. The image of N has no tubular neighborhood.

Proof of the tubular neighborhood theorem, verison 3. Note first that at the points of the zero section the tangent space to the normal bundle $T_{(x,0)}(TN^g)$ is isomorphic to $T_xN \oplus (T_xN)^g \simeq$

T_xM. Note also that $d(\exp)(x,0) : T_xN \oplus (T_xN)^g \longrightarrow T_xM$ is given by $(v,w) \longmapsto v+w$. Hence $d(\exp)(x,0)$ is an isomorphism for all $x \in N$. Therefore exp is a local diffeomorphism in a neighborhood of every point of the zero section, i.e. for any $x \in N$, there exists neighborhoods O_x of x in TN^g and O'_x of x in M such that $\exp : O_x \longrightarrow O'_x$ is a diffeomorphism.

To finish the proof we need a topological lemma.

Lemma 44. *Suppose $N \subset M$ embedded submanifold and suppose there is a smooth map $\psi : TN^g \longrightarrow M$ such that*

1. *$\psi(x) = x$ for all $x \in N$, and*
2. *ψ is a local diffeomorphism in a neighborhood of any $x \in N$.*

Then there exists neighborhoods U_0 of N in the normal bundle TN^g, U of N in M such that $\psi : U_0 \longrightarrow U$ is a diffeomorphism.

Proof. See Bröker & Jänich, *Introduction to Differential Topology*, Lemma 12.6. □

This finishes the proof of version 3 of the tubular neighborhood theorem. □

Note that since the "musical" isomorphism $g^\sharp : TM \to T^*M$ identifies the metric normal bundle TN^g with the conormal buncle $\nu^*(N)$, the first version of the tubular neighorhood theorem, Theorem 36, follows as well. □

Exercise 7. Let $L \subset (M,\omega)$ be an embedded Lagrangian submanifold. Show that $\nu^*(L) \simeq T^*L$. If you get stuck see the next lecture.

8. Lecture 8. Proof of the Lagrangian embedding theorem. Almost complex structures

Recall that if (M,ω) symplectic manifold, an embedding $i : L \longrightarrow M$ is Lagrangian iff $\dim L = \frac{1}{2}\dim M$ and $i^*\omega = 0$.

Lemma 45. *If $i : L \longrightarrow (M,\omega)$ is a Lagrangian embedding, then the conormal bundle of L in M is isomorphic to the tangent bundle: $\nu^*(L) \simeq TL$.*

Proof. Recall that $\omega^\sharp : TM \longrightarrow T^*M$ defined by $\omega^\sharp(x,v) = \omega(x)(v,\cdot)$ is an isomorphism of vector bundles. For all $x \in L$ and for all $v,w \in T_xL$, we have $0 = \omega(x)(v,w) = \langle \omega^\sharp(x)(v), w\rangle$. Therefore $\omega^\sharp(T_xL)$ is contained in the set $\{\xi \in T_x^*M \mid \xi|_{T_xL} = 0\}$, i.e. $\omega^\sharp(TL) \subset TL^\circ$. On the other hand, $\dim T_xL^\circ = \dim M - \dim L = \dim L$. Hence $\omega^\sharp(T_xL) = T_xL^\circ$. Therefore $\omega^\sharp : TL \longrightarrow TL^\circ = \nu^*(L)$ is an isomorphism. □

Theorem 46 (Lagrangian embedding, version 1)**.** *Let $(M_0,\omega_0), (M_1,\omega_1)$ be two symplectic manifolds. $i_0 : L \longrightarrow M_0, i_1 : L \longrightarrow M_1$ be two Lagrangian embeddings with respect to ω_0 and ω_1 respectively. Then there exist neighborhoods U_0 of $i_0(L)$ in M_0, U_1 of $i_1(L)$ in M_1 and a diffeomorphism $\varphi : U_0 \longrightarrow U_1$ such that $\varphi^*\omega_1 = \omega_0$ and $\varphi(i_0(x)) = i_1(x)$ for all $x \in L$.*

Here is an equivalent version.

Theorem 47 (Lagrangian embedding, version 2)**.** *If $i : L \longrightarrow (M, \omega)$ is a Lagrangian embedding, then there exist neighborhoods U of $i(L)$ in M, U_0 of L in T^*L and a diffeomorphism $\varphi : U_0 \longrightarrow U$ such that $\varphi(x) = i(x)$ for all $x \in L$ and such that $\varphi^*\omega = \omega_{T^*L}$ where ω_{T^*L} is the standard symplectic form on the cotangent bundle T^*L.*

We will see later on that Lagrangian manifolds arise in the study of completely integrable systems. Here is another way in which Lagrangian manifolds come up. Suppose (M_1, ω_1) and (M_2, ω_2) are two symplectic manifolds with $\dim M_1 = \dim M_2$. Suppose $f : M_1 \longrightarrow M_2$ a smooth map. We have two obvious projections

$$\begin{tikzcd} & M_1 \times M_2 \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ M_1 & & M_2 \end{tikzcd}$$

Define $\omega = \pi_1^*\omega_1 - \pi_2^*\omega_2$. (This form is also written as $\omega_1 \oplus (-\omega_2)$ or simply as $\omega_1 - \omega_2$). The form ω is symplectic (check this!).

Recall that the graph of f is the set $\mathrm{graph}(f) = \{(m, f(m)) \in M_1 \times M_2\}$; it is a submanifold of $M_1 \times M_2$. Clearly its dimension $\dim \mathrm{graph}(f)$ is $\dim M_1$ which is half the dimension of the product $\dim(M_1 \times M_2)$. The tangent space to the graph $T_{(m,f(m))}(\mathrm{graph}(f))$ is the subspace $\{(v, u) \in T_mM_1 \times T_{f(m)}M_2 \,|\, u = df_m(v)\}$. Now for $(v, df(m)(v)), (w, df(m)(w) \in T_{(m,f(m))}(\mathrm{graph}(f))$ we have

$$\begin{aligned} \omega((m, f(m))((v, dfv), (w, dfw)) &= (\pi_1^*\omega_1 - \pi_2^*\omega_2)((m, f(m))((v, dfv), (w, dfw)) \\ &= \omega_1(m)(v, w) - \omega_2(f(m))(dfv, dfw) \\ &= \omega_1(m)(v, w) - (f^*\omega_2)(m)(v, w), \end{aligned}$$

where we wrote dfv for $df(m)v$ and dfw for $df(m)w$. Hence $\mathrm{graph}(f)$ is a Lagrangian submanifold of $(M_1 \times M_2, \pi_1^*\omega_1 - \pi_2^*\omega_2)$ if and only if $\omega_1 - f^*\omega_2 = 0$, i.e., iff f is symplectic.

Proof of Theorem 46. By Lemma 45 and by the tubular neighborhood theorem, we may assume that $M_0 = M_1 = TL$ and that ω_0, ω_1 are two symplectic forms on TL with $\omega_0|_L = 0$ and $\omega_1|_L = 0$. We would like to find neighborhoods U_0 and U_1 of L in TL, and a diffeomorphism $\varphi : U_0 \longrightarrow U_1$ such that $\varphi|_L = id$, and $\varphi^*\omega_1 = \omega_0$.

Example 22. To appreciate the kind of a problem that we are confronting consider $L = \mathbb{R}^2$. Then $TL = \mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^4$. Let (x_1, x_2, y_1, y_2) be coordinates on TL. Suppose $\omega_0 = dx_1 \wedge dy_1 + dx_2 \wedge dy_2$ and $\omega_1 = dx_1 \wedge dy_2 + dx_2 \wedge dy_1$. These forms are symplectic and $\omega_i|_L = 0$ for $i = 0, 1$.

The proof of the theorem is a modification of Moser's deformation argument. We first prove it under an extra assumption:

$$\omega_1(x) = \omega_0(x), \quad \forall x \in L, \tag{6}$$

i.e. for all $x \in L$ the forms $\omega_0(x)$ and $\omega_1(x)$ agree on $T_{(x,0)}(TL)$. Later on we will see that such an assumption is justified (Proposition 53).

Let us now assume that (6) holds and let us look for a one-form τ such that $\tau(x) = 0$ for all $x \in L$ and such that $\omega_1 - \omega_0 = d\tau$.

Consider $\rho_t : TL \longrightarrow TL$ defined by $\rho_t(x, v) = (x, tv)$. Then $\rho_1 = id$, and for all $(x, v) \in TL$ we have $\rho_0(x, v) = (x, 0)$ and $\rho_t(x, 0) = (x, 0)$ for all t. Therefore $\rho_0^*\omega_i = 0$ and $\rho_1^*\omega_i = \omega_i$ for $i = 1, 2$. We now compute, as in the proof of the Poincaré lemma:

$$\omega_1 - \omega_0 = \rho_1^*(\omega_1 - \omega_0) - \rho_0^*(\omega_1 - \omega_0) = \int_0^1 (\frac{d}{dt}\rho_t^*(\omega_1 - \omega_0))\, dt.$$

Let $R(x, v) = \frac{d}{dt}\big|_0 \rho_t(x, v)$; it is (the analog of) the radial vector field. Then

$$\begin{gathered}\omega_1 - \omega_0 = \int_0^1 \rho_t^*(L_R(\omega_1 - \omega_0))\, dt = \\ \int_0^1 (\rho_t^*(d\iota(R) + \iota(R)d)(\omega_1 - \omega_0))\, dt = \\ \int_0^1 \rho_t^* d\iota(R)(\omega_1 - \omega_0)\, dt = \\ \int_0^1 d\rho_t^*(\iota(R)(\omega_1 - \omega_)))\, dt = \\ d\int_0^1 \rho_t^*(\iota(R)(\omega_1 - \omega_0))\, dt.\end{gathered}$$

We define $\tau = \int_0^1 \rho_t^*(\iota(R)(\omega_1 - \omega_0))\, dt$.

For all $x \in L$ we have $(\omega_1 - \omega_0)(x) = 0$ by the extra assumption. Since $\rho_t(x) = (x)$ for all $x \in L$, we have $R(x) = 0$. Hence $\iota(R)(\omega_1 - \omega_0))(x) = 0$ and so $[\rho_t^*(\iota(R)(\omega_1 - \omega_0))](x) = 0$. Consequently $\tau(x) = \left(\int_0^1 \rho_t^*(\iota(R)(\omega_1 - \omega_0))\, dt\right)(0) = 0$ for all $x \in L$.

Next, let $\omega_t = t\omega_1 + (1 - t)\omega_0$, $t \in [0, 1]$. For $x \in L$ and for all $t \in [0, 1]$ we have $\omega_t(x) = t\omega_1(x) + (1 - t)\omega_0(x) = \omega_1(x)$ by (6). Hence the form $\omega_t(x)$ is nondegenerate for all t and all $x \in L$. Consequently there is a neighborhood W of L in TL on which the form ω_t is nondegenerate for all t (here we used that $[0, 1]$ is compact). Therefore on W we can find a time-dependent vector field X_t such that $\iota(X_t)\omega_t = -\tau$: we set $X_t := ((\omega_t)^\sharp)^{-1}(-\tau)$. Then $X_t(x, 0) = 0$ for all $x \in L$. Hence the flow φ_t of X_t exists for all $t \in [0, 1]$ on a perhaps smaller neighborhood $W' \subset W$. Consequently $d\iota(X_t)\omega_t = -d\tau = -(\omega_1 - \omega_0) = -\frac{d}{dt}\omega_t$. It follows that $L_{X_t}\omega_t + \frac{d}{dt}\omega_t = 0$ and hence $0 = \varphi_t^* L_{X_t}\omega_t + \varphi_t^* \frac{d}{dt}\omega_t = \frac{d}{dt}(\varphi_t^*\omega_t) = 0$. Consequently $\varphi_t^*\omega_t = \varphi_0^*\omega_0 = \omega_0$ for all t. We finally conclude that $\varphi_1^*\omega_1 = \omega_0$ where the open embedding φ_1 is defined on some neighborhood of L in TL.

This finishes the proof of Theorem 46 modulo the simplifying assumption that (6) holds. □

We now start the work required to show that the assumption is justified.

Almost Complex Structure.

Definition 48. Let V be a real vector space. A linear map $J : V \longrightarrow V$ such that $J^2 = -id$ is called a **complex structure** on the vector space V.

Example 23. Let $V = \mathbb{C}^n$. The map $J(z) = \sqrt{-1}z$ is a complex structure.

If a real vector space V has a complex structure J, then V can be made into a complex vector space by defining $(a + \sqrt{-1}b)v = av + bJv$ for all $v \in V$, $a, b \in \mathbb{R}$.

Example 24. The point of this example is to observe that a symplectic vector space always has a complex structrue. In the next lecture we will see a proof that does not use the exitence of symplectic basis.

Let (V, ω) be symplectic vector space, let $e_1, \dots, e_n, f_1, \dots, f_n$ be a symplectic basis (so that $\omega = \sum e_i^* \wedge f_i^*$ in terms of the dual basis $e_1^*, \dots, e_n^*, f_1^*, \dots, f_n^*$). Define $J : V \to V$ on the basis by $Je_i = f_i$, $Jf_i = -e_i$. Then clearly $J^2 = -id$.

Note that $\omega(e_i, Je_i) = \omega(e_i, f_i) = +1$ and that $\omega(f_i, Jf_i) = \omega(f_i, -e_i) = +1$ hence for all $v \in V$ we have $\omega(v, Jv) \geq 0$ and $= 0$ if and only if $v = 0$. Let $g(v, w) = \omega(v, Jw)$. We claim that g is a positive definite inner product. Since $\omega(Je_i, Jf_i) = \omega(f_i, -e_i) = \omega(e_i, f_i)$ we have $\omega(Jv, Jw) = \omega(v, w)$ for all $v, w \in V$. Hence $g(w, v) = \omega(w, Jv) = \omega(Jw, J^2v) = \omega(Jw, -v) = \omega(v, Jw) = g(v, w)$, i.e. g is symmetric.

We conclude that every symplectic vector space (V, ω) has a complex structure J such that $g(v, w) = \omega(v, Jw)$ is a positive definite inner product. Such a complex structure is called **compatible** with the symplectic form ω.

Remark 49. Let (V, ω) be a symplectic vector space and let $J : V \to V$ is be a complex structure. It is not hard to see that $g(v, w) := \omega(v, Jw)$ is a symmetric bilinear form if and only if J is symplectic, i.e., $J^*\omega = \omega$.

Indeed suppose g is symmetric. Then $\omega(v, Jw) = g(w, v) = g(v, w) = \omega(w, Jv) = \omega(Jv, -w) = \omega(Jv, J^2w)$ for any $v, w \in V$. Since J is onto, any $u \in V$ is of the form $u = Jw$ for some w. Hence $\omega(v, u) = \omega(v, Jw) = \omega(Jv, J^2w) = \omega(Jv, Ju)$.

Conversely, suppose $\omega(v, u) = \omega(Jv, Ju)$ for all $v, u \in V$. Then $g(w, v) = \omega(v, Jw) = \omega(Jv, J^2w) = \omega(Jv, -w) = \omega(w, Jv) = g(w, v)$.

Definition 50. An **almost complex structure** J on a manifold M is a vector bundle map $J : TM \longrightarrow TM$ such that $J^2 = -id$.

Clearly every complex manifold has an almost complex structure. Note however that there are manifolds with almost complex structures which are not complex manifolds.

We will prove in the next lecture:

Theorem 51. *Every symplectic manifold (M, τ) has an almost complex structure J compatible with τ, i.e. there exists an almost complex structure J on M such that $g(\cdot, \cdot) := \omega(\cdot, J\cdot)$ is a Riemannian metric on M*

We end this lecture with a prototypical application of the existence of compatible (almost) complex structures.

Lemma 52. *Let (V, ω) be a symplectic vector space, let $L \subset V$ be a Lagrangian subspace and let J a complex structure compatible with the symplectic form ω. Then the subspace JL is also Lagrangian, $L \oplus JL = V$ and the linear map $\psi : L \oplus JL \longrightarrow L^* \oplus L$ defined by $\psi(v \oplus w) = \omega(w, \cdot) \oplus v$ satisfies $\psi^*\omega_0 = \omega$, where ω_0 is the standard form on $L \oplus L^*$ (c.f. Proposition 18).*

Proof. Since $\omega(Jv, Jw) = \omega(v, w)$ for all $v, w \in V$, the subspace JL is Lagrangian. (Recall that by Remark 49 the bilinear form $g(v, w) := \omega(v, Jw)$ is symmetric if and only if $\omega(Jv, Jw) = \omega(v, w)$).

It remains to show that $L \cap JL = \{0\}$. Suppose $u \in L \cap JL$. Then $u = Jv$ for some $v \in L$. We have $\omega(v, u) = \omega(v, Jv) = g(v, v) \geq 0$. But $u, v \in L$ so we must have $\omega(v, u) = 0$. Therefore $g(v, v) = 0$, hence $v = 0$ and so $u = J0 = 0$ as well. □

Homework Problem 9. Let M be a compact manifold. Suppose that ω_0 and ω_1 are two symplectic forms on M such that $\omega_1 - \omega_0 = d\tau$ for some $\tau \in \Omega^1(M)$. Prove that if the form $\omega_t = t\omega_1 + (1-t)\omega_0$ is symplctic for all $t \in [0, 1]$ then there is a diffeomorphism $f : M \longrightarrow M$ such that $f^*\omega_1 = \omega_0$.

Homework Problem 10 (de Rham cohomology)**.** Let M be a manifold. As usual let $\Omega^k(M)$ denote the set of differential k-forms on M and let $d : \Omega^k(M) \longrightarrow \Omega^{k+1}(M)$ denote the exterior differentiation. The k-th deRham cohomology group $H^k(M)$ is by definition the quotient $ker\{d : \Omega^k(M) \longrightarrow \Omega^{k+1}(M)\}/im\{d : \Omega^{k-1}(M) \longrightarrow \Omega^k(M)\}$. For $\tau \in \Omega^k(M)$with $d\tau = 0$, let $[\tau]$ denote the class of τ in $H^k(M)$.

1. (Functoriality) Show that given a smooth map $f : N \longrightarrow M$, the map in cohomology $H(f) :: H^k(M) \longrightarrow H^k(N)$ given by $H(f)([\tau]) = [f^*\tau]$ is well-defined. Show also that $H(f \circ g) = H(g) \circ H(f)$.
2. Prove that if M is connected, then $H^0(M) = \mathbb{R}$.
3. Prove that

$$H^k(S^1) = \begin{cases} \mathbb{R}, & k = 0, 1 \\ 0, & k > 1 \end{cases} \tag{7}$$

   Hint: The map $\Omega^1(S^1) \longrightarrow \mathbb{R}$, $\alpha \mapsto \int_{S^1} \alpha$ may be useful.
4. (Smooth homotopy invariance) Two smooth maps $f_0, f_1 : M \longrightarrow N$ are smoothly **homotopic** iff there is a smooth map $F : [0, 1] \times M \longrightarrow N$ such that $F(1, x) = f_1(x)$ and $F(0, x) = f_0(x)$. Show that if f_1 and f_0 are homotopic, then $H(f_0) = H(f_1)$. Hints: Define $i_s : M \longrightarrow \mathbb{R} \times M$ by $i_s(m) = (s, m)$. Then $F(s, x) = F(i_s(x))$. Show that there are maps $Q_k : \Omega^k(\mathbb{R} \times M) \longrightarrow \Omega^{k-1}(M)$ such that $d(Q_k\tau) + Q_{k+1}(d\tau) = i_1^*\tau - i_0^*\tau$ for any $\tau \in \Omega^k(\mathbb{R} \times M)$ by considering $\int_0^1 \frac{d}{dt}(i_t^*\tau)\, dt$. Conclude that $H(i_1) = H(i_0)$. Finish the proof by observing that $f_s(x) = (F \circ i_s)(x), s = 0, 1$ and use functoriality.

9. Lecture 9. Almost complex structures and Lagrangian embeddings

The goal of this lecture is to finish the proof of the Lagrangian embedding theorem. We accomplish it by proving the existence of an almost complex structure compatible with a given symplectic form (Theorem 51) and by proving Proposition 53 below.

Proof of Theorem 51. We first consider the proof in the setting of vector spaces. Let V be a vector space and τ a skew-symmetric nondegenerate bilinear form. Choose a positive definite inner product g on V. We have two isomorphisms:

$$\tau^{\#} : V \to V^*, \quad v \mapsto \tau(v, \cdot)$$

and

$$g^{\#} : V \to V^*, \quad v \mapsto g(v, \cdot).$$

Let $A = (g^{\#})^{-1} \circ \tau^{\#}$. Then for any $v, w \in V$ we have

$$g(Av, w) = \langle g^{\#} Av, w\rangle = \langle \tau^{\#} v, w\rangle = \tau(v, w) = -\tau(w, v) = -g(Aw, v) = -g(v, Aw),$$

i.e., $A = -A^*$ where the adjoint is taken relative to the metric g. Therefore $-A^2 = AA^*$ is diagonalizable and all eigenvalues are positive. Let P be the positive square root of $-A^2$. For example we can define P by

$$P = \frac{1}{2\pi\sqrt{-1}} \int_\gamma (-A^2 - z)^{-1} \sqrt{z}\, dz,$$

where $\sqrt{z}$ is defined via the branch cut along the negative real axis and γ is a contour containing the spectrum of $-A^2$. It follows that P commutes with A and that

$$(AP^{-1})^2 = A^2 P^{-2} = A^2(-A^2) = -1.$$

The map $J = AP^{-1}$ is the desired complex structure.

Note that the same argument works if we consider a symplectic manifold (M, τ). We choose a Riemannian metric g on M and consider a vector bundle map $A = (g^{\#})^{-1} \circ \tau^{\#}$. We define $P : TM \to TM$ by essentially the same formula: for $x \in M$ the map $P_x : T_xM \to T_xM$ on the fiber above x is given by

$$P_x = \frac{1}{2\pi\sqrt{-1}} \int_{\gamma_x} (-A_x^2 - z)^{-1} \sqrt{z} dz,$$

Note that since the spectrum of A_x varies with the base point x and since we don't assume that the base is compact, we have to let the contour γ_x vary with x as well to make sure that the spectrum of $-A_x^2$ lies inside γ_x. The map P so defined is a smooth vector bundle map that commutes with A and we set the complex structure J to be AP^{-1}. $\square$

Proposition 53. *Let ω_1 and ω_2 be two symplectic forms on a manifold M. Suppose $L \subset M$ is a submanifold, which is Lagrangian for both ω_1 and ω_2. Then there exist neighborhoods U_1, U_2 of L in M and a diffeomorphism $F : U_1 \longrightarrow U_2$ such that $F\big|_L = id$ and such that $(F^*\omega_2)(x) = \omega_1(x)$ for all $x \in L$.*

Proof. Suppose (V,ω) is a symplectic vector space, $L \subset V$ is a Lagrangian subspace and $J : V \longrightarrow V$ is a complex structure, compatible with ω. Then the image JL of L under J is perpendicular to L with respect to the inner product $g(\cdot,\cdot) = \omega(\cdot, J\cdot)$: $JL = L^g$. Therefore $V = L \oplus JL$. Moreover, since $\omega(J\cdot, J\cdot) = \omega(\cdot,\cdot)$, JL is a Lagrangian subspace of V. Recall that the map $A : L \oplus JL \to L^* \oplus L$ given by $A(l,l') = \omega(\cdot, l') \oplus l$ has the property that $A^*\omega_0 = \omega$ where ω_0 is the canonical symplectic form on $L^* \oplus L$ (cf. Example 8 and Proposition 18).

Suppose now that ω_1, ω_2 are two symplectic forms on a vector space V, and the subspace $L \subset V$ is a Lagrangian with respect to both forms. We want to construct an isomorphism $F : V \to V$ with $F|_L = id_L$ such that $F^*\omega_2 = \omega_1$.

Let J_1 be a complex structure on V compatible with ω_1 and let J_2 be a complex structure on V compatible with ω_2. We define F to be the composition $(L \oplus J_1 L, \omega_1) \xrightarrow{\simeq} (L \oplus L^*, \omega_0) \xrightarrow{\simeq} (L \oplus J_2 L, \omega_2)$. The map F has the desired property.

We now adapt this proof to the setting of manifolds. Choose almost complex structures J_1, J_2 compatible with the symplectic forms ω_1, ω_2 respectively. Let $g_i(\cdot,\cdot) = \omega_i(\cdot, J_i\cdot)$ $(i = 1,2)$ be the corresponding Riemannian metrics so that the subbundle $J_i TL$ is the perpendicular subbundle TL^{g_i} $(i = 1,2)$. By the tubular neighborhood theorem there are neighborhoods U_i of L in M and U_i^0 of L in TL^{g_i} and such that the exponential maps $\exp_i : U_i^0 \longrightarrow U_i$ determined by the metrics g_i are diffeomorphisms $(i = 1,2)$.

We also have vector bundle isomorphisms $\varphi_i : J_i(TL) \longrightarrow T^*L$ defined by $\varphi_i(x,v) = (x, \omega_i(x)(\cdot, v))$ for $v \in J_i(TL)_x$. Define $f : J_1TL \to J_2TL$ by $f = \varphi_2^{-1} \circ \varphi_1$. Denote by $f(x)$ the restriction of f to the fiber above x; then $f(x) : J_1T_xL \to J_2T_xL$. As in the vector space case for every point $x \in L$ the map $id \oplus f(x) : T_xM = T_xL \oplus J_1T_xL \to T_xL \oplus J_2T_xL = T_xM$ satisfies $[(id \oplus f(x))^*\omega_2](x) = \omega_1(x)$.

Define $F : U_1 \to M$ by $F = \exp_2 \circ f \circ (\exp_1|_{U_1})^{-1}$:

$$\begin{tikzcd} J_1TL \arrow[r, "f"] & J_2TL \arrow[d, "\exp_2"] \\ U_1 \arrow[u, "(\exp_1)^{-1}"] \arrow[r, dotted, "F"] & M \end{tikzcd}$$

By construction we have that $F|_L = id_L$. We are done if we can show that for any point $x \in L$ the differential $dF(x) : T_xM \to T_xM$ satisfies $(dF(x)^*\omega_2)(x) = \omega_1(x)$. We therefore need to compute the map $dF(x)$. Note that since the differentials $d\exp_i(0) : T_xM \to T_xM$, $i = 1,2$, are the identity maps, we need only to compute $df(x)$ and prove that it is symplectic.

Now suppose $E \longrightarrow B, E' \longrightarrow B$ are two vector bundles over B and suppose $A : E \longrightarrow E'$ a vector bundle map. Let's compute the differential dA of A at the points of the zero section. We have $T_{(x,0)}E = T_xB \oplus E_x$. The map A restricted to the zero section is the identity, and A restricted to the fiber E_x is a linear map which we denote by $A(x) : E_x \to E'_x$. Therefore $dA(x,0) : T_xB \oplus E_x \to T_xB \oplus E'_x$ is simply $id \oplus A(x)$. Applying this argument to the map f

we see that $df(x) = id \oplus f(x)$. Since $(id \oplus f(x))^*\omega_2(x) = \omega_1(x)$ by construction of f, we are done. $\square$

This finishes the proof of the Lagrangain embedding theorem.

10. Lecture 10. Hamilton's principle. Euler-Lagrange equations

10.1. **Classical system of N particles in $\mathbb{R}^3$.** .

Consider a mechanical system consisting of N particles in $\mathbb{R}^3$ subject to some forces. Let $x_i \in \mathbb{R}^3$ denote the position vector of the ith particle. Then all possible positions of the system are described by N-tuples $(x_1, \ldots, x_N) \in (\mathbb{R}^3)^N$. The space $(\mathbb{R}^3)^N$ is an example of a **configuration space**. The time evolution of the system is described by a curve $(x_1(t), \ldots, x_N(t))$ in $(\mathbb{R}^3)^N$ and is governed by Newton's second law:

$$m_i \frac{d^2 x_i}{dt^2} = F_i(x_1, \ldots, x_N, \dot{x}_1, \ldots, \dot{x}_N, t)$$

where F_i denotes the force on ith particle (which depends on the positions and velocities of all N particles and on time), $\dot{x}_i = \frac{dx_i}{dt}$, and m_i denotes the mass of the ith particle.

Let us now re-lable the variables. Let q_{3i}, q_{3i+1} and q_{3i+2} be respectively the first, the second and the third coordinate of the vector x_i, $i = 1, \ldots, N$. The configuaration space is then $\mathbb{R}^n$, $n = 3N$ and the equations of motion take the form

$$m_\alpha \frac{d^2 q_\alpha}{dt^2} = F_\alpha(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t), \quad 1 \leq \alpha \leq n. \tag{8}$$

Let us now suppose that the forces are time-independent and conservative, that is, that there exists a function $V : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $F_\alpha(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t) = F_\alpha(q_1, \ldots, q_n) = -\frac{\partial V}{\partial q_\alpha}(q_1, \ldots, q_n)$.

Example 25. For example if N particles interact by gravitational attraction, $V(x_1, \ldots, x_N) = -\gamma \sum_{i \neq j} \frac{m_i m_j}{\|x_i - x_j\|}$, where γ is a universal constant.

Then equation (8) takes the form

$$m_\alpha \frac{d^2 q_\alpha}{dt^2} = -\frac{\partial V}{\partial q_\alpha}(q_1, \ldots, q_n), \quad 1 \leq \alpha \leq n. \tag{9}$$

We now rewrite equation (9) as a first order ODE by doubling the number of variables. Let us call the new variables v_α:

$$\begin{cases} m_\alpha \frac{dv_\alpha}{dt} = -\frac{\partial V}{\partial q_\alpha}(q_1, \ldots, q_n) \\ \frac{dq_\alpha}{dt} = v_\alpha \end{cases} \tag{10}$$

A solution $(q(t), v(t))$ of the above equation is a curve in the tangent bundle $T\mathbb{R}^n$ such that $\dot{q}(t) = v(t)$. The tangent bundle $T\mathbb{R}^n$ is an example of a **phase space**.

We now define a function $L(q,v) = \frac{1}{2}\sum_\alpha m_\alpha v_\alpha^2 - V(q)$ on the phase space $T\mathbb{R}^n$. It is the difference of the kinetic energy and the potential energy. We will see shortly that we can re-write (10) as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_\alpha}\right) - \frac{\partial L}{\partial q_\alpha} = 0, \quad 1 \le \alpha \le n \tag{11}$$

The function L is called the **Lagrangian** of the system (the name has almost nothing to do with Lagrangian submanifolds). The equation (11), which we claim is equivalent to Newton's law of motion, is an example of the Euler-Lagrange equations. Let us now check that equations (10) and (11) are, indeed, the same. Since $\frac{\partial}{\partial v_\alpha}(\frac{1}{2}\sum_\beta m_\beta v_\beta^2) = m_\alpha v_\alpha$ and $\frac{\partial}{\partial q_\alpha}(\frac{1}{2}\sum m_\beta v_\beta^2 - V) = -\frac{\partial V}{\partial q_\alpha} = F_\alpha$ we get $0 = \frac{d}{dt}(\frac{\partial}{\partial v_\alpha}L) - \frac{\partial L}{\partial q_\alpha} = \frac{d}{dt}(m_\alpha v_\alpha) + \frac{\partial V}{\partial q_\alpha}$. Hence $m_\alpha \frac{dv_\alpha}{dt} = -\frac{\partial V}{\partial q_\alpha}$.

So far introducing the Lagrangian did not give us anything new. We now show that it does indeed allow us to look at Newton's law from another point of view, and that the new point of view has interesting consequences.

10.2. **Variational formulation.** Let $L \in C^\infty(T\mathbb{R}^n)$ be a Lagrangian (i.e. a smooth function). Let $q^{(0)}, q^{(1)}$ be two points in $\mathbb{R}^n$. Consider all possible twice continuously differentiable (C^2) paths $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ with $\gamma(a) = q^{(0)}, \gamma(b) = q^{(1)}$. Denote by $\mathcal{P}(q^{(0)}, q^{(1)})$ the space of all such paths. The Lagrangian L defines a map $A_L : \mathcal{P}(q^{(0)}, q^{(1)}) \longrightarrow \mathbb{R}$ by $A_L(\gamma; q^{(0)}, q^{(1)}) = \int_a^b L(\gamma(t), \dot\gamma(t))\, dt$. This map is called an **action**.

Hamilton's principle: physical trajectories between two points $q^{(0)}, q^{(1)})$ of the system governed by the Lagrangian L are critical points of the action functional $A_L(\cdot; q^{(0)}, q^{(1)})$.

Proposition 54. *Hamilton's principle implies Euler-Lagrange equations and hence Newton's law of motion.*

Proof. Let $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ be a critical point (path, trajectory) of an action functional A_L; let $\gamma(t,\epsilon)$ be a family of paths depending on $\epsilon \in \mathbb{R}$ such that $\gamma(t,0) = \gamma(t)$ and such that for all ϵ we have $\gamma(a,\epsilon) = \gamma(a)$ and $\gamma(b,\epsilon) = \gamma(b)$ (i.e., we fix the end points).

Let $y(t) = \frac{\partial}{\partial\epsilon}\big|_{\epsilon=0}\gamma(t,\epsilon)$. Note that $y(a) = 0$ and $y(b) = 0$. Also $\frac{\partial}{\partial\epsilon}\big|_{\epsilon=0}(\frac{\partial}{\partial t}\gamma(t,\epsilon)) = \frac{\partial}{\partial t}(\frac{\partial}{\partial\epsilon}\big|_{\epsilon=0}\gamma(t,\epsilon)) = \dot y(t)$.

Conversely, given a curve $y : [a,b] \to \mathbb{R}^n$ with $y(a) = y(b) = 0$, we can find a family of paths $\gamma(t,\epsilon)$ with fixed end points such that $\gamma(t,0) = \gamma(t)$ and $\frac{\partial}{\partial\epsilon}\big|_{\epsilon=0}(\frac{\partial}{\partial t}\gamma(t,\epsilon)) = \dot y(t)$.

Now

$$
\begin{aligned}
0 &= \frac{d}{d\epsilon}\big|_{\epsilon=0} A_L(\gamma(t,\epsilon)) = \frac{d}{d\epsilon}\big|_{\epsilon=0} \int_a^b L(\gamma(t,\epsilon),\dot\gamma(t,\epsilon))\,dt \\
&= \int_a^b \frac{d}{d\epsilon}\big|_0 L\,dt = \sum_\alpha \int_a^b (\frac{\partial L}{\partial q_\alpha}\frac{\partial q_\alpha}{\partial \epsilon}\big|_{\epsilon=0} + \frac{\partial L}{\partial v_\alpha}\frac{\partial v_\alpha}{\partial \epsilon}\big|_{\epsilon=0})\,dt \\
&= \sum_\alpha \int_a^b (\frac{\partial L}{\partial q_\alpha} y_\alpha + \frac{\partial L}{\partial v_\alpha}\dot y_\alpha)\,dt \qquad (12) \\
&= \sum_\alpha \{\int_a^b \frac{\partial L}{\partial q_\alpha} y_\alpha\,dt + \frac{\partial L}{\partial v_\alpha} y_\alpha\big|_a^b - \int_a^b \frac{d}{dt}(\frac{\partial L}{\partial v_\alpha})\,y_\alpha\,dt\} \quad \text{(integration by parts)} \\
&= \sum_\alpha \int (\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}(\frac{\partial L}{\partial v_\alpha}))y_\alpha\,dt
\end{aligned}
$$

where $y_\alpha(t)$ are arbitrary functions on $[a,b]$ with $y_\alpha(a) = 0 = y_\alpha(b)$. Recall:

Lemma 55. *If $f \in C^1([a,b])$ and for any $y \in C^1$ with $y(a) = y(b) = 0$, we have $\int_a^b f(t)y(t)\,dt = 0$, then $f(t) \equiv 0$.*

We conclude that for any index α we must have $\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}(\frac{\partial L}{\partial v_\alpha}) = 0$. That is to say, Hamilton's principle implies the Euler-Lagrange equations. □

Note that we have proved that given a Lagrangian there is a vector field on $T\mathbb{R}^n$ whose integral curves are the critical curves of the corresponding action.

Constrained systems and d'Alembert principle.
Let us start with listing some examples of constrained systems.

Example 26 (Spherical pendulum)**.** The system consists of a massive particle in $\mathbb{R}^3$ connected by a very light rod of lenght ℓ to a universal joint. So the configuration space is a sphere S^2 of radius ℓ. The phase space is TS^2. This is an example of a holonomic constraint.

Example 27 (Free rigid body)**.** The system consists of N point masses in $\mathbb{R}^3$ maintaining fixed distances between each other: $\|x_i - x_j\| = \text{const}_{ij}$. We will see later on in Lecture 17 that the configuration space is $E(3)$, the Euclidean group of distance preserving transformations of $\mathbb{R}^3$. It is not hard to show that $E(3)$ consists of rotations and translations. In fact we can represent $E(3)$ as a certain collection of matrices:

$$E(3) \simeq \{\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{4^2} \mid A^T A = I, \det A = \pm 1, v \in \mathbb{R}^3\}$$

The group $E(3)$ is a 6-dimesional manifold, hence the phase space for a free rigid body is $TE(3)$. The free rigid body is also a holonomic system.

Example 28 (A quater rolling on a rough plane in an upright position (without slipping))**.** The configuration space is $(\mathbb{R}^2 \times S^1) \times S^1$, where the elements of $\mathbb{R}^2$ keep track of the point of contact of the quater with the rough plane, the points in the first S^1 keeps track of the orientation of the plane of the quater and points of the second S^1 keep track of the orientation of the design on the quater. The phase space of the system is smaller than $T(\mathbb{R}^2 \times S^1 \times S^1)$ becasue the point of contact of the quater with the plane has to be stationary. This is an example of non-holonomic constraints, since the constraints on position do not determine the constraints on velocity: the roll-no-slip condition is extra.

This leads us to a definition. Constraints are **holonomic** if the constraints on possible velocities are determined by the constraints on the configurations of the system. In other words if the constraints confine the configurations of the system to a submanifold M of $\mathbb{R}^n$ and the corresponding phase space is TM, then the constraints are holonomic.

We will study only holonomic systems with an added assumption: **constraint forces do no work.**

d'Alembert's principle: If constraint forces do no work, then the true physical trajectory of the system are extremals of the action functional of the free system restricted to the paths lying in the constraint submanifold.

Note that this principle is very powerful: we no longer need to know anything about the constraining forces except for the fact that they limit the possible configurations to a constraint submanifold. Let us see what kind of equation of motion d'Alembert's principle produces.

Let $M \subseteq \mathbb{R}^n$ be a submanifold and let $L : T\mathbb{R}^n \supseteq TM \longrightarrow \mathbb{R}$ be a Lagrangian for a free system. By d'Alembert's principle we should find paths $\gamma : [a, b] \longrightarrow M$ such that γ is critical for

$$A_L : \{\sigma : [a, b] \longrightarrow M \mid \sigma(a) = q^{(0)}, \sigma(b) = q^{(1)}\} \to \mathbb{R}$$
$$A_L(\sigma) = \int_a^b L(\sigma, \dot{\sigma})\, dt$$

Suppose the end points $q^{(0)}$ and $q^{(1)}$ lie in some coordinate patch on M. Let $(q_1, \ldots, q_n)$ be the coordinates on the patch and let $(q_1, \ldots, q_n, v_1, \ldots, v_n)$ be the corresponding coordinates on the correspondign patch in TM. The same argument as before (cf. Proposition 54) gives us Euler-Lagrange equations $\frac{d}{dt}(\frac{\partial L}{\partial v_\alpha}) - \frac{\partial L}{\partial q_\alpha} = 0$! Note that these equations represent a vector field on a coordinate patch in the tangent bundle TM.

Example 29 (Planar pendulum)**.** The system consists of a heavy particle in $\mathbb{R}^3$connected by a very light rod of lenght ℓ to a fixed point. Unlike the spherical pendulum the rod is only allowed to pivot in a fixed vertical plane. The configuration space M is the circle $S^1 \subset \mathbb{R}^2$ of radius ℓ. The Lagrangian of the free system is $L(x, v) = \frac{1}{2}m(v_1^2 + v_2^2) - mgx_2$ where m is the mass of the particle, x_1, x_2 coordinates on $\mathbb{R}^2$, x_1, x_2, v_1, v_2 the corresponding coordinates on

$T\mathbb{R}^2$ and g is the gravitational acceleration ($9.8\ m/s^2$). Let us now compute the equations of motion for the constraint system.

Consider the embedding $S^1 \longrightarrow \mathbb{R}^2, \varphi \mapsto (\ell \sin\varphi, -\ell\cos\varphi)$. The correpsonding embedding $TS^1 \longrightarrow T\mathbb{R}^2$ is given by $(\varphi, v_\varphi) \mapsto (\ell\sin\varphi, -\ell\cos\varphi, \ell\cos\varphi v_\varphi, \ell\sin\varpi v_\varphi)$. Therefore the constraint Lagrangian is given by $L(\varphi, v_\varphi) = \frac{1}{2}m(\ell^2\cos^2\varphi v_\varphi^2 + \ell^2\sin^2\varphi v_\varphi^2) + mg\ell\cos\varphi = \frac{1}{2}m\ell^2 v_\varphi^2 + mg\ell\cos\varphi$. The corresponding Euler-Lagrange equation is $\frac{d}{dt}(\frac{\partial L}{\partial v_\varphi}) - \frac{\partial L}{\partial\varphi} = 0$, i.e., $m\ell^2\frac{dv_\varphi}{dt} + mg\ell\sin\varphi = 0$. Therefore $\ddot\varphi = -\frac{g}{\ell}\sin\varphi$ is the equation of motion.

Homework Problem 11. Show that de Rham cohomology is a ring: Let M be a manifold, and let $\omega \in \Omega^k(M), \tau \in \Omega^l(M)$ be two closed forms. Show that $\omega\wedge\tau$ is closed and that its cohomology class depends only on the classes of ω and τ, i.e. show that the definition $[\omega]\wedge[\tau] = [\omega\wedge\tau]$ makes sense.

Homework Problem 12. Show that if a form $\omega \in \Omega^2(M)$ is exact then $\omega^n = \omega\wedge\cdots\wedge\omega$ (n times) is also exact. Conclude that if a form $\omega\in\Omega^2(M)$ is symplectic and the manifold M is compact, then the cohomology class $[\omega]$ in $H^2(M)$ is not zero. Hint: Consider $\int_M \omega^n$ where $n = \frac{1}{2}\dim M$.

Homework Problem 13. Look up the cohomology groups $H^k(S^n)$. Use it to answer the following question: for what values of n, does S^n admit a symplectic form?

11. Lecture 11. Legendre transform

We start with a brief digression. Let $L : V \to \mathbb{R}$ be a smooth function on a vector space V and let $v_1, \dots v_n$ be coordinates on V (we can choose them, for example, to be a basis of V^*). For each $v \in V$ consider the matirix $(\frac{\partial^2 L}{\partial v_i \partial v_j}(v))$. It can be interpreted as a quadratic form $d^2L(v)$ on V as follows: for $u = \sum u_i \frac{\partial}{\partial v_i}$ and $w = \sum w_i\frac{\partial}{\partial v_i}$ we define

$$d^2L(v)(u,w) = \sum_{i,j}\frac{\partial^2 L}{\partial v_i\partial v_j}(v)u_i w_j.$$

Note that the form $d^2L(v)$ has a coordinate-free definition: by chain rule for any $u, w \in V$

$$d^2L(v)(u,w) = \left.\frac{\partial^2}{\partial s\partial t}L(v+su+tw)\right|_{(0,0)}.$$

Consequently the matrix $(\frac{\partial^2 L}{\partial v_i\partial v_j}(v))$ is invertible if and only if the quadratic form $d^2L(v)$ is nondegenerate.

Recall that given a Lagrangian $L : TM \longrightarrow \mathbb{R}$ and two points $m_1, m_2 \in M$, the corresponding action $A_L : \{\ C^1$ paths connecting m_1 to $m_2\} \longrightarrow \mathbb{R}$ is defined by

$$A_L(\sigma) = \int_a^b L(\sigma(t), \dot\sigma(t))\, dt$$

Suppose that the points m_1, m_2 lie in a coordinate patch U with coordinates $x_1, \dots, x_n$. Let $x_1, \dots, x_n, v_1, \dots, v_n$ be the corresponding coordinates on $TU \subset TM$. We saw that a path $\gamma : [a, b] \to U$, $\gamma(a) = m_1$, $\gamma(b) = m_2$, is critical for the action A_L if and only if the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_i}(\gamma(t), \dot\gamma(t))\right) - \frac{\partial L}{\partial x_i}(\gamma, \dot\gamma) = 0, \quad 1 \le i \le n$$

hold. Now $\frac{d}{dt}(\frac{\partial L}{\partial v_i}(\gamma, \dot\gamma)) = \sum_j (\frac{\partial^2 L}{\partial x_j \partial v_i}\dot\gamma_j + \frac{\partial^2 L}{\partial v_j \partial v_i}\ddot\gamma_j)$, hence the Euler-Lagrange equations read

$$\sum_j \frac{\partial^2 L}{\partial v_i \partial v_j}\ddot\gamma_j = \frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i}\dot\gamma_j \quad 1 \le i \le n.$$

Now assume that L is a **regular** Lagrangian, that is to say that the matrix $(\frac{\partial^2 L}{\partial v_i \partial v_j}(x, v))$ is invertible for all $(x, v) \in TU$. Equivalently assume that for all $x \in U$ the form $d^2 L\big|_{T_x M}(v)$ is nondegenerate for all $v \in T_x M$. Then there exists an inverse metrix $(M_{ki}) = (M_{ki}(x, v))$:

$$\sum_i M_{ki}\frac{\partial^2 L}{\partial v_i \partial v_j} = \delta_{kj}.$$

Under this assumption

$$\sum_{i,j} \underbrace{M_{ki}\frac{\partial^2 L}{\partial v_i \partial v_j}}_{\delta_{kj}}\ddot\gamma_j = \sum_i M_{ki}(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i}\dot\gamma_j),$$

hence

$$\ddot\gamma_k = \sum_i M_{ki}(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i}\dot\gamma_j) \tag{13}$$

Exercise 8. Let g be a Riemannian metric on M and let $L(x, v) = \frac{1}{2}g(x)(v, v)$. Check that this Lagrangian is regular. What does (13) look like for this L?

We can rewrite (13) as first order system in $2n$ variables.

$$\begin{aligned} \dot{x_j} &= v_j \\ (\ddot x_k =)\dot v_k &= \sum_i M_{ki}(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i}v_j) \end{aligned} \tag{14}$$

Note that in physics literature the coordinates x_i's are usually called q_i's and the corresponding coordinates v_i's are usually called $\dot q_i$'s. The confusing point here is that the dot above q_i does not stand for anything; $\dot q_i$ is simply a name of a coordinate.

Equation (14) means that we have a vector field X_L on TU:

$$X_L(x, v) = \sum_j v_j \frac{\partial}{\partial x_j} + \sum_{k,i} M_{ki}(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i}v_j)\frac{\partial}{\partial v_k}$$

The vector field X_L is called the **Euler-Lagrange vector field**. One can show that

Proposition 56. *X_L is a well-defined vector field on the tangent bundle TM, i.e. it transforms correctly under the change of variables.*

Rather than proving the proposition directly we will use a different, indirect, approach which has its own merits. Given a function $L : TM \longrightarrow \mathbb{R}$ it makes sense to restrict it to a fiber T_xM and compute the differential of the restriction at a point $v \in T_xM$. Now the differential $d(L\big|_{T_xM})(v)$ is naturally an element of the dual space T_x^*M. This gives us a map $\mathcal{L} = \mathcal{L}(L) : TM \longrightarrow T^*M$, $\mathcal{L}(x, v) = d(L\big|_{T_xM})(v)$. The map is called the **Legendre transform** associated with a Lagrangian L. Note that for any $v, w \in T_xM$ we have $\langle \mathcal{L}(x, v), w\rangle = d(L\big|_{T_xM})(v), w\rangle = \frac{d}{dt}\big|_0 L(x, v + tw)$, which is often a good way to compute the Legendre transform. Although the diagram

$$\begin{tikzcd} TM \arrow[rr, "\mathcal{L}"] \arrow[dr] & & T^*M \arrow[dl] \\ & M & \end{tikzcd}$$

commutes, the transform $\mathcal{L}(L)$ need not to be a map of vector bundles since $\mathcal{L}$ restricted to each fiber need not be linear.

Example 30. Consider the manifold $M = \mathbb{R}$ and the Lagrangian $L(x, v) = e^v$ (this Lagrangian has no physical meaning). Then the Legendre transform $\mathcal{L} : T\mathbb{R} \to T^*R \simeq \mathbb{R}^2$ is given by $\mathcal{L}(x, v) = (x, e^v)$

Example 31. Let g be a Riemannian metric on on a manifold M and let $L(x, v) = \frac{1}{2}g(x)(v, v)$ ("kinetic energy"). Let us compute the Legendre transform.

$$\langle \mathcal{L}(x, v), w\rangle = \frac{d}{dt}\Big|_0 \left(\frac{1}{2}g(x)(v + tw, v + tw)\right)$$

$$= \frac{d}{dt}\Big|_0 \left(\frac{1}{2}g(x)(v, v) + tg(x)(v, w) + \frac{1}{2}t^2 g(x)(w, w)\right) = g(x)(v, w).$$

Hence $\mathcal{L}(x, v) = g(x)(v, \cdot) = g(x)^\sharp$.

Remark 57. One can define a Legendre transform for a function on an arbitrary vector bundles. Namely, if $L : E \to \mathbb{R}$ is a smooth function on a vector bundle $E \to B$ it makes sense to define the Legendre transform $\mathcal{L}(L)$ as a map from E to its dual bundle E^* (a bundle with fibers dual to the fibers of E) by

$$\mathcal{L}(x, \cdot) : E_x \to (E_x)^* =: E_x^*, \quad \mathcal{L}(x, e) = d(L\big|_{E_x})(e).$$

Let us now compute the Legendre transform in coordinates. Let $L : TM \to \mathbb{R}$ be a Lagrangian. Let $x_1, \ldots, x_n$ be coordinates on M, and $x_1, \ldots, x_n, v_1, \ldots, v_n$ be the corresponding coordiantes on TM. Then it is easy to see that

$$\mathcal{L}(x_1, \dots, x_n, v_1, \dots, v_n) = (x_1, \dots, x_n, \frac{\partial L}{\partial v_1}(x,v), \dots, \frac{\partial L}{\partial v_n}(x,v)).$$

Consequently the differential of $\mathcal{L}$ is given by

$$d(\mathcal{L}) = \begin{pmatrix} \begin{matrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{matrix} & 0 \\ & \\ \frac{\partial^2 L}{\partial x_i \partial v_j} & \frac{\partial^2 L}{\partial v_i \partial v_j} \\ & \end{pmatrix}$$

It follows that a Lagrangian L is regular if and only if the differential of the associated Legender transform $d(\mathcal{L}(L))$ is one-to-one, i.e., if the Legender transform $\mathcal{L}$ is a local diffeomorphism (we used the inverse function theorem).

Theorem 58. *Suppose $L : TM \longrightarrow \mathbb{R}$ is a Lagrangian such that for all $(x,v) \in TM$ the quadratic form $d^2(L\big|_{T_xM})(v)$ is positive definite. Then*

1. *The image $\mathcal{O}$ of the Legendre transform $\mathcal{L} = \mathcal{L}(L) : TM \longrightarrow T^*M$ is open and the map $\mathcal{L} : TM \to \mathcal{O}$ is a diffeomorphism.*
2. *There is a smooth function $H : \mathcal{O} \longrightarrow \mathbb{R}$ such that $d\mathcal{L}(X_L) = X_H$ where X_H is the Hamiltonian vector field of H and X_L is the Euler-Lagrangian vector field of L. Moreover,*
$$H(\mathcal{L}(x,v)) = \langle \mathcal{L}(x,v), v) \rangle - L(x,v).$$
3. *The inverse $\mathcal{L}(L)^{-1}$ of the Legendere transform of L is the Legendre transform $\mathcal{L}(H)$ associated with H.*

As a preparation to proving the theorem let us first consdier the Legendre transform of a function on a vector space. Let $f : V \longrightarrow \mathbb{R}$ be a smooth function on a vector space V. We define the "Legendre transform" $\mathcal{L}(f) : V \to V^*$ by $\mathcal{L}(f)(v) = df(v)$.

Assume now that for all $v \in V$ the form $d^2 f(v)$ is positive definite. Such functions f are called (strictly) **convex**.

In the simplest case $\dim V = 1$, i.e., $V = \mathbb{R}$. Then the quadratic form $d^2 f(v)$ is simply the second derivative $f''(v)$, so f is convex if and only if $f''(v) > 0$ for all v. Note that if f is convex and if $f'(x) = 0$ for some $x \in \mathbb{R}$, then x is the **unique** minimum of f. Of course such x need not exist.

Example 32. Let $f(v) = e^v$. Then $f''(v) > 0$ but f has no critical points.

Continuing with the simplest case of $f : \mathbb{R} \to \mathbb{R}$ we see that $\mathcal{L}(v) = f'(v)$ and

$$\begin{aligned} \xi = \mathcal{L}(v) &\Leftrightarrow \xi = f'(v) \\ &\Leftrightarrow \frac{d}{dv}(f(v) - \xi \cdot v) = 0 \\ &\Leftrightarrow v \text{ is a critical point of } f_\xi(v) := f(v) - \xi \cdot v \end{aligned}$$

Since ${f_\xi}''(v) = f''(v)$ we see that if f is strictly convex then the transform $\mathcal{L}$ is one-to-one.

Now let's drop the assumption on the dimension of V and prove that if $f : V \to \mathbb{R}$ is convex then the transform $\mathcal{L}(f) : V \longrightarrow V^*$ is one-to-one. Let us show first that if f is convex, then the critical points of f are unique:

Suppose v_1 and v_2 are two distinct critical points of f. Consider a function $h(t) := f(tv_1 + (1-t)v_2)$. The function h is is convex, and $t = 0, 1$ are critical points of h. This contradicts the dimension 1 case we have considered. $\square$.

Now $\xi = \mathcal{L}(v_0)$ iff $\xi = df(v_0)$ iff v_0 is a critical point of $f_\xi(v) := f(v) - \xi(v)$. If f is convex than f_ξ is also convex. Hence, since the critical points of f_ξ are unique, we conclude that the transform $\mathcal{L}(f)$ is **globally** one-to-one.

Since for a convex function the transform $\mathcal{L}$ is a local diffeomorphism, the image $\mathcal{O}$ of $\mathcal{L}(f)$ is an open subset of V^*. Combining this with the observation that the transform is globally injective, we conclude further that $\mathcal{L}(f) : V \longrightarrow \mathcal{O} \subset V^*$ is a diffeomorphism.

It follows that if $L : TM \longrightarrow \mathbb{R}$ is **fiber convex** (that is, if $d^2(L\big|_{T_xM})(v)$ is positive definite for all $(x, v) \in TM$), then the image $\mathcal{O} = \mathcal{L}(TM)$ is open and that $\mathcal{L}(L) : TM \longrightarrow \mathcal{O} \subseteq T^*M$ is a diffeomorphism. This proves part 1 of Theorem 58.

Let us now go back to the vector space case. Consider again a smooth function $f \in C^\infty(V)$ wich is convex: $d^2f(v)$ is positive definte for all $v \in V$. Let us assume for simplicity of notation that the image of $\mathcal{L}(f) : V \longrightarrow V^*$ is all of V^*.

Lemma 59. *Let $f : V \to \mathbb{R}$ be a smooth convex function on a vector space V such that $\mathcal{L}(f)(V) = V^*$. Then $\mathcal{L}(f)^{-1} = \mathcal{L}(H)$ for some $H \in C^\infty(V^*)$. In fact, $H \circ \mathcal{L}(f)(v) = \langle \mathcal{L}(f)(v), v\rangle - f(v)$. Here $\langle \cdot, \cdot \rangle : V^* \times V \longrightarrow \mathbb{R}$ denotes the canonical pairing.*

Proof. Consider $\Lambda = \mathrm{graph}(\mathcal{L}(f)) \subseteq V \times V^*$. Choose a basis of V and the dual basis of V^*. This gives us coordinates $x_1, \dots, x_n$ on V and dual coordinates $\xi_1, \dots, \xi_n$ on V^*. Recall that $\alpha_V = \sum \xi_i \, dx_i$ is the tautological one-form on $T^*V \simeq V \times V^*$ and $\alpha_{V^*} = \sum x_i \, d\xi$ is the tautological one-form on $T^*(V^*) \simeq V \times V^*$. Now $\alpha_V + \alpha_{V^*} = \sum \xi_i \, dx_i + \sum x_i \, d\xi_i = d(\sum x_i \xi_i) = d\langle \cdot, \cdot \rangle(x, \xi)$. The graph $\Lambda = \{(v, df(v)) : v \in V\}$ is the image of the map $df : V \longrightarrow T^*V$ Since $d(df) = 0$, the graph Λ is Lagrangian in $(T^*V, d\alpha_V)$ (cf. homework problem 7). Since the transform $\mathcal{L}(f)$ is a diffeomorphism, the graph Λ is the image of some map $\mu : V^* \longrightarrow T^*(V^*) \simeq V \times V^*$. Since Λ is Lagrangian for $d\alpha_V = -d\alpha_{V^*}$ we have $d\mu = d(\mu^* \alpha_{V^*}) = 0$. By Poincaré Lemma (Lemma 28) $\mu = dH$ for some function $H \in C^\infty(V^*)$. By construction, $\mathcal{L}(H) = \mathcal{L}(f)^{-1}$.

It remains to compute the functon H. Let

$$\begin{tikzcd} & V \times V^* \arrow[dl, "p_1"'] \arrow[dr, "p_2"] & \\ V & & V^* \end{tikzcd}$$

be the obvious projections. Then $df \circ p_1\big|_\Lambda = id_\Lambda$ and $dH \circ p_2\big|_\Lambda = id_\Lambda$. Therefore $\alpha_V\big|_\Lambda = (df \circ p_1\big|_\Lambda)^* \alpha_V = p_1^*((df)^* \alpha_V) = p_1^* df$. Similarly $\alpha_{V^*} = p_2^* dH$. Therefore $(p_1\big|_\Lambda)^* df + (p_2\big|_\Lambda)^* dH =$

$(\alpha_V + \alpha_{V^*})\big|_\Lambda = d\langle\cdot,\cdot\rangle\big|_\Lambda$. Hence $(p_1^* f + p_2^* H)\big|_\Lambda = \langle\cdot,\cdot\rangle\big|_\Lambda + c$, where c is some constant which we may take to be zero. We conclude that $f(v) + H(\mathcal{L}(f)(v)) = \langle \mathcal{L}(f)(v), v\rangle$. $\square$

Note that same proof works through even if the Legendre transform $\mathcal{L}(f)$ is not surjective, i.e. if $\mathcal{L}(f) : V \longrightarrow \mathcal{O} \subseteq V^*$ is a diffeomorphism for some open subset $\mathcal{O}$ of V^*.

If L is a fiber-convex Lagrangian as in the hypotheses of Theorem 58 we can apply the above argument fiber by fiber to the maps $\mathcal{L}(L)(x,\cdot) : T_xM \to T_x^*M$. We conclude that the smooth function $H : \mathcal{O} \to \mathbb{R}$ defined by $H(\mathcal{L}(L)(x,v)) = \langle \mathcal{L}(L)(x,v), v)\rangle - L(x,v)$ has the property that $\mathcal{L}(H) = \mathcal{L}(L)^{-1}$. This proves Theorem 58 except for the claim that $d\mathcal{L}(L)(X_L) = X_H$.

Remark 60. Recall that the physics notation for $(x_1,\dots,x_n,v_1,\dots,v_n)$ is $(q_1,\dots,q_n,\dot q_1,\dots,\dot q_n)$. In physics literature $p_i := \frac{\partial L}{\partial q_i}$ are callled (generalized) momenta. They are considered as coordinates on the cotangent bundle and as functions on the tangent bundle. The defintition of the Hamiltonian H then takes the form $H(q,p) = \sum p_i \dot q_i - L(q,\dot q)$.

12. Lecture 12. Legendre transform and some examples

It remians to prove $d\mathcal{L}(L)(X_L) = X_H$ where X_L is the Euler-Lagrange vector field of the Lagrangian L, $d\mathcal{L} : T(TM) \longrightarrow T(\mathcal{O})$ is the differential of the Legendre transform $\mathcal{L} = \mathcal{L}(L)$ and X_H is the Hamiltonian vector field of of the function H defined in Theorem 58. Our proof is a computation in coordinates.

Recall that the Euler-Lagrange vector field X_L is given by

$$X_L(x,v) = \sum_i v_i \frac{\partial}{\partial x_i} + \sum_k B_k \frac{\partial}{\partial v_k}$$

where $B_k := \sum_i M_{ki}(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i} v_j)$ and the matrix (M_{ki}) is defined by $\sum_i M_{ki} \frac{\partial^2 L}{\partial v_i \partial v_j} = \delta_{kj}$.

It is enough to show that $\iota(X_L)\mathcal{L}^*\omega = d(\mathcal{L}_L^* H)$. In coordinates, $\mathcal{L}(x_i, v_i) = (x_i, \frac{\partial L}{\partial v_i})$ and $\omega = \sum dx_i \wedge dp_i$. Consequently $\mathcal{L}^*\omega = \sum dx_i \wedge d(\frac{\partial L}{\partial v_i})$. Therefore

$$\begin{aligned}
d(\mathcal{L}^* H) &= d(\langle \mathcal{L}(x,v), v\rangle - L(x,v)) \\
&= \sum d(\frac{\partial L}{\partial v_i} v_i - L) \\
&= \sum \left(\frac{\partial L}{\partial v_i} dv_i + v_i d(\frac{\partial L}{\partial v_i}) - \frac{\partial L}{\partial x_i} dx_i - \frac{\partial L}{\partial v_i} dv_i \right) \\
&= \sum \left(v_i d(\frac{\partial L}{\partial v_i}) - \frac{\partial L}{\partial x_i} dx_i \right)
\end{aligned}$$

On the other hand

$$\begin{aligned}
\iota(X_L)\mathcal{L}^*\omega &= \sum_{i,k} \iota(v_k\frac{\partial}{\partial x_k} + B_k\frac{\partial}{\partial v_k})dx_i \wedge d(\frac{\partial L}{\partial v_i}) \\
&= \sum_{i,k}\left(v_k\langle\frac{\partial}{\partial x_k}, dx_i\rangle d(\frac{\partial L}{\partial v_i}) + v_k\langle\frac{\partial}{\partial x_k}, d(\frac{\partial L}{\partial v_i})\rangle(-dx_i) + 0 + B_k\langle\frac{\partial}{\partial v_k}, d(\frac{\partial L}{\partial v_i})\rangle(-dx_i)\right) \\
&= \sum_{i,k}\left(v_k\delta_{ik}d(\frac{\partial L}{\partial v_i}) - \frac{\partial^2 L}{\partial x_k\partial v_i}v_k dx_i - B_k\frac{\partial^2 L}{\partial v_k\partial v_i}dx_i\right) \\
&= \sum_i v_i d(\frac{\partial L}{\partial v_i}) - \sum_{k,i}\frac{\partial^2 L}{\partial x_k\partial v_i}v_k dx_i - \sum_i(\frac{\partial L}{\partial x_i} - \sum_j\frac{\partial^2 L}{\partial x_j\partial v_i}v_j)dx_i \\
&= \sum_i v_i d(\frac{\partial L}{\partial v_i}) - \frac{\partial L}{\partial x_i}dx_i
\end{aligned}$$

Therefore $d(\mathcal{L}^*H) = \iota(X_L)\mathcal{L}^*\omega$ and hence $d\mathcal{L}(X_L) = X_H$. This finishes the proof of Theorem 58.

Note that we have also proved that the Euler-lagrange vector field of a fiber-convex Lagrangian **is** a well-defined vector field. But the main content of the theorem is that it links the variational and Hamiltonain formulations of mechanics. Note also that it follows that the integral curves of the geodesic flow used to prove the tubular neighborhood theorem do project down to geodesics (cf. Lecture 7).

Let us now look at some examples. Recall that if we have N particles in $\mathbb{R}^3$ with masses m_k, $1 \leq k \leq N$, subject to conservative forces encoded in a potential V, then the Lagrangian of the system is given by $L(\vec{x}_1, \ldots, \vec{x}_N, \vec{v}_1, \ldots, \vec{v}_N) = \frac{1}{2}m_k\|\vec{v_k}\|^2 - V(\vec{x}_1, \ldots, \vec{x}_N)$. If we relable the particles and identify $(\mathbb{R}^3)^N$ with $\mathbb{R}^n$, $n = 3N$, then the Lagrangian has the form

$$L(x, v) = \frac{1}{2}\sum_{i,j} g_{ij}v_iv_j - V(x)$$

where (g_{ij}) is a fixed symmetric positive definite matrix. If a submanifold $M \hookrightarrow \mathbb{R}^n$ is a constraint for our system, then d'Alembert's principle tells us that the Legrangian L_M for the constrained system is $L_M = i^*L$, that is

$$L_M(x, v) = \frac{1}{2}(i^*g)(x)(v, v) - V(x)$$

where i^*g is a Riemannian metric induced by the embedding i of M into $\mathbb{R}^n$. This motivates

Definition 61. A classical mechanical system on a manifold M is a Lagrangian $L : TM \longrightarrow \mathbb{R}$ of the form $L(x, v) = \frac{1}{2}g(x)(v, v) - V(x)$, where g a Riemannian metric on M and V is a function on M.

Given a classical mechanical system with a Lagrangian as above what is the corresponing Hamiltonian on T^*M?

It is easy to see that the Legendre transform $\mathcal{L} : TM \longrightarrow T^*M$ is given by $\mathcal{L}(x, v) = g(x)^\sharp(v)$ (cf. Example 31). Consequently $H(\mathcal{L}(x, v)) = H(g^\sharp(x)v) = \langle g^\sharp(x)v, v\rangle - L(x, v) = g(x)(v, v) - \frac{1}{2}g(x)(v, v) + V(x) = \frac{1}{2}g(x)(v, v) + V(x)$. Therefore $H(x, p) = \frac{1}{2}g(x)((g^\sharp(x))^{-1}p, (g^\sharp(x))^{-1}p) + V(x) = \frac{1}{2}g^*(x)(p, p) + V(x)$ where g^* is the dual metric on T^*M.

Recall that if if $g^{ij}(x) := g^*(x)(dx_i, dx_j)$, then $g^{ij}g_{jk} = \delta^i_k$. We can now compute some concrete examples.

Example 33 (Planar pendulum)**.** The unconstrained Lagrangian of a heavy particle in $\mathbb{R}^2$ is $L(x, v) = \frac{m}{2}(v_1^2 + v_2^2) - mgx_2 = \frac{m}{2}(dx_1^2 + dx_2^2)(v, v) - mgx_2$, where g is the gravitational constant and m is the mass of the particle. Our constraint is a circle $M = S^1$ of radius ℓ, and the embedding $i : S^1 \longrightarrow \mathbb{R}^2$ can be chosen to be $i : \varphi \mapsto (\ell \sin\varphi, -\ell\cos\varphi)$. Consequently the constrained Lagrangian is given by

$$\begin{aligned}
L(\varphi, v_\varphi) &= \frac{m}{2}(i^*g)(v_\varphi\frac{\partial}{\partial\varphi}, v_\varphi\frac{\partial}{\partial\varphi}) - i^*V(x)\\
&= \frac{m}{2}(d(\ell\sin\varphi)^2 + d(-\ell\cos\varphi)^2)(v_\varphi\frac{\partial}{\partial\varphi}, v_\varphi\frac{\partial}{\partial\varphi}) + \ell(\cos\varphi)mg\\
&= \frac{m}{2}l^2(\cos^2\varphi d\varphi^2 + \sin^2\varphi d\varphi^2)(v_\varphi\frac{\partial}{\partial\varphi}, v_\varphi\frac{\partial}{\partial\varphi}) + \ell(\cos\varphi)mg\\
&= \frac{m\ell^2}{2}v_\varphi^2 + mg\ell\cos\varphi
\end{aligned}$$

It follows that in coordinates $(\varphi, p_\varphi) \in S^1 \times \mathbb{R}$ the Hamiltonian is given by $H(\varphi, p_\varphi) = \frac{1}{2}\frac{1}{ml^2}p_\varphi^2 - mg\ell\cos\varphi$. The symplectic form, of course, is the standard form $\omega = d\varphi \wedge dp_\varphi$. Therefore the Hamiltonian vector field X_H is $\frac{1}{ml^2}p_\varphi\frac{\partial}{\partial\varphi} - mg\ell\sin\varphi\frac{\partial}{\partial p_\varphi}$. Thus the equations of motion are

$$\begin{aligned}
\dot{\varphi} &= \frac{1}{m\ell^2}p_\varphi\\
\dot{p_\varphi} &= -mg\ell\sin\varphi
\end{aligned}$$

We now start exploiting the fact that our equations of motion are Hamiltonian.

Definition 62. Let (M, ω) be a symplectic manifold. A function $f \in C^\infty(M)$ is a **conserved quantity** (a.k.a. a first integral) for a Hamiltonian $H \in C^\infty(M)$ iff f is constant on the the integral curves $\gamma(t)$ of the Hamiltonian vector field X_H of H: $f(\gamma(t)) =$ constant.

The following theorem is trivial but has profound consequences.

Theorem 63. *Let (M, ω) be a symplectic manifold and H a Hamiltonian on M. The function H is a conservd quantity for H. In particular the integral curves of the Hamiltonian vector field of H lie entirely in the level hypersurfaces $H = c$.*

Proof. Let γ be an integral curve of the Hamiltonian vector field of H. Then $H(\gamma(t)) =$ const if and only if $0 = \frac{d}{dt}H(\gamma(t))$. Now $\frac{d}{dt}H(\gamma(t)) = X_H(H) = \iota(X_H)dH = \iota(X_H)(\iota(X_H)\omega) = \omega(X_H, X_H) = 0$ since ω is a skew-symmetric tensor. $\square$

Example 34. For planar pendulum the level sets of H are one-dimensional. Hence the connected components of the regular level sets of H are of integral curves of X_H. Therefore if $(\varphi(t), p_\varphi(t))$ is an integral curve, we have $H(\varphi, p_\varphi) = \frac{1}{2}\frac{1}{m\ell^2}p_\varphi^2 - mg\ell\cos\varphi \equiv E$ for some constant E. Hence $p_\varphi = \pm\sqrt{2m\ell^2}(E + mg\ell\cos\varphi)^{1/2}$

Example 35 (Spherical pendulum)**.** For a spherical pendulum the unconstrained lagrangian is $L(x, v) = \frac{1}{2}m(v_1^2 + v_2^2 + v_3^2) - mgx_3 = \frac{m}{2}(dx_1^2 + dx_2^2 + dx_3^2)(v_i\frac{\partial}{\partial x_i}, v_i\frac{\partial}{\partial x_i}) - mgx_3$. Our constraint is a round sphere $S^2 \subset \mathbb{R}^3$ of radius ℓ. Hence the constrained Lagrangian is given by $L(q, \dot{q}) = \frac{m}{2}g(x)(\dot{q}, \dot{q}) - mgx_3$ where g is the round metric on S^2. The corresponding Hamiltonian is $H(q, p) = \frac{1}{2m}g^*(x)(p, p) + mgx_3$ where g^* is the metric dual to the round metric. We know that H is a conserved quantity. We will see in the next few lectures that there is another conserved quantity — angular momentum about the x_3-axis. The reason for the existence of the second conserved quantity is the rotational symmetry of our system.

Homework Problem 14. Let (M, g) be a Riemannian manifold, $N \subset M$ a submanifold. Then $TN \subset TM$ since N is a submanifold, it inherits a metric g_N from M. Thus we hae two isomorphisms $g^\sharp : TM \longrightarrow T^*M$ and $g_N^* : TN \longrightarrow T^*N$. Show that the composition

$$\varphi : T^*N \xrightarrow{(g_N^\sharp)^{-1}} TN \hookrightarrow TM \xrightarrow{g^\sharp} T^*M$$

is a symplectic embedding, i.e. $\varphi^*(\omega_{T^*M}) = \omega_{T^*N}$.

13. Lecture 13. Constants of motion. Lie and Poisson algebras

Recall that if (M, ω) is a symplectic manifold and $f \in C^\infty(M)$ is a smooth function on M, then the Hamiltonian vector field X_f of f is defined by $\iota(X_f)\omega = df$, that is, X_f is a section of the tangent bundle making the diagram

$$\begin{tikzcd}
TM \arrow[rr, "\omega^\sharp", "\simeq"'] & & T^*M \\
& M \arrow[ul, "X_f"] \arrow[ur, "df"'] &
\end{tikzcd}$$

commute.

Definition 64. Let (M, ω) be a symplectic manifold and let $f \in C^\infty(M)$ is a smooth function on M. A function h is a **constant of motion** for f, (equivalently h is **preserved** by f, h is a **first integral** of f) iff $X_f(h) = 0$.

Lemma 65. *Let (M, ω) be a symplectic manifold, $f, h \in C^\infty(M)$ be two smooth functions. Then f is preserved by h if and only if h is preserved by f.*

Proof.

$$\begin{aligned} X_h(f) &= \langle X_h, df\rangle \\ &= \langle X_h, \iota(X_f)\omega\rangle \\ &= \omega(X_f, X_h) \\ &= -\omega(X_h, X_f) \\ &= -\langle X_f, \iota(X_h)\omega\rangle \\ &= -\langle X_f, dh\rangle \\ &= -X_f(h) \end{aligned}$$

Therefore $X_h(f) = 0$ if and only if $X_f(h) = 0$. □

Definition 66. The function $\{f, h\} := X_f(h) = \langle X_f, dh\rangle = -\langle X_h, df\rangle$ is called the **Poisson bracket** of f and h defined by the symplectic form ω.

Thus Lemma 65 asserts that the functions f and h are constants of motion for each other if and only if their Poisson bracket $\{f, h\}$ is zero..

Example 36. Consider a particle in $\mathbb{R}^3$ in a central force field. That is, consider the manifold $M = T^*\mathbb{R}^3$ with coordinates $(q_1, q_2, q_3, p_1, p_2, p_3)$ and with the canonical the symplectic form $\omega = \sum dq_i \wedge dp_i$. A central force Hamiltonian is a function of the form $h(q,p) = \frac{1}{2m}\sum p_i^2 + V(||q||^2)$. Consider the vector field $Z = q_1\frac{\partial}{\partial q_2} - q_2\frac{\partial}{\partial q_1} + p_1\frac{\partial}{\partial p_2} - p_2\frac{\partial}{\partial p_1}$. We then have

$$Z(h) = \frac{1}{m}(p_1p_2 - p_2p_1) + V'(||q||^2)2(q_1q_2 - q_2q_1) = 0.$$

Moreover, the vector field Z is Hamiltonian: $\iota(Z)\omega = \iota(q_1\frac{\partial}{\partial q_2} - q_2\frac{\partial}{\partial q_1} + p_1\frac{\partial}{\partial p_2} - p_2\frac{\partial}{\partial p_1})\sum dq_i \wedge dp_i = q_1dp_2 - q_2dp_1 - p_1dq_2 + p_2dq_1 = d(q_1p_2) - d(q_2p_1) = d(q_1p_2 - q_2p_1)$. So $Z = X_{j_3}$ where $j_3(q,p) := q_1p_2 - q_2p_1$. Consequently the function j_3 is conserved.

This example is meant to provoke a number of questions:

1. Where did this vector field Z come from and why is it Hamiltonian?
2. Are there any other conserved quantities?

It is easy to check that $j_2(q,p) = q_1p_3 - q_3p_1$ is also conserved and that the Poisson bracket $j_1 := \{j_2, j_3\}$ of j_3 and j_2 is conserved as well.

One can also notice that the Hamiltonian iss rotationally symmetric. We will see that it is the rotational symmetry of the problem which is responsible for the presence of the three conserved quantities.

To make sense of all this, we first introduce the notion of Lie algebras. We will then follow by introducing Lie groups, group actions, and Hamiltonian group actions.

13.1. **Lie algebras.**

Definition 67. A (real) **Lie algebra** A is a (real) vector space (finite or infinite dimensional) together with a bilinear map $[\cdot,\cdot] : A \times A \longrightarrow A$, called the **Lie bracket**, such that for all $X, Y, Z \in A$

1. $[X, Y] = -[Y, X]$ anti-symmetry
2. $[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]]$ Jacobi identity

Example 37. Let $A = \mathbb{R}^3$. Define the bracket $[\cdot, \cdot] : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ by $[x, y] := x \times y$, the cross-product of x and y. It is not hard to check that A is a Lie algebra.

Example 38. Let V be any real vector space. Define $[x, y] = 0$ for all $x, y \in V$. Then $(V, [\cdot, \cdot])$ is a Lie algebra, called an **abelian** Lie algebra.

Example 39. Let A be any associative algebra over the reals (for example A can be the algebra of $n \times n$ matrices with multiplication being the matrix multiplication). Then A can be made into a Lie algebra by defining

$$[a, b] := ab - ba$$

(Check that A with this bracket is indeed a Lie algebra).

Example 40. Let M be a manifold and let A be the vector space of smooth vector fields on M. Take the bracket on A to be the Lie bracket of vector fields: for any vector fields X and Y and for any smooth function f let $[X, Y](f) = X(Y(f)) - Y(X(f))$. With this bracket A forms a Lie algebra.

Definition 68. Let (M, ω) be a symplectic manifold. A vector field X on M is called **symplectic** if the Lie derivative of the symplectic form ω with respect to X is zero: $L_X \omega = 0$.

Remark 69. Let ϕ_t be the flow of a vector field X on a symplectic manifold (M, ω). The vector field X is a symplectic if and only if $\phi_t^* \omega = \omega$ for all time t.

Note also that if a vector field X on a symplectic manifold (M, ω) is symplectic then

$$0 = L_X \omega = d\iota(X)\omega + \iota(X) d\omega = d\iota(X)\omega + \iota(X)0 = d\iota(X)\omega$$

since ω is closed. Thus a vector field X is symplectic if and only if the form $\iota(X)\omega$ is closed. Since for any Hamiltonian vector field X_f we have, by definition, $\iota(X_f)\omega = df$, we conclude that

Proposition 70. *Any Hamiltonian vector field is symplectic.*

The converse is not true: not every symplectic vector field is Hamiltonian. Indeed consider the torus $S^1 \times S^1$ with "coordinates" θ_1, θ_2. The form $d\theta_1 \wedge d\theta_2$ is a globally defined symplectic form. The vector field $\frac{\partial}{\partial \theta_1}$ is symplectic but not Hamiltonian since $d\theta_2$ is not an exact form on the torus.

We assert that symplectic vector fields on a symplectic manifold form a Lie algebra. In fact, more is true.

Lemma 71. *The Lie bracket of two symplectic vector fields is Hamiltonian.*

Proof. Let Z and Y be two symplectic vector fields on a symplectic manifold (M, ω). On the one hand

$$L_Z(\iota(Y)\omega) = \iota(L_Z Y)\omega + \iota(Y)(L_Z \omega) = \iota([Z, Y])\omega.$$

On the other hand,

$$L_Z(\iota(Y)\omega) = d(\iota(Z)(\iota(Y)\omega)) + \iota(Z)d(\iota(Y)\omega) = d(\iota(Z)(\iota(Y)\omega)) + 0 = d(\omega(Y,Z))$$

Therefore,

$$\iota([Z,Y])\omega = d\omega(Y,Z), \tag{15}$$

that is, the Lie bracket of Z and Y is the Hamiltonian vector field of the function $\omega(Y,Z)$. ☐

Definition 72. Let L be a Lie algebra. A vector subspace A is a **Lie subalgebra** if for any $X, Y \in A$ the Lie bracket $[X,Y]$ is also in A.

A linear map ψ from a Lie algebra A_1 to a Lie algebra A_2 is a **Lie algebra morphism (a map of Lie algebras)** if for any $x, y \in A_1$ we have

$$\psi([x,y]) = [\psi(x), \psi(y)].$$

A Lie subalgebra I of an algebra L is an **ideal** if for any $X \in I$ and any $Y \in L$ the Lie bracket $[X,Y]$ is in I.

Thus the Lemma 71 asserts that the symplectic vector fields on a symplectic manifold form a Lie algebra and that Hamiltonian vector fields form an ideal inside the Lie algebra of symplectic vector fields.

Remark 73. Let X_f and X_g denote, as usual, the Hamiltonian vector fields of the functions $f, g \in C^\infty(M)$ on a symplectic manifold (M,ω) respectively. Then it follows from equation (15) that

$$\iota([X_g, X_f])\omega = d\omega(X_f, X_g) = d(\langle \iota(X_f)\omega, X_g\rangle) = d(\langle df, X_g\rangle) = d(X_g(f)) = d(\{g,f\}).$$

Therefore

$$[X_g, X_f] = X_{\{g,f\}}. \tag{16}$$

This strongly suggests that on a symplectic manifold (M,ω) the Poisson bracket on the space of smooth functions $C^\infty(M)$ makes it into a Lie algebra, and that map from $C^\infty(M)$ to the Lie algebra of the Hamiltonian vector fields $\mathrm{Ham}(M,\omega)$ given by $f \mapsto X_f$ is a map of the Lie algebras.

Lemma 74. *Let (M,ω) be a symplectic manifold. The algebra of smooth functions forms a Lie algebra with the bracket given by the Poisson bracket.*

Proof. It is easy to see that the Poisson bracket $\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ is skew-symmetric and bilinear. The interesting part is the Jacobi identity. Let $f, g, h \in C^\infty(M)$ be any three functions and let X_f, X_g and X_h denote the corresponding Hamiltonian vector fields. Now

$$\begin{aligned}
\{\{f,g\},h\} &= X_{\{f,g\}}(h) \quad \text{by definition of the Poisson bracket}\\
&= [X_f, X_g](h) \quad \text{by equation (16)}\\
&= X_f(X_g(h)) - X_g(X_f(h))\\
&= \{f,\{g,h\}\} - \{g,\{f,h\}\}.
\end{aligned}$$

This implies that $\{f,\{g,h\}\} = \{\{f,g\},h\}+\{g,\{f,h\}\}$, which proves that the Poisson bracket on a symplectic manifold satisfies the Jacobi identity . Consequently $(C^\infty(M), \{\cdot,\cdot\})$ is a Lie algebra. □

Corollary 75. *Let (M,ω) be a symplectic manifold. The map from the Poisson algebra of smooth functions $C^\infty(M)$ to the Lie algebra of symplectic vector fields $\Xi(M,\omega)$ given by $f \mapsto X_f$ is a Lie algebra map.*

Corollary 76. *Suppose two smooth functions f, g on a symplectic manifold (M,ω) are constants of motion for a function h. Then the Poisson bracket $\{f,g\}$ is also a constant of motion for h.*

Proof. If $\{h,g\} = \{h,f\} = 0$ then by the Jacobi identity $\{h,\{f,g\}\} = \{\{h,f\},g\}+\{f,\{h,g\}\} = \{0,g\} + \{f,0\} = 0 + 0 = 0$. □

Remark 77. We note that the Poisson bracket on a symplectic manifold defined by the symplectic form has one more property: for any functions f, g and h

$$\{f,gh\} = X_f(gh) = (X_f g)\,h + g\,X_f(h) = \{f,g\}h + g\{f,h\}$$

i.e. $\{\cdot,\cdot\}$ is a **bi-derivation** with respect to ordinary multiplication of functions.

The observation motivates the following abstract definition.

Definition 78. Let A be any commutative algebra over the reals. A bilinear map $\{\cdot,\cdot\} : A \times A \to A$ is a **Poisson bracket** iff

1. $\{\cdot,\cdot\}$ is skew-symmetric: $\{a,b\} = -\{b,a\}$ for any $a, b \in A$,
2. $\{\cdot,\cdot\}$ satisfies the Jacobi identity: $\{a,\{b,c\}\} = \{\{a,b\},c\} + \{b,\{a,c\}\}$ for any a, b and c in A.
3. $\{\cdot,\cdot\}$ is a bi-derivation: $\{a,bc\} = \{a,b\}c + b\{a,c\}$ for any a, b and c in A.

A commutative algebra A together with a Poisson bracket is called a **Poisson algebra**.

Thus the algebra of functions on a symplectic manifold (M,ω) with the bracket defined by $\{f,g\} = X_f(g)$ is a Poisson algebra. Other Poisson algebras arise as (sub)algebras of the algebras of smooth functions on a manifold. The Poisson bracket, however, need not come from a symplectic form (see Homework Problem 16).

Homework Problem 15. Let (M,ω) be a symplectic manifold and let $\{,\}$ be the corresponding Poisson bracket. Show that in coordinates

1. $\{f,g\} = \sum \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}\{x_i,x_j\}$
2. Let $P_{ij} = \{x_i,x_j\}$, let $\omega_{ij} = \omega(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$. Show that $\sum_j P_{ij}\omega_{jk} = \delta_{ik}$
3. What condition should a collection of functions P_{ij} satisfy for $(f,g) := \sum \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial x_j}P_{ij}$ to be a Lie algebra bracket?

Homework Problem 16. Let $\mathfrak{g}$ be a finite dimensional Lie algebra with bracket $[\cdot,\cdot]$, and let $\mathfrak{g}^*$ denote vector space dual, i.e. $\mathfrak{g}^* = \{\ell : \mathfrak{g} \longrightarrow \mathbb{R} \,|\, \ell \text{ linear }\}$. Then we have a pairing

$\mathfrak{g}^* \times \mathfrak{g} \longrightarrow \mathbb{R}, (\ell, X) \mapsto \langle \ell, X \rangle \equiv \ell(X)$. Now let $f, h \in C^\infty(\mathfrak{g}^*)$ be two functions. We define a bracket $\{f, h\}(\ell)$ as follows: the covectors $df(\ell) \in T^*_\ell \mathfrak{g}^* \simeq (\mathfrak{g}^*)^* \simeq \mathfrak{g}$ and similarly $dh(\ell) \in T^*_\ell \mathfrak{g}^* \simeq \mathfrak{g}$. Hence $[df_\ell, dh_\ell]$ makes sense as an element of $\mathfrak{g}$. Define $\{f, h\}(\ell) = \langle \ell, [df_\ell, dh_\ell] \rangle$. Clearly $\{f, h\}(\ell) = -\{h, f\}(\ell)$. Show that $\{\cdot, \cdot\}$ is in fact a Poisson bracket, i.e. that

1. $\{f, gh\} = \{f, g\}h + g\{f, h\}$ and
2. $\{f, \{g, h\}\} = \{\{f, g\}, h\} + \{g, \{f, h\}\}$

for all $f, g, h \in C^\infty(\mathfrak{g}^*)$.
Hint: Prove (1) first.

The following observation may be useful for proving (2). Let $x_1, \dots, x_n$ be a basis of $\mathfrak{g}$. They are coordinate functions on $\mathfrak{g}^*$. Since $\mathfrak{g}$ is a Lie algebra, $[x_i, x_j] = \sum_k C^k_{ij} x_k$ for some constants C^k_{ij}. These constants are not arbitrary; the Jacobi identity $[x_i, [x_j, x_k]] = [[x_i, x_j], x_k] + [x_j, [x_i, x_k]]$ gives relations. Show that $\{x_i, x_j\} = \sum C^k_{ij} x_k$. Use (1) to show that

$$\{f, g\} = \sum \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_j} \{x_i, x_j\}. \tag{17}$$

14. Lecture 14. Lie groups: a crash course

The material in this section is sketchy. Most proofs are only outlined. We refer the reader to [Warner] or [Spivak] for details.

Definition 79. A Lie group G is a group and a manifold such that

1. the group multiplication map $\mu : G \times G \longrightarrow G$, $(a, b) \mapsto ab$ is C^∞.
2. the inverse map $\mathrm{inv} : G \longrightarrow G$, $a \mapsto a^{-1}$ is C^∞.

Remark 80. If G is a Lie group then for all $a \in G$ the maps $R_a : G \to G$, $g \mapsto ga$ and $L_a : G \to G$, $g \mapsto ag$ (the right and left multiplication by a) are diffeomorphisms.

Example 41. Any finite dimensional vector space is a Lie group under vector addition.

Example 42. The **general linear group** $\mathrm{GL}(n, \mathbb{R})$ of all $n \times n$ nonsingular real matrices is a Lie group under matrix multiplication. Note that $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$ defined by the equation $det A \neq 0$.

More generally, if V is a finite dimensional vector space over $\mathbf{R}$ the group $GL(V)$ of all invertible linear maps is a Lie group. Similarly, if V is a vector space over $\mathbf{C}$ the group of invertible complex-linear maps is also a Lie group.

Example 43. The circle $S^1 = \{e^{i\theta} : \theta \in \mathbf{R}\}$, the group of complex numbers of norm 1, is a Lie group under complex multiplication.

To each Lie group G there corresponds a Lie algebra $\mathfrak{g}$. The correspondence is defined as follows.

Definition 81. Let ξ be a vector field on a Lie group G. χ is *left-invariant* if for any $a \in G$

$$dL_a(\xi) = \xi. \tag{18}$$

where L_a as above denotes the left multiplication by $a \in G$.

It is not very difficult to show that any left-invariant vector field is automatically smooth (see [Warner, p 85]). If ξ and η are two left-invariant vector field then $[\xi, \eta]$ is also left-invariant since $dL_a([\xi, \eta]) = [dL_a(\xi), dL_a(\eta)]$. Therefore the left-invariant vector fields form a Lie subalgebra of the Lie algebra of all vector fields on our Lie group. We shall denote this Lie algebra by $\mathfrak{g}$. Equation (18) implies that a left-invariant vector field ξ is determined by its value at the identity, $(dL_a)(e)(\xi(e)) = \xi(a)$. Conversely, given a vector ξ_e in T_eG the vector field ξ defined by $\xi(a) = (dL_a)(e)(\xi_e)$ is left-invariant. Thus the map sending a left-invariant vector field to its value at the identity is bijective. From now on we identify $\mathfrak{g}$ with the tangent space at the identity.

Definition 82. A Lie group H is a **Lie subgroup** of G if there is a map $i : H \to G$ such that :

1. i is a group homomorphism and
2. i is a one-to-one immersion.

Remark 83. A Lie subgroup could be a dense in the group: consider for example a line in $\mathbb{R}^2$ with irrational slope. The image of the line is dense in the two torus $S^1 \times S^1 = \mathbb{R}^2/\mathbb{Z}^2$.

Theorem 84. *There is a bijective correspondence between connected Lie subgroups of a Lie group G and Lie subalgebras of its Lie algebra $\mathfrak{g}$.*

Sketch of proof. Let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{g}$ and $\xi_1, \ldots, \xi_h$ be the basis of $\mathfrak{h}$. The elements ξ_i's are left-invariant vector fields on G. For every $a \in G$ the vectors $\xi_1(a), \ldots, \xi_h(a)$ are a basis for an h-dimensional subspace of the tangent space T_aG thus giving rise to an h-dimensional distribution on G. This distribution is involutive. Indeed, since $\mathfrak{h}$ is a Lie subalgebra

$$[\xi_i, \xi_j] = \sum_k c_{ij}^k \xi_k,$$

where $c_{ij}^k \in \mathbb{R}$. By Frobenius integrability theorem for every $a \in G$ there is a maximal connected integral submanifold N_a passing through a. Set $H = N_e$, the integral submanifold passing through the identity element of G. Let b be a point in H. Since the distribution is invariant under left translations, $L_{b^{-1}}(H)$ is also an integral submanifold passing through e. Hence for any $c \in H$ $b^{-1}c \in H$. It follows that H is an abstract subgroup of G. In fact H is a Lie (sub)group (cf [Warner, p 94]). $\square$

Exercise 9. Consider the general linear group $\mathrm{GL}(n, \mathbb{R})$. Since the group is an open subset of $\mathbb{R}^{n^2}$, its Lie algebra $\mathfrak{gl}(n, \mathbb{R})$ is $\mathbb{R}^{n^2}$. Show that the bracket $[\cdot, \cdot]$ on $\mathfrak{gl}(n, \mathbb{R})$ is given by $[A, B] = AB - BA$. Hint: If $g \in \mathrm{GL}(n, \mathbb{R})$ and $A \in T_{id}\,\mathrm{GL}(n, \mathbb{R})$ show that $dL_g(A) = gA$ (matrix multiplication).

14.1. **Homomorphisms.**

Definition 85. Let G and H be two Lie groups. A map $\rho : G \to H$ a **Lie group morphism** iff ρ is a group homomorphism and is smooth.

Example 44. The determinant map $det : GL(n, \mathbb{R}) \to \mathbb{R}^\times$ is a Lie group morphism ($\mathbb{R}^\times$ is a Lie group under ordinary multiplication). The determinant map is smooth because it is polynomial.

Given a Lie group homomorphism $\rho : G \to H$, we have an associated linear map $\delta\rho : \mathfrak{g} \to \mathfrak{h}$ defined by $\delta\rho(\xi) = d\rho_e(\xi)$ (recall that we have identified $\mathfrak{g}$ with T_eG, the tangent space at the identity, and $\mathfrak{h}$ with T_eH).

Exercise 10. Check that $\delta\rho$ is a morphism of Lie algebras.

If G is simply connected we also have the converse:

Theorem 86. *Let G and H be Lie groups. Suppose that G is simply connected. Given a Lie algebra homomorphism $\tau : \mathfrak{g} \to \mathfrak{h}$ there exists a unique Lie group homomorphism $\rho : G \to H$ such that $\delta\rho = \tau$.*

Proof. The product $G \times H$ is a Lie group with Lie algebra $\mathfrak{g} \times \mathfrak{h}$. Let $\mathfrak{k}$ be the graph of τ, $\mathfrak{k} = \{(\xi, \tau(\xi)) : \xi \in \mathfrak{g}\}$. Since τ is a Lie algebra map, $\mathfrak{k}$ is a Lie subalgebra of $\mathfrak{g} \times \mathfrak{h}$. The the connected subgroup K of $G \times H$ corresponding to $\mathfrak{k}$ should be the graph of the homomorphism that we are trying to construct.

Let us prove this. Let $\pi_1 : G \times H \to G$ be the projection on the first factor. Since π_1 is a Lie group homomorphism, $\phi = \pi_1|_K : K \to G$ is a Lie group homomorphism. Moreover, $d\phi(e) : T_eK \to T_eG$ is surjective. It follows that $d\phi(k)$ is surjective for any $k \in K$. Indeed, for any $a \in K$,

$$\phi \circ L_k\,(a) = \phi(ka) = \phi(k)\phi(a) = L_{\phi(k)} \circ \phi\,(a),$$

so by the chain rule

$$d\phi(k) \circ (dL_k)(e) = (dL_{\phi(k)})(e) \circ d\phi(e).$$

Since $L_{\phi(k)}$ and L_k are diffeomorphisms, $d\phi(k)$ is surjective if and only if $d\phi(e)$ is surjective. Since $\dim K = \dim G$, the map ϕ is a local diffeomorphism. In particular $\phi(K)$ is open in G. Since G is connected, $\phi(K) = G$. It follows that $\phi : K \to G$ is a covering map. But G is simply connected! So ϕ is a diffeomorphism. The map $\rho : \pi_2 \circ \phi^{-1} : G \to H$, where $\pi_2 : G \times H \to H$ is the projection on the second factor, is the desired Lie group homomorphism from G to H.

The uniqueness of ρ follows from the uniqueness of the subgroup corresponding to a given Lie subalgebra. $\square$

14.2. **The exponential map.**

Example 45. Consider the map $\exp : \mathfrak{gl}(n, \mathbb{R}) \longrightarrow \mathrm{GL}(n, \mathbb{R})$ defined by $A \mapsto e^A := \sum_{n=0}^{\infty} \frac{1}{n!} A^n$ which converges for all $A \in \mathfrak{gl}(n, \mathbb{R})$. One can check that $e^{tA} \cdot e^{sA} = e^{(t+s)A}$. So for a given vector $A \in \mathfrak{gl}(n, \mathbb{R})$, the image $\{e^{tA} \,|\, t \in \mathbb{R}\}$ of exp is a one-dimensional subgroup of $\mathrm{GL}(n, \mathbb{R})$. Moreover the curve $\gamma_A(t) = e^{tA}$ satisfies $\frac{d}{dt}\big|_t \gamma_A(t) = e^{tA}A = (dL_{(e^{tA})})(A)$. Thus $\gamma_A(t)$ is an integral curve of the left invariant vector field corresponding to A.

The example motivates the following construction. Let G be a Lie group and $\mathfrak{g}$ denote its Lie algebra. Fix a nonzero vector $X \in \mathfrak{g}$. Then the map

$$\mathbb{R} \to \mathfrak{g}, \quad t \mapsto tX$$

is a Lie algebra homomorphism (the bracket on $\mathbb{R}$ is necessarily zero). Since $\mathbb{R}$ is simply connected there exists by Theorem 86, a unique group homomorphism $\gamma_X : \mathbb{R} \to G$ with $\frac{d}{dt}\big|_{t=0} \gamma_X(t) = X$. Since γ_X is a homomorphism, $\gamma_X(a+t) = \gamma_X(a)\gamma_X(t)$. Hence

$$\frac{d}{dt}\bigg|_{t=a} \gamma_X(t) = \frac{d}{dt}\bigg|_{t=0} \gamma_X(t+a) = \frac{d}{dt}\bigg|_{t=0} \gamma_X(a)\gamma_X(t) = dL_{\gamma_X(a)}(X).$$

We conclude that γ_X is an integral curve of the left invariant vector field determined by $X \in \mathfrak{g}$. Since the solutions of a differential equation depend smoothly on the parameters, it follows that the map

$$\mathbb{R} \times \mathfrak{g} \to G, \quad (t, X) \mapsto \gamma_X(t)$$

is smooth.

We now define the **exponential map** $\exp : \mathfrak{g} \to G$ by

$$\exp(X) = \gamma_X(1).$$

By the above discussion it is a smooth map.

Since $\frac{d}{dt}\big|_{t=0} \gamma_X(ct) = cX$ it follows that $\exp(cX) = \gamma_X(c)$. From this one can easily deduce that $d\exp(0) : T_0\mathfrak{g} = \mathfrak{g} \to T_eG = \mathfrak{g}$ is the identity map. Hence exp is a local diffeomorphism near 0 in $\mathfrak{g}$.

Exercise 11. Show that for the general linear group the two definitions of the exponential map agree.

We note that the exponential map has a universal property: if G and H are two Lie groups, $f : G \to H$ is a morphism of Lie groups, and $\delta f : \mathfrak{g} \to \mathfrak{h}$ the corresponding morphism of Lie algebras, the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\delta f"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "f"] & H \end{tikzcd}$$

commutes. It follows that if G is any Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$ then the exponential map $\exp : \mathfrak{g} \to G$ is also given by the formula

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

The exponential map lies at the heart of the following very powerful theorem.

Theorem 87 (closed subgroup)**.** *Let H be an abstract subgroup of a Lie group G, which is also a closed subset of G. Then H is a Lie group, its Lie algebra $\mathfrak{h}$ is the set*

$$\{X \in \mathfrak{g} \mid \exp(tX) \in H \quad \textit{for all} \quad t \in \mathbb{R}\},$$

and H is a Lie subgroup of G.

We will not prove this theorem. The interested reader can find a proof in [Warner] or [Spivak].

Example 46 (Symplectic group)**.** Let (V, ω) be a symplectic vector space. The subset $\mathrm{Sp}(V, \omega)$ of $\mathrm{GL}(V)$ consisting of all elements A with $A^*\omega = \omega$ is closed in $\mathrm{GL}(V)$ (why?) and forms a subgroup. By the closed subgroup theorem it is a Lie subgroup called the **symplectic group**, and its Lie algebra $\mathfrak{sp}(V, \omega)$ consists of all linear maps $X : V \to V$ such that

$$\exp(tX)^*\omega = \omega \quad \text{ for all } \ t \in R.$$

Differentiating the above equation with respect to t we get

$$\mathfrak{sp}(V, \omega) = \{X \in \mathfrak{gl}(V) \mid \omega(Xv, w) + \omega(v, Xw) = 0 \quad \text{for all}\, v, w \in V\}.$$

Example 47 (Orthogonal group)**.** Let V be a vector space and let g be a positive definite bilinear form on V. The subset $\mathrm{O}(V, g)$ of $\mathrm{GL}(V)$ consisting of all elements A with $A^*g = g$ is closed in $\mathrm{GL}(V)$ and forms a subgroup called the **orthogonal group**. By the closed subgroup theorem it is a Lie group, and its Lie algebra $\mathfrak{o}$ consists of all linear maps $X : V \to V$ such that

$$\exp(tX)^*g = g \quad \text{ for all } \ t \in R.$$

Differentiating the above equation with respect to t we get

$$\mathfrak{o}(V, g) = \{X \in \mathfrak{gl}(V) \mid g(Xv, w) + g(v, Xw) = 0 \quad \text{for all}\, v, w \in V\}.$$

Homework Problem 17. Let (V, ω) be a symplectic vector space. Show that the map

$$V \times V \to \mathfrak{sp}(V, \omega)^*, \quad (v, u) \mapsto \{X \mapsto \frac{1}{2}\omega(Xv, u)\}$$

gives rise to a vector space isomorphism $S^2(V) \to \mathfrak{sp}(V, \omega)^*$, where $S^2(V)$ denotes the symmetric tensors in $V \otimes V$.

Homework Problem 18. Let (V, g) be a vector space with an inner product. Show that the map

$$V \times V \to \mathfrak{o}(V, \omega)^*, \quad (v, u) \mapsto \{X \mapsto g(Xv, u)\}$$

gives rise two a vector space isomorphism $\Lambda^2(V) \to \mathfrak{o}(V, \omega)^*$, where $\Lambda^2(V)$ denotes the skew-symmetric tensors in $V \otimes V$.

15. Lecture 15. Group actions

Definition 88. An **action** of a group G on a set N is a map $G \times N \longrightarrow N$, $(g, x) \mapsto g \cdot x$ such that

1. $1_G \cdot x = x$ for all $x \in N$, where 1_G denotes the identity element of G;
2. $(g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x)$ for all $g_1, g_2 \in G$ and all $x \in N$.

Definition 89. A **smooth action** of a Lie group G on a manifold N is an action $G \times N \to N$ which is also a smooth map.

Example 48. Let V be a (finite dimensional) vector space. The general linear group $\mathrm{GL}(V)$ acts on V: for $A \in \mathrm{GL}(V)$, $v \in V$ we define $A \cdot v := A(v)$.

Example 49. Denote by $E(n)$ ("E" for "Euclidean") the group of rigid motions of $\mathbb{R}^n$: $E(n) = \{f : \mathbb{R}^n \longrightarrow \mathbb{R}^n \mid \|f(x) - f(y)\| = \|x - y\| \text{ for all } x, y \in \mathbb{R}^n\}$. Clearly the Euclidean group $E(n)$ acts on $\mathbb{R}^n$:

$$E(n) \times \mathbb{R}^n \to \mathbb{R}^n, \quad (f, x) \mapsto f(x).$$

It is not entirely obvious that $E(n)$ is a Lie group. To show that it is a Lie group, we observe first that $E(n)$ is generated by translations and rotations. Indeed if $f \in E(n)$ and if $v = f(0)$ then $T(w) := f(w) - v$ is also a rigid motion but $T(0) = 0$. Now if $T \in E(n)$, and $T(0) = 0$, then, since T sends a triangle with vertices at the origin to a congruent triangle, T has to be a rotation.

Next observe that we can identify $E(n)$ with a subgroup of $\mathrm{GL}(n+1, \mathbb{R})$ as follows. First identify $\mathbb{R}^n$ with a subset of $\mathbb{R}^{n+1}$: $\mathbb{R}^n = \{\begin{pmatrix} v \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1} \mid v \in \mathbb{R}^n\}$. Then $E(n) = \{\begin{pmatrix} A & w \\ 0 & 1 \end{pmatrix} : A \in \mathrm{O}(n), w \in \mathbb{R}^n\}$ for $\begin{pmatrix} A & w \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v \\ 1 \end{pmatrix} = \begin{pmatrix} Av + w \\ 1 \end{pmatrix}$.

Since $E(n)$ is the product $\mathrm{O}(n) \times \mathbb{R}^n$, $E(n)$ is a Lie group. Clearly it is a closed subgroup of $\mathrm{GL}(n+1, \mathbb{R})$ and the action of $E(n)$ on $\mathbb{R}^n$ is smooth.

Example 50. Let G be a Lie group. Then G acts on itself in three different ways:

1. Left multiplication: $L : G \times G \longrightarrow G$ $L(g, x) = gx =: L_g x$
2. Right multiplication (by inverse): $R : G \times G \longrightarrow G$, $R(g, x) = xg^{-1}$.
3. Conjugation: $C : G \times G \longrightarrow G$, $C(g, x) = gxg^{-1} := c_g(x)$.

One can think of a smooth action of a Lie group G on a manifold N as a "smooth" group homomorphism $\rho : G \to \mathrm{Diff}(N)$, where $\mathrm{Diff}(N)$ is the group of diffeomorphisms of N and "smooth" should be defined is such a way that the map $G \times N \to N$ given by $(g, x) \mapsto \rho(g)(x)$ is smooth.

Given a map of Lie groups $\tau : G \to H$ we get a corresponding map of Lie algebras $\delta\tau : \mathfrak{g} \to \mathfrak{h}$. By analogy, given an action $\rho : G \to \mathrm{Diff}(N)$ we should get a map of corresponding Lie algebras $\delta\rho$ from the Lie algebra $\mathfrak{g}$ to the Lie algebra of $\mathrm{Diff}(N)$), which, for various reasons is taken to be the Lie algebra of vector fields $\chi(N)$.

Here is one reason. Recall that for $X \in \mathfrak{g}$ the curve $t \mapsto \exp tX$ in G is a one-parameter subgroup so that

$$\exp(s+t)X = \exp sX \exp tX.$$

If a Lie group G acts on a manifold N then for every $X \in \mathfrak{g}$ we have a collection of diffeomorphisms $\phi_t^X(x) := (\exp tX)\cdot x$, which is a one parameter group since

$$\begin{aligned} \phi_{t+s}^X(x) = & \quad (\exp(t+s)X)\cdot x \\ = & \quad ((\exp tX)(\exp sX))\cdot x \\ = & \quad (\exp tX)\cdot((\exp sX)\cdot x) \\ = & \quad \phi_t^X(\phi_s^X(x)). \end{aligned}$$

Denote the vector field corresponding to the one parameter group $\{\phi_t^X\}$ by X_N, that is,

$$X_N(x) := \frac{d}{dt}\Big|_{t=0}\phi_t^X(x) = \frac{d}{dt}\Big|_{t=0}((\exp tX)\cdot x).$$

It therefore makes sense to define $\delta\rho(X) = X_N$. Unfortunately the map $\delta\rho$ is an **anti**-Lie algebra map:

$$[X,Y]_N = -[X_N, Y_N].$$

This annoying fact is not obvious and will discussed later.

Example 51. The group $G = \mathbb{R}^n$ acts on $\mathbb{R}^n$ by "left multiplication": $(g,v) \mapsto g+v$. The Lie algebra of $\mathbb{R}^n$ is $\mathbb{R}^n$ with the trivial bracket. Consequently the exponential map $\exp : \mathrm{Lie}(\mathbb{R}^n) \longrightarrow \mathbb{R}^n$ is the identity map. Therefore $X_{R^n}(v) = \frac{d}{dt}(\exp tX)\cdot v = \frac{d}{dt}(tX+v) = X$, a constant vector field.

Example 52. The group $G = \mathrm{SO}(3)$ acts on $M = \mathbb{R}^3$ in the standard fashion: $(A,v) \mapsto Av$. The Lie algebra of $\mathfrak{so}(3)$ consists of 3×3 skew-symmetric matrices. For $\xi \in \mathfrak{so}((3)$ the corresponding induced vector field $\xi_{\mathbb{R}^3}(v) = \frac{d}{dt}\big|_0(e^{t\xi}\cdot v) = \xi v$.

Lifted actions. If a group G acts on a manifold M, then there exists a "lifted" action of G on the cotangent bundle T^*M of M. Namely, given a group homomorphism $\tau : G \longrightarrow \mathrm{Diff}(M)$, $g \mapsto \tau_g \in \mathrm{Diff}(M)$ we get group homomorphism $\widetilde{\tau} : G \longrightarrow \mathrm{Diff}(T^*M)$ defined as follows.

Since τ_g is a diffeomorphism of M, $\widetilde{\tau}_g := (d(\tau_g^{-1}))^T$ is a diffeomorphim of T^*M. Moreover, it is easy to check that $\widetilde{\tau}_g \circ \widetilde{\tau}_h = (\tau_g \circ \tau_h)^{\sim} = \widetilde{\tau}_{gh}$.

Note that a lifted action preserves the Liouville one-form α_M:

$$(\widetilde{\tau}_g)^*\alpha_M = \alpha_M$$

for all $g \in G$ and hence preserves the symplectic form $\omega_{T^*M} = d\alpha_M$.

Definition 90. An action τ of G on a symplectic manifold (M,ω) is **symplectic** if $\tau_g^*\omega = \omega$ for all $g \in G$.

Example 53. Given an action of a Lie group G on a manifold N, the corresponding lifted action of G on (T^*N, ω_{T^*N}) is symplectic.

Example 54. Let M be the two-torus $T^2 = S^1 \times S^1$. The two form $\omega = d\theta_1 \wedge d\theta_2$ is symplectic. The torus $G = S^1 \times S^1$ acts on M: the action is given by $(\lambda_1, \lambda_2) \cdot (\theta_1, \theta_2) = (\theta_1 + \lambda_1, \theta_2 + \lambda_2)$; it is symplectic.

Lemma 91. *Suppose an action of a Lie group G on a symplectic manifold (M, ω) is symplectic. Then the image of the corresponding anti-Lie algebra map $\mathfrak{g} \to \chi(M)$ is contained in the subalgebra of the symplectic vector fields.*

Proof. For any $X \in \mathfrak{g}$ we have $(\tau_{\exp tX})^* \omega = \omega$ for all $t \in \mathbb{R}$. Differentiating both sides with respect to t we get $(\tau_{\exp tX})^*(L_{X_M}\omega) = 0$. Hence $L_{X_M}\omega = 0$, i.e., the induced vector field X_M is symplectic. $\square$

Recall that $L_{X_M}\omega = 0$ if and only if the form $\iota(X_M)\omega$ is closed.

Definition 92. A symplectic action τ of G on (M, ω) is **Hamiltonian** if there is an anti-Lie algebra map $\gamma : \mathfrak{g} \longrightarrow (C^\infty(M), \{\cdot, \cdot\})$, $X \mapsto \varphi^X$ such that $d\varphi^X = \iota(X_M)\omega$, i.e., the diagram

$$\begin{tikzcd}
X \arrow[rrd] & & \mathfrak{g} \arrow[r, "\delta\tau", "\text{anti-Lie}"'] \arrow[dr, dotted, "\gamma"'] & \chi(M,\omega) & X_f \\
& & \varphi^X & C^\infty(M) \arrow[u, "\text{Lie}"'] & f \arrow[u]
\end{tikzcd}$$

commutes.

Note that the map γ being an anti-Lie algebra map means:

1. $\gamma : \mathfrak{g} \longrightarrow C^\infty(M)$ is linear,
2. $-\{\varphi^X, \varphi^Y\} = \varphi^{[X,Y]}$ for all $X, Y \in \mathfrak{g}$.

We will see later on that such a map γ need not be unique.

Example 55. The symplectic action of Example 54 is **not** Hamiltonian (why?).

A good source of examples of Hamiltonian actions is the lifted actions.

Proposition 93. *Suppose $\tau : G \longrightarrow$ Diff(Q) is a smooth action of a Lie group G on a manifold Q. The corresponding lifted action $\widetilde{\tau} : G \longrightarrow$ Diff(T^*Q) is Hamiltonian.*

Proof. Since the lifted action preserves the Liouville one-form α_Q, we have $L_{X_{T^*Q}}\alpha_Q = 0$ for any $X \in \mathfrak{g}$. Hence $0 = d\iota(X_{T^*Q})\alpha_Q + \iota(X_{T^*Q})d\alpha_Q = d\iota(X_{T^*Q})\alpha_Q + \iota(X_{T^*Q})\omega_{T^*Q}$. Consequently

$$\iota(X_{T^*Q})\omega_{T^*Q} = -d(\iota(X_{T^*Q})\alpha_Q).$$

Here we may define $\gamma(X) = \varphi^X$ to be $-\iota(X_{T^*Q})\alpha_Q$. It remians to show that $\{\varphi^X, \varphi^Y\} = -\varphi^{[X,Y]}$ for all $X, Y \in \mathfrak{g}$. Now

$$\begin{aligned}
\{\varphi^X, \varphi^Y\} &= L_{X_{T^*Q}}\varphi^Y \\
&= L_{X_{T^*Q}}(-\iota(Y_{T^*Q})\alpha) \\
&= -\iota(L_{X_{T^*Q}}Y_{T^*Q})\alpha - \iota(Y_{T^*Q})L_{X_{T^*Q}}\alpha \\
&= -\iota([X_{T^*Q}, Y_{T^*Q}])\alpha + 0 \\
&= -\iota(-[X,Y]_{T^*Q})\alpha \\
&= -\varphi^{[X,Y]}
\end{aligned}$$

□

Example 56. Let (M, ω) be a compact symplectic manifold with the property that its first de Rham cohomology group vanishes: $H^1_{\mathrm{DR}}(M) = 0$ ((for example we may take $M = S^2$). Then any symplectic action on (M, ω) is Hamiltonian.

Recall what the condition $H^1_{\mathrm{DR}}(M) = 0$ means: if θ is a closed one-form on M then $\theta = df$ for some function f. Hence if $\tau : G \longrightarrow \mathrm{Diff}(M, \omega)$ is a symplectic action, then $d\iota(X_M)\omega = 0$ implies that $\iota(X_M)\omega = d\varphi^X$ for some function φ^X, which we know up to a constant.

We fix the constant by requiring that $0 = \int_M \varphi^X \omega^n$, where $n = \frac{1}{2}\dim M$. To show that the action is indeed Hamiltonian with this choice of constants it is enough to check that if for some $X, Y \in \mathfrak{g}$ we have $\int_M \varphi^X \omega^n = 0$ and $\int_M \varphi^Y \omega^n = 0$, then $\int_M \{\varphi^X, \varphi^Y\}\omega^n = 0$.

Now $\{\varphi^X, \varphi^Y\}\omega^n = (L_{X_M}\varphi^Y)\cdot\omega^n = L_{X_M}(\varphi^Y\omega^n)$ because $L_{X_M}\omega = 0$. Let ψ_t denote the flow of X_M. Then $\int_M \{\varphi^X, \varphi^Y\}\omega^n = \int_M L_{X_M}(\varphi^Y\omega^n) = \int \frac{d}{dt}\big|_0 \psi_t^*(\varphi^Y\omega^n) = \frac{d}{dt}\big|_0 \int_M \psi_t^*(\varphi^Y\omega^n) = \frac{d}{dt}\big|_0 \int_M \varphi^Y\omega^n = 0$

Let us consider a concrete special case of the above general example.

Example 57. Let $M = S^2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid \sum x_i^2 = 1\}$, and let $\omega = (x_1 dx_2 \wedge dx_3 + x_2 dx_3 \wedge dx_1 + x_3 dx_1 \wedge dx_2)\big|_{S^2}$, standard area form. The group S^1 acts on S^2 by

$$e^{2\pi i\theta} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

It is not difficult to compute that the induced vector field $(\frac{\partial}{\partial\theta})_{S^2} = x_2\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_2}$ and that $\iota((\frac{\partial}{\partial\theta})_{S^2})\omega = dx_3\big|_{S^2}$. Consequently $\varphi^{\frac{\partial}{\partial\theta}} = (x_3 + \mathrm{const})\big|_{S^2}$.

Another important source of examples of Hamiltonian group actions are linear group actions on symplectic vector spaces.

Definition 94. A **representation** ρ of a Lie group G on a vector space V is a Lie group homomorphism $\rho : G \longrightarrow \mathrm{GL}(V)$.

Note that a representation is a smooth action. Suppose (V, ω) is a symplectic vector space. A representation $\rho : G \longrightarrow \mathrm{GL}(V)$ is **symplectic** if $\rho(g)^*\omega = \omega$ for all $g \in G$. In other words a symplectic representation is a Lie group homomorphism $\rho : G \to \mathrm{Sp}(V, \omega)$.

Lemma 95. *A symplectic representation $\rho : G \to \mathrm{Sp}(V, \omega)$ is a Hamiltonian action of G on the symplectic manifold (V, ω).*

Proof. For $X \in \mathfrak{g}$ denote the induced vector field by X_V. It is easy to check that $X_V(v) = d\rho(X)v$ where $d\rho(V) \in \mathfrak{sp}(V, \omega) \subset \mathfrak{gl}(V) = \mathrm{Hom}(V, V)$. By Poincaré lemma $\iota(X_V)\omega = d\varphi^X$ for some function φ^X which we can normalize by requiring that $\varphi^X(0) = 0$. We leave it as an exercise to check that $\varphi^X(x) = \frac{1}{2}\omega(X_V x, x)$ and that the map $\mathfrak{g} \to C^\infty(V)$ defined by $X \mapsto \varphi^X$ is an anti-Lie algebra map. $\square$

16. Lecture 16. Moment map

We start the lecture by proving a result that justifies the definition of a Hamiltonian group action. Recall that a function f is a constant of motion for a function h on a symplectic manifold if f is constant along the integral curves of the Hamiltonian vector field of h (see Definition 64). A result due to E. Noether shows that symmetries are responsible for conservation laws. More precisely we have

Theorem 96 (E. Noether)**.** *Consider a Hamiltonian action of a Lie group G on a symplectic manifold (M, ω). Let $\gamma : \mathfrak{g} \to C^\infty(M)$, $\gamma : X \mapsto \varphi^X$ be a corresponding anti-Lie algebra map.*

Let $h \in C^\infty(M)$ be a G-invariant smooth function. Then for any $X \in \mathfrak{g}$ the function $\gamma(X) = \varphi^X$ is a constant of motion for h.

Proof. Since h is G-invariant, for any $X \in \mathfrak{g}$ we have $h((\exp tX) \cdot x) = h(x)$ for all $t \in \mathbb{R}$ and all $x \in M$. It follows by differentiation with respect to t that $X_M(h) = 0$.

By definition of φ and the Poisson bracket, $\{\varphi^X, h\} = X_M(h) = 0$. Thus h is a constant of motion for φ^X. Hence by skew symmmetry of the Poisson bracket (cf. Lemma 65) φ^X is a constant of motion for h. $\square$

Remark 97. Note that Noether's theorem above holds under a weaker assumption. Namely it is enough to only assume that for every element of the Lie algebra X the induced vector field X_M is Hamiltonian; it is not necessary to assume that the map $\gamma : \mathfrak{g} \to C^\infty(M)$ is an anti-Lie algebra map.

However, in many cases the map γ is an anti-Lie algebra map (for example, for lifted actions). We'll see that this extra property of γ has an important consequence — equivariance of a corresponding moment map.

Remark 98. Noether originally proved the theorem for symmetric Lagrangian systems.

Example 58. We now attempt to demystify Example 36.

Consider a particle in $\mathbb{R}^3$ in a central force field. That is, consider the manifold $M = T^*\mathbb{R}^3$ with coordinates $(q_1, q_2, q_3, p_1, p_2, p_3)$ and with the canonical the symplectic form $\omega = \sum dq_i \wedge dp_i$. A central force Hamiltonian is a function of the form $h(q, p) = \frac{1}{2m} \sum p_i^2 + V(||q||^2)$.

The action of the group $\mathrm{O}(3)$ on $\mathbb{R}^3$ preserves the standard inner product. Therefore the potential V is $\mathrm{O}(3)$-invariant. The kinetic energy $\frac{1}{2m} \sum p_i^2$ is invariant under the lifted action of $\mathrm{O}(3)$ on $T^*\mathbb{R}^3$, hence the whole Hamiltonian is invariant under the lifted action. The Lie algebra $\mathfrak{o}(3)$ of O(3), which consists of 3×3 skew-symmetric matrices, is three-dimensional. Therefore we have three constants of motion: j_1, j_2 and j_3 (cf. Example 36), corresponding to a basis of $\mathfrak{o}(3)$.

Which matrices in $\mathfrak{o}(3)$ give rise to these constants of motion?

If you are familiar with elementary physics, you may notice that the three functions j_1, j_2 and j_3 of Examples 36 and 58 are components of one vector value function $j : T^*R^3 \to \mathbb{R}^3$, the angular momentum.

Example 59. The above example can be generalized as follows. Consider a Riemannian manifold (Q, g). Suppose a Lie group G acts on Q and preserves the metric g. Let V be a G invariant function of Q. Then the Hamiltonian $H(q,p) = \frac{1}{2} g^*(q)(p,p) + V(q)$ on T^*Q is invariant under the lifted action of G. Therefore, for any $X \in \mathfrak{g}$ we get an constant of motion φ^X of H.

Note that these constants of motion are not independent: if $X_1, \dots X_r$ is a basis of the Lie algebra of G, then any $X \in \mathfrak{g}$ is a linear combination of the X_i's and hence φ^X is a linear combination of φ^{X_i}'s.

To make sense of these constants of motion it is useful to introduce the analog of the notion of angular momentum for an arbitrary Hamiltonian group action. The analog is the notion of a moment map.

Definition 99. Consider a Hamiltonian action of a Lie group G on a symplectic manifold (M, ω). Let $\gamma : \mathfrak{g} \to C^\infty(M)$, $X \mapsto \varphi^X$ be a corresponding anti-Lie algebra map. A **moment map** $\Phi : M \to \mathfrak{g}^*$ corresponding to the action is defined by

$$\langle \Phi(m), X \rangle = \varphi^X(m) = \gamma(X)(m)$$

for $X \in \mathfrak{g}$ and $m \in M$, where $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ is the canonical pairing.

Remark 100. We can think $\Phi : M \to \mathfrak{g}^*$ as the transpose of γ provided we think of the manifold M as a subset of the linear dual of $C^\infty(M)$: each $m \in M$ defines a linear evaluation map $ev_m : C^\infty(M) \to \mathbb{R}$ by $ev_m(f) = f(m) =$"$lam, f\rangle$". Then $\langle \Phi(m), X \rangle = \gamma(X)(m) =$"$\langle m, \gamma(X) \rangle$".

Remark 101. Note that a moment map corresponding to a given action need not be unique. Recall that for a symplectic action of G on $(M.\omega)$ to be Hamiltonian we require that there exists an anti-Lie algebra map $\gamma : \mathfrak{g} \longrightarrow (C^\infty(M)$ with the property that $d\gamma(X) = \iota(X_M)\omega$ for every $X \in \mathfrak{g}$. Suppose $\gamma' : \mathfrak{g} \longrightarrow (C^\infty(M)$ is another anti-Lie algebra map satisfying $d\gamma'(X) = \iota(X_M)\omega$. Then $d(\gamma(X) - \gamma'(X)) = 0$. Consequently the map $c : \mathfrak{g} \to \mathbb{R}$ defined $c(X) = \gamma(X) - \gamma'(X)$ is linear (here we tacitly assumed that M is connected).

The requirement that both γ and γ' are anti-Lie algebra maps amounts to

$$\{\gamma(X) + c(X), \gamma(Y) + c(Y)\} = -\gamma([X,Y]) - c([X,Y])$$

for all $X, Y \in \mathfrak{g}$. Therefore, since $\{\gamma(X) + c(X), \gamma(Y) + c(Y)\} = \{\gamma(X), \gamma(Y)\} = -\gamma([X,Y])$, we must have $c([X,Y]) = 0$.

Conversely, given an anti-Lie algebra map $\gamma : \mathfrak{g} \to C^\infty(M)$ with $\iota(X_M)\omega = d\gamma(X)$ and an element $c \in \mathfrak{g}^*$ such that $c([X, Y]) = 0$ for any $X, Y \in \mathfrak{g}$, the map $\gamma' = \gamma + c$ is another anti-Lie algebra map satisfying $\iota(X_M)\omega = d\gamma'(X)$ for all $X \in \mathfrak{g}$.

Moment maps in general are difficult to compute directly from the definition. Three cases from an exception: lifted actions on the cotangent bundles, linear actions and actions on coadjoint orbits.

Example 60 (Lifted actions)**.** Suppose $\tau : G \longrightarrow \mathrm{Diff}(Q)$ is a smooth action of a Lie group G on a manifold Q. Let $\widetilde{\tau} : G \longrightarrow \mathrm{Diff}(T^*Q)$ denote the corresponding lifted action. We have seen (Proposition 93 that the lifted action of G on $(T^*Q, \omega = \omega_{T^*Q})$ is Hamiltonian:

$$\iota(X_{T^*Q})\omega = -d(\iota(X_{T^*Q})\alpha_Q).$$

where α_Q is the Liouville one-form. That is, we saw that we should define $\varphi^X = -\iota(X_{T^*Q})\alpha_Q$.

Let us compute the corresponding moment map. Let $p \in T^*_q Q$ be a covector. We claim that $\varphi^X(q, p) = -\langle p, X_Q(q)\rangle$. Indeed, since the action on T^*Q is lifted, we have $d\pi(X_{T^*Q})(q, p) = X_Q(q)$ where $\pi : T^*Q \to Q$ is the standard projection. Therefore by definition of α_Q we have $-\iota(X_{T^*Q})\alpha_Q(q, p) = \langle \alpha_Q(q, p), X_{T^*Q}(q, p)\rangle = \langle p, d\pi(X_{T^*Q}(q, p))\rangle = \langle p, X_Q(q)\rangle$. We conclude that the moment map $\Phi : T^*Q \to \mathfrak{g}^*$ is defined by

$$\langle \Phi(m), X\rangle(q, p) = \langle p, X_Q(q)\rangle$$

for all $X \in \mathfrak{g}$ and all $(q, p) \in T^*Q$.

In a number of special case, one can make this formula more concrete. For example let G be any connected Lie group, $Q = G$ and let the action of G on Q be given by left multiplication: $(g, a) \mapsto ga = L_g(a)$. Let us compute the induced vector fields $X_Q(a)$:

$$X_Q(a) = \frac{d}{dt}\Big|_0 L_{\exp tX}(a) = \frac{d}{dt}\Big|_0 (\exp tX) \cdot a = \frac{d}{dt}\Big|_0 R_a(\exp tX) = (dR_a)_1(X),$$

where R_a denotes right multiplication by a. Therefore $\varphi^X(a, \eta) = \langle \eta, dR_a(X)\rangle = \langle dR_a^T \eta, X\rangle$ where $dR_a : T_1G \longrightarrow T_aG$ and $(dR_a)^T : T^*_aG \longrightarrow T^*_1G = \mathfrak{g}^*$. We conclude that in this case the moment map $\Phi : T^*G \to \mathfrak{g}^*$ is given by

$$\Phi(a, \eta) = dR_a^T \eta.$$

We can specialize this example further by considering $G = \mathbb{R}^n$. Then $R_a(q) = a + q$ and hence $dR_a = id$. Therefore the moment map $\Phi : T^*\mathbb{R}^n = \mathbb{R}^n \times (\mathbb{R}^n)^* \to (\mathbb{R}^n)^*$ is simply the projection $\Phi(q, p) = p$, i.e., it is the linear momentum.

Example 61. Let (V, ω) be a (finite dimensional) symplectic vector space. The "birth certificate" representation of the symplectic group $id : \mathrm{Sp}(V, \omega) \to \mathrm{Sp}(V, \omega)$ is, by Lemma 95, a Hamiltonian action of $\mathrm{Sp}(V, \omega)$ on (V, ω). The map $\gamma : \mathfrak{sp}(V, \omega) \to C^\infty(V)$ can be chosen to be $X \mapsto \varphi^X(v) = \frac{1}{2}\omega(Xv, v)$. Therefore the moment map $\Phi : V \to \mathfrak{sp}(V, \omega)^*$ is given by

$$\langle \Phi(v), X\rangle = \frac{1}{2}\omega(Xv, v).$$

Recall that we have an isomorphism between the dual of the Lie algebra of the symplectic group and the symmetric 2-tensors on V (Homework 17) :

$$\begin{aligned} \psi : S^2(V) &\to \mathfrak{sp}(V,\omega)^* \\ v \odot w &\mapsto \{X \mapsto \frac{1}{2}\omega(Xv, w)\}, \end{aligned} \tag{19}$$

where $v \odot w$ denotes the symmetric tensor product of v and w. Therefore

$$\langle \Phi(v), X\rangle = \langle \psi(v), X\rangle.$$

Thus, under the identification ψ of $\mathfrak{sp}(V,\omega)^*$ with $S^2(V)$ the moment map Φ is given by

$$\Phi(v) = v \odot v.$$

We'd like to generalize this example to arbitrary symplectic representations $\rho : K \to \mathrm{Sp}(V,\omega)$.

Lemma 102. *Suppose we have a Hamiltonian action of a Lie group G on a symplectic manifold (M,ω). Let $\Phi : M \to \mathfrak{g}^*$ denote a corresponding moment map. Let $\rho : K \to G$ be a Lie group homomorphism. Then the action of K on M defined by*

$$(k, m) \mapsto \rho(k) \cdot m$$

is also Hamiltonian and a corresponding moment map Ψ is given by

$$\Psi = (\delta\rho)^T \circ \Phi,$$

where $\delta\rho : \mathfrak{k} \to \mathfrak{g}$ is the Lie algebra map induced by ρ and $(\delta\rho)^T : \mathfrak{g}^* \to \mathfrak{k}^*$ is its transpose.

Proof. For any $X \in \mathfrak{k}$ the induced vector field X_M is defined by $X_M(m) = \frac{d}{dt}\big|_0(\rho(\exp tX)) \cdot m$. Now, by the universal property of the exponential map $\rho(\exp tX) = \exp(\delta\rho(tX))$. Hence $X_M(m) = (\delta\rho(X))_M(m)$ where the right hand side denotes the vector field induced by $\delta\rho(X) \in \mathfrak{g}$. By definition of the moment map

$$d\langle \Psi, X\rangle = \iota(X_M)\omega = \iota((\delta\rho(X))_M)\omega = d\langle \Phi, \delta\rho(X)\rangle.$$

Finally, since $\delta\rho : \mathfrak{k} \to \mathfrak{g}$ is a map of Lie algebras, the map $\mathfrak{k} \to C^\infty(M)$ given by $X \mapsto \langle \Phi, \delta\rho(X)\rangle$ is an anti-Lie algebra map. $\square$

It follows from Lemma 102 above that a moment map Φ for a representation $\rho : K \to \mathrm{Sp}(V,\omega)$ is given by $\Phi(v) = ((\delta\rho)^T \circ \psi)(v \odot v)$ where $\psi : S^2(V) \to \mathfrak{sp}(V,\omega)^*$ is defined in equation (19) above.

Adjoint and coadjoint representations. Recall that a Lie group G acts on itself by conjugation: for all $g \in G$ we have a map $c_g : G \longrightarrow G$ defined by $c_g(a) = gag^{-1}$, and $c_g \circ c_h = c_{gh}$. We therefore have the **Adjoint** representation $Ad : G \longrightarrow \mathrm{GL}(\mathfrak{g})$ defined by $Ad(g)(X) = d(c_g)(1)X$ where $1 \in G$ is the identity element. Since $c_g \circ c_h = c_{gh}$, $Ad(g) \circ Ad(h) = Ad(gh)$, so Ad is a group homomorphism, i.e., Ad is indeed a representation. The action of a Lie group G on its Lie algebra $\mathfrak{g}$ by the Adjoint representation is called the **Adjoint action**.

Given a representation ρ of a group G on a vector space V, there exists an associated representation $\rho^\dagger$ of G on the dual vector space V^*: $\rho^\dagger(g)\ell := \ell \circ \rho(g^{-1})$ for $\ell \in V^*$ and $g \in G$. Thus associated to the Adjoint representation $Ad : G \to \mathrm{GL}(\mathfrak{g})$ we have the **coadjoint** representation $Ad^\dagger : G \to \mathrm{GL}(\mathfrak{g}^*)$:

$$\langle Ad^\dagger(g)f, X\rangle := \langle f, Ad(g^{-1})X\rangle \tag{20}$$

for $f \in \mathfrak{g}^*$, $X \in \mathfrak{g}$ and $g \in G$.

Definition 103. Let $G \times Q \to Q$ be an action of a group G on a set Q. An **orbit** of an element $x \in Q$ is the set

$$G \cdot x := \{y \in Q \mid y = g \cdot x \text{ for some}\, g \in G\}.$$

It is an important fact that the orbits of a coadjoint action (**coadjoint orbits** for short) are naturally symplectic manifolds. We'll see this in the next lecture.

Definition 104. Let G be a group acting on two sets Q_1 and Q_2. A map $f : Q_1 \to Q_2$ is **equivariant** if $f(g \cdot q) = g \cdot f(q)$ for all $g \in G$ and all $q \in Q_1$.

Theorem 105. *Let (M, ω) be a symplectic manifold and let $G \times M \to M$ be a Hamiltonian actions of a* **connected** *Lie group G. Then a corresponding moment map $\Phi : M \to \mathfrak{g}^*$ is equivariant: for any $g \in G$ and any $m \in M$*

$$\Phi(g \cdot m) = Ad^\dagger(g)\Phi(m).$$

To prove the theorem we need a number of technical results.

Lemma 106. *Let G be a connected Lie group. Then any neighborhood U of the identity generates G as a group: for any $g \in G$ there are $u_1, \dots, u_k \in U$ (k depends on g) such that*

$$g = u_1 \dots u_k.$$

Proof. Let $U^r = \{u_1 \dots u_r \mid u_i \in U, \quad 1 \leq i \leq r\}$, the set consisting of products of r elements of U. Then $U^1 = U$. Also $U^2 = \cup_{x\in U} xU$, hence is open. It follows by induction that the sets U^r are open for all r. Consequently the set $V = \bigcup_{l=1}^\infty U^l$ is a connected open subgroup of G. If V is not all of G then the cosets of V partition G into a union of open sets, which contradicts connectedness of G. Therefore $G = V = \bigcup_{l=1}^\infty U^l$. □

Lemma 107. *Let G be a connected Lie group. Then any element g of G is of the form*

$$g = \exp X_1 \exp X_2 \dots \exp X_k$$

for some X_1, X_2, … $X_k \in \mathfrak{g}$ (k depends on g).

Proof. Since the exponential map is a local diffeomorphism, there exists a neighborhood of the identity U in G such that all $u \in U$ are of the form $\exp X$, $X \in \mathfrak{g}$. Now apply Lemma 106. □

Proposition 108. *Let $f : M_1 \longrightarrow M_2$ be a smooth map between two manifolds. Let Y_i be a vector field on M_i, $i = 1, 2$. Let φ_t^i denote the flow of Y_i. If $df(Y_1) = Y_2 \circ f$, then $f(\varphi_t^1(x)) = \varphi_t^2(f(x))$, for all $t \in \mathbb{R}$ and $x \in M_1$.*

Proof. Fix $x \in M_1$. Let $\gamma_1(t) = f(\varphi^1_t(x))$ and let $\gamma_2(t) = \varphi^2_t(f(x))$. Bother curves pass through $f(x)$ at $t = 0$. The curve γ_2 is integral curve of Y_2. So to prove that the two curves are equal it is enough to check that γ_1 is also an integral curve of Y_2. Now $\frac{d}{dt}\gamma_1(t) = \frac{d}{dt}f(\varphi^1_t(x)) = df(\frac{d}{dt}\varphi^1_t(x)) = df(Y_1(\varphi^1_t(x))) = Y_2(f(\varphi^1_t(x))) = Y_2(\gamma_1(t))$. Therefore $\gamma_1(t) = \gamma_2(t)$, i.e. $f(\varphi^1_t(x)) = \varphi^2_t(f(x))$. □

Theorem 109. *Let G be a Lie group with the Lie algebra $\mathfrak{g}$. Let $ad : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ be the map of Lie algebras corresponding to the Adjoint representation $Ad : G \to \mathrm{GL}(\mathfrak{g})$. Then for all $X, Y \in \mathfrak{g}$*

$$ad(X)Y = [X, Y],$$

where $[\cdot, \cdot]$ is the bracket on $\mathfrak{g}$.

Proof. This amounts to unwinding a string of definitions. See [Warner, p. 112]. □

Remark 110. The map of Lie algebras $ad : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is called the **adjoint** representation.

Note that by definition $ad(X)(Y) = \frac{d}{dt}\big|_{t=0} Ad(\exp tX)Y$, which is the value of the vector field $X_{\mathfrak{g}}$ on $\mathfrak{g}$ induced by the adjoint action. In other words the theorem above asserts that $X_{\mathfrak{g}}(Y) = [X, Y]$.

Corollary 111. *Let G be a Lie group, $X \in \mathfrak{g}$. Then the vector field $X_{\mathfrak{g}^*}$ induced on $\mathfrak{g}^*$ by the coadjoint action satisfies*

$$\langle X_{\mathfrak{g}^*}(f), Y\rangle = -\langle f, [X, Y]\rangle$$

for $f \in \mathfrak{g}^*$ and $Y \in \mathfrak{g}$. That is,

$$X_{\mathfrak{g}^*}(f) = -ad(X)^T f.$$

Proof. For $X, Y \in \mathfrak{g}$, $f \in \mathfrak{g}^*$ and $t \in \mathbb{R}$ we have

$$\langle Ad^\dagger(\exp tX)f, Y\rangle = \langle f, Ad(\exp(-tX))Y\rangle.$$

Now differentiate both sides with respect to t at $t = 0$. □

Proof of Theorem 105. It follows from Lemma 107 and Proposition 108 that it is enough to show: for $X \in \mathfrak{g}$ and $m \in M$

$$d\Phi(X_M(m)) = X_{\mathfrak{g}^*}(\Phi(m)). \tag{21}$$

Now for any $Y \in \mathfrak{g} = (\mathfrak{g}^*)^*$

$$\begin{aligned}
\langle d\Phi(X_M(m)), Y\rangle &= Y \circ d\Phi(X_M(m)) \\
&= d(Y \circ \Phi)(X_M(m)) \quad \text{since } Y : \mathfrak{g}^* \to \mathbb{R} \text{ is linear} \\
&= X_M(Y \circ \Phi)(m) = X_M(\langle \Phi, Y\rangle)(m) \quad \text{since } Y \circ \Phi(m) = \langle \Phi(m), Y\rangle \\
&= \{\langle \Phi, X\rangle, \langle \Phi, Y\rangle\}(m) \quad \text{since } X_M(\psi) = \{\langle \Phi, X\rangle, \psi\} \text{ for any function } \psi \\
&= \langle \Phi(m), -[X, Y]\rangle \quad \text{since } \{\phi^X, \phi^Y\} = \phi^{-[X,Y]} \\
&= \langle X_{\mathfrak{g}^*}(\Phi(m)), Y\rangle \quad \text{by Corollary 111.}
\end{aligned}$$

Therefore $d\Phi(X_M(m)) = X_{\mathfrak{g}^*}(\Phi(m))$ and we are done. □

We finish the lecture with a corollary of Lemma 107.

Corollary 112. *Let (M, ω) be a symplectic manifold with a Hamiltonian action of a connected Lie group G. Let $\Phi : M \to \mathfrak{g}^*$ be a corresponding moment map.*

The moment map carries enough information to recover the action of G on (M, ω).

Proof. By Lemma 107 it is enough to recover the actions of the elements of the form $\exp X$, $X \in \mathfrak{g}$. But the curves $t \mapsto (\exp tX) \cdot x$, $x \in M$, are integral curves of the Hamiltonian vector field of $\langle \Phi, X \rangle$. Therefore if we know the moment map Φ, we also know $\exp X \cdot x$ for $X \in \mathfrak{g}$, $x \in M$. $\square$

The corollary can be sharpenned as follows. Let $\gamma : \mathfrak{g} \to C^\infty(M)$ be an anti-Lie algebra map from a Lie algebra $\mathfrak{g}$ to the Poisson algebra on a symplectic manifold (M, ω). It is a difficult theorem (which we won't attempt even to sketch a proof of) that given a Lie algebra $\mathfrak{g}$, there exists a connected and simply-connected Lie group G whose Lie algebra is $\mathfrak{g}$. The map γ allows us to define the action of elements of G of the form $\exp X$, $X \in \mathfrak{g}$. The fact that γ is an anti-Lie algebra map together with simple connectedness of G guarantees that these actions cohere into an action of G.

Homework Problem 19. Consider the action of S^1 on $\mathbb{C}^n$ given by $e^{i\theta} \cdot (z_1, \dots, z_n) = (e^{i\theta} z_1, \dots, e^{i\theta} z_n)$. Show that the action preserves the imaginary part of the Hermitian inner product (which is a symplectic form). What is a corresponding moment map? Is it unique?

17. Lecture 17. Coadjoint orbits

In this section we study coadjoint oribts. We start by collecting more examples of Lie groups and their Lie algebras.

Example 62. Let $G = \mathrm{GL}(n, \mathbb{R})$. The conjugation is given by $c_g(a) = gag^{-1}$. Hence $Ad(g)(X) = \frac{d}{dt}\big|_0 g \exp tX g^{-1} = gXg^{-1}$. Therefore, $[Y, X] = ad(Y)X = \frac{d}{dt}\big|_0 (\exp tY) X (\exp(-tY)) = YX - XY$.

Example 63. Let $G = \mathrm{GL}(n, \mathbb{C})$, the invertible linear maps on $\mathbb{C}^n$. Note that $\mathrm{GL}(n, \mathbb{C})$ is a subgroup of $\mathrm{GL}(2n, \mathbb{R})$ since $\mathbb{C}^n = \mathbb{R}^{2n}$ as real vector spaces. Same argument as in the previous example shows that the Lie bracket on the Lie algebra $\mathfrak{gl}(n, \mathbb{C})$ is also given by the commmutator $[Y, X] = YX - XY$.

Example 64 (Unitary group)**.** The unitary group $U(n)$ is the group of unitary matrices. That is,

$$U(n) = \{A \in \mathrm{GL}(n, \mathbb{C}) \mid AA^* = I\}$$

where A^* is the conjugate transpose of A. It is not hard to see that the Lie algebra $\mathfrak{u}(n)$ of $U(n)$ is

$$\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) \mid X^* + X = 0\}$$

the space of skew-Hermitian matrices, and that the Lie bracket on $\mathfrak{u}(n)$ is the commutator.

Note that there is a way to identify the $\mathfrak{u}(n)$ with its dual $\mathfrak{u}(n)^*$. Namely, $(X, Y) = -\operatorname{tr} XY$ is a symmetric pairing invariant under conjugation. Since any skew-Hermitian matrix is conjugate to a diagonal matrix, and since $-\operatorname{tr} X^2 > 0$ for a non-zero skew-Hermitian diagonal matrix X, the pairing is a positive definite inner product. Since the inner product is invariant under conjugation, the corresponding isomorphism $\psi : \mathfrak{u}(n) \to \mathfrak{u}(n)^*$, $X \mapsto (X, \cdot)$ intertwines the Adjoint and the coadjoint actions: $\psi(Ad(g)X) = Ad^\dagger(\psi(X))$.

Example 65. Let $G = \mathbb{R}^n$ with the group operation given by vector addition. Since $\mathbb{R}^n$ is abelian, $c_g(a) = a$ for all $g \in \mathbb{R}^n$. Therefore $Ad(g) = id$ and hence $[X, Y] = ad(X)Y = 0$ for all $X, Y \in \mathfrak{g} = \mathbb{R}^n$.

Example 66 (Heisenberg group)**.** Let (V, ω) be a symplectic vector space. Let $\mathcal{H} = V \times \mathbb{R}$. We make $\mathcal{H}$ into a Lie group by defining the multiplication to be $(v_1, t_1) \cdot (v_2, t_2) = (v_1 + v_2, \frac{1}{2}\omega(v_1, v_2) + t_1 + t_2)$. The element $(0, 0)$ is clearly the identity element. Show that the Lie bracket on the Lie algebra $\mathfrak{H} \simeq V \times \mathbb{R}$ of the Heisenberg group is given by $[(X, \lambda), (Y, \mu)] = (0, \omega(X, Y))$. Hint: Theorem 109.

Definition 113. Suppose a group G acts on a set Q. The **isotropy subgroup** of an element $q \in Q$ is the set

$$G_q := \{g \in G \mid g \cdot q = q\}.$$

It is easy to see that G_q is indeed a subgroup. Note also that the evaluation map $ev_q : G \to Q$ given by $ev_q(g) = g \cdot q$ induces a bijection between left cosets G/G_q and the orbit $G \cdot q = \{y \in Q \mid y = g \cdot q \text{ for some} g \in G\}$.

Moreover, if G is a Lie group and the action $G \times Q \to Q$ is continuous then the evaluation map ev_q is continuous. Consequently the isotropy group $G_q := (ev_q)^{-1}(q)$ is closed. Furthermore it follows from the closed subgroup theorem (Theorem 87) that the isotropy subgroup G_q is a Lie subgroup of G.

Suppose G is a Lie group and $H \subset G$ is a closed subgroup. Then the coset space G/H inherits from G a topology making the projection $\pi : G \to G/H$, $\pi(g) = gH$ continuous.

Proposition 114. *Let G be a Lie group and H be a closed subgroup of G. Then the coset space G/H is a smooth manifold, $T_{1H}(G/H) \simeq \mathfrak{g}/\mathfrak{h}$ and the action of G on G/H given by $(g, aH) \mapsto gaH$ is smooth.*

Sketch of a proof. Choose a subspace $\mathfrak{m}$ of $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ as vector spaces. Consider the map $\mathfrak{h} \oplus \mathfrak{m} \longrightarrow G$ defined by $(X, Y) \mapsto \exp X \exp Y$. This is a diffeomorphism near $0 \oplus 0 \in \mathfrak{h} \oplus \mathfrak{m}$. For a small enough neighborhood U of 0 in $\mathfrak{m}$ the set $\exp(U)$ is a model for G/H, i.e. $\pi : \exp(U) \longrightarrow \exp(U)H \subseteq G/H$ is a homeomorphism. Thus $\pi \circ \exp : U \to G/H$ gives coordinates on G/H near $1H$. Then for any $g \in G$ the map $\pi \circ L_g \circ \exp : U \to G/H$ gives coordinates near gH. Note that $\mathfrak{m} \simeq \mathfrak{g}/\mathfrak{h}$ and that $\mathfrak{m} \simeq T_{1H}(G/H)$. □

Proposition 115. *Let G be a Lie group acting smoothly on a manifold M. Then for every $x \in M$ the map $\phi : G/G_x \to M$, $\phi(gG_x) = g \cdot x$ is a one-to-one immersion (where G_x denotes the isotropy group of x). Consequently if the group G is compact, the orbits of G are embedded submanifolds of M.*

Proof. It is easy to check that the map ϕ is well-defined and is one-to-one. The smoothness of ϕ follows from the fact that for a fixed $x \in M$ the map $G \to M$, $g \mapsto g \cdot x$ is smooth and from the definition of the smooth structure on G/G_x.

Recall that the group G acts on G/G_x by $g \cdot aG_x = gaG_x$. Since $\phi(g \cdot aG_x) = g \cdot \phi(aG_x)$, i.e., since ϕ is equivariant, and since the action of G on G/G_x has only one orbit, it is enough to check that the differential $d\phi$ is injective at one point. For example, it is enough to check that $d\phi(1G_x) : T_{1G_x}(G/G_x) \longrightarrow T_xM$ is one-to-one.

Note that for $X \in \mathfrak{g}$

$$\begin{aligned} X_M(x) = 0 &\iff \frac{d}{dt}\Big|_{t=0}(\exp tX) \cdot x = 0 \\ &\iff d((\exp sX) \cdot (\,))\left(\frac{d}{dt}\Big|_{t=0}(\exp tX) \cdot x\right) = 0 \quad \text{for all } s \\ &\iff \frac{d}{dt}\Big|_{t=s}(\exp tX) \cdot x = 0 \quad \text{for all } s \\ &\iff \exp sX \in G_x \quad \text{for all } s \\ &\iff X \in \mathfrak{g}_x, \text{ the Lie algebra of } G_x, \quad \text{by the closed subgroup theorem.} \end{aligned}$$

Consequently the map

$$\mathfrak{g}/\mathfrak{g}_x \to T_xM \quad X + \mathfrak{g}_x \mapsto X_M(x)$$

is well-defined and is injective.

Since $d\pi : \mathfrak{g} = T_1G \to T_{1G_x}(G/G_x)$ induces the isomorphism of $\mathfrak{g}/\mathfrak{g}_x$ and of the tangent space $T_{1G_x}(G/G_x)$, and since $\phi \circ \pi(g) = g \cdot x$ it follows that $d\phi(1G_x)$ is injective. $\square$

Note that the above argument shows that for an orbit $G \cdot x$, the tangent space $T_x(G \cdot x)$ satisfies

$$T_x(G \cdot x) = \{X_M(x) \mid X \in \mathfrak{g}\}.$$

In particular if $G \cdot f$ is a coadjoint orbit of G, then $T_f(G \cdot f) = \{X_{\mathfrak{g}^*}(f) \mid X \in \mathfrak{g}\} = \{-ad(X)^T f \mid X \in \mathfrak{g}\}$.

Theorem 116 (Kirillov-Kostant-Souriau)**.** *Let G be a Lie group. The coadjoint orbits $\mathcal{O} = Ad^\dagger(G)f_0$ carry a natural symplectic form $\omega_{\mathcal{O}}$ defined as follows :*

$$\omega_{\mathcal{O}}(f)(X_{\mathfrak{g}^*}(f), Y_{\mathfrak{g}^*}(f)) = \langle f, [X, Y]\rangle,$$

for $X, Y \in \mathfrak{g}$ and $f \in \mathcal{O}$. Moreover the action of G on $(\mathcal{O}, \omega_{\mathcal{O}})$ is Hamiltonian and a corresponding moment map is the inclusion $\mathcal{O} \hookrightarrow \mathfrak{g}^*$.

Proof. We first show that $\omega = \omega_{\mathcal{O}}$ is well-defined. For this we need to check that if $X_{\mathfrak{g}^*}(f) = 0$ then $\langle f, [X, Y]\rangle = 0$ as well. But $\langle f, [X, Y]\rangle = \langle f, ad(X)Y\rangle = \langle ad(X)^T f, Y\rangle = \langle -X_{\mathfrak{g}^*}(f), Y\rangle$ by Corollary 111. So ω is well-defined.

The smoothness of ω follows from the fact that $\omega(f)(X_{\mathfrak{g}^*}(f), Y_{\mathfrak{g}^*}(f)) = \langle -X_{\mathfrak{g}^*}(f), Y\rangle = \langle X_{\mathfrak{g}^*}, -d\phi^Y\rangle(f)$ where $\phi^Y \in C^\infty(\mathcal{O})$ is the smooth function defined by $\phi^Y(f) = \langle f, Y\rangle = Y|_{\mathcal{O}}(f)$.

To prove that the form $\omega_{\mathcal{O}}$ is non-degenerate we check that $\omega(f)(X_{\mathfrak{g}^*}(f), Y_{\mathfrak{g}^*}(f)) = 0$ for all $Y \in \mathfrak{g}$ implies that $X_{\mathfrak{g}^*}(f) = 0$. Now

$$0 = \omega(f)(X_{\mathfrak{g}^*}(f), Y_{\mathfrak{g}^*}(f)) = \langle ad^T(X)f, Y\rangle$$

for all $Y \in \mathfrak{g}$ implies that $0 = ad^T(X)f = -X_{\mathfrak{g}^*}(f)$. Therefore the form $\omega_{\mathcal{O}}$ is non-degenerate.

Next we check that the form is G-invariant, i.e., that $(Ad^\dagger(g))^*\omega = \omega$ for any $g \in G$. On the one hand, by definition

$$((Ad^\dagger(g))^*\omega)(f)(X_{\mathfrak{g}^*}(f), Y_{\mathfrak{g}^*}(f)) = \omega(Ad^\dagger(g)(f))(dAd^\dagger(g)X_{\mathfrak{g}^*}(f), dAd^\dagger(g)Y_{\mathfrak{g}^*}(f))$$

and

$$\begin{aligned} dAd^\dagger(g)(X_{\mathfrak{g}^*}(f)) &= \frac{d}{dt}\Big|_0 Ad^\dagger(g)(Ad^\dagger(\exp tX)f) \\ &= \frac{d}{dt}\Big|_0 Ad^\dagger(g\exp tX)f \\ &= \frac{d}{dt}\Big|_0 Ad^\dagger(g(\exp tX)g^{-1})Ad^\dagger(g)f \\ &= \frac{d}{dt}\Big|_0 Ad^\dagger(\exp(tAd(g)X))Ad^\dagger(g)f \\ &= (Ad(g)X)_{\mathfrak{g}^*}(Ad^\dagger(g)f) \end{aligned}$$

Therefore

$$\begin{aligned} &((Ad^\dagger(g))^*\omega)(f)(X_{\mathfrak{g}^*}(f), Y_{\mathfrak{g}^*}(f)) \\ &= \omega(Ad^\dagger(g)f)\Big((Ad(g)X)_{\mathfrak{g}^*}(Ad^\dagger(g)f), (Ad(g)Y)_{\mathfrak{g}^*}(Ad^\dagger(g)f)\Big) \\ &= \langle Ad^\dagger(g)f, [Ad(g)X, Ad(g)Y]\rangle \\ &= \langle Ad^\dagger(g)f, Ad(g)[X,Y]\rangle \quad \text{since } Ad(g): \mathfrak{g} \to \mathfrak{g} \text{ is a Lie algebra map} \\ &= \langle f, [X,Y]\rangle \end{aligned}$$

which proves G-invariance.

Finally to prove that the form ω is closed, it is enough to show $\iota(Y_{\mathfrak{g}^*})d\omega = 0$ for all $Y \in \mathfrak{g}$. Since ω is G-invariant $0 = L_{Y_{\mathfrak{g}^*}}\omega = d\iota(Y_{\mathfrak{g}^*})\omega + \iota(Y_{\mathfrak{g}^*})d\omega = dd\varphi^Y + \iota(Y_{\mathfrak{g}^*})d\omega$. Hence $\iota(Y_{\mathfrak{g}^*})d\omega = 0$.

To summarize: we proved ω is well-defined, smooth, symplectic, invariant. We also showed that for all $Y \in \mathfrak{g}$ we have $d\varphi^Y = \iota(Y_{\mathfrak{g}^*})\omega$ where $\varphi^Y = Y\big|_{\mathcal{O}}$. Consequently a moment map Φ satisfies $\langle \Phi(f), Y\rangle = \varphi^Y(f) = \langle f, Y\rangle$. Therefore $\Phi(f) = f$.

It remains to show that Φ is equivariant. Now $\{\varphi^X, \varphi^Y\}(f) = \langle X_{\mathfrak{g}^*}(f), Y\rangle = \langle -ad(X)^T f, Y\rangle = -\langle f, [X,Y]\rangle = -\varphi^{[X,Y]}(f)$. □

We now give an example to show that adjoint and coadjoint orbits are very different.

Example 67 (Heisenberg group)**.** Recall the definition of the Heisenberg group: one starts with a symplectic vector space (V, ω) and sets $\mathcal{H} == V \times \mathbb{R}$ with multiplication give by $(v,t)\cdot(u,s) = (v+u, \frac{1}{2}\omega(v,u) + t + s)$. Recall also that $(v,t)\cdot(0,0) = (v,t)$ so $(0,0) = 1$, $(v,t)\cdot(-v,-t) = (0,0)$ and that $c_{(v,t)}(u,s) = (v,t)\cdot(u,s)\cdot(-v,-t) = (v+u, \frac{1}{2}\omega(v,u) +$

$t+s)\cdot(-v,-t) = (u, \frac{1}{2}\omega(v+u,-v)+\frac{1}{2}(v,u)+t+s-t) = (u, s+\omega(v,u))$. Consequently $Ad(v,t)(X,x) = (X, x+\omega(v,X))$. It follows that

$$Ad(G)(X,x) = \begin{cases} \{X\}\times\mathbb{R} & \text{if } X\neq 0,\\ \{(0,x)\} & \text{if } X = 0.\end{cases}$$

Note that the dimensions of adjoint orbits are zero and one. Note also that $ad(Y,y)(X,x) = (0,\omega(Y,X))$, i.e., $[(Y,y),(X,x)] = (0,\omega(Y,X))$

Let us now compute the coadjoint action and coadjoint orbits. Let $(Y^*,y^*)\in\mathfrak{g}^* = V^*\times\mathbb{R}^*$, $(X,x)\in\mathfrak{g} = V\times\mathbb{R}$. $\langle Ad^\dagger(v,t)(Y^*,y^*),(X,x)\rangle = \langle (Y^*,y^*), Ad(-v,-t)(X,x)\rangle = \langle (Y^*,y^*),(X,x-\omega(v,X))\rangle = \langle Y^*,X\rangle - (y^*\circ\omega^\sharp(v))(X) + \langle y^*,x\rangle = \langle (Y^*-y^*\circ\omega^\sharp(v),y^*),(X,x)\rangle$. Therefore

$$Ad^\dagger(v,t)(Y^*,y^*) = (Y^*-y^*\circ\omega^\sharp(v),y^*).$$

We conclude:

$$Ad^\dagger(G)(Y^*,y^*) = \begin{cases} \{(Y^*,0)\} & \text{if } y^* = 0,\\ V^*\times\{y^*\} & \text{if } y^*\neq 0.\end{cases}$$

Homework Problem 20. Let G be the subgroup of $\mathrm{GL}(2,\mathbb{R})$ consisting of the matrices

$$\{\begin{pmatrix} a & b\\ 0 & 1\end{pmatrix} \mid a,b\in\mathbb{R}, a\neq 0\}$$

This is the so called "$ax+b$" group, the group of affine motions of the line. Show that the adjoint orbits are either one- or zero-dimensional. Compute the coadjoint orbits and the corresponding symplectic forms.

18. Lecture 18. Reduction

Consider a system of particles in $\mathbb{R}^3$ interacting via a potential which is invariant under translations: $V(\vec{x}_1+\vec{v},\dots,\vec{x}_N+\vec{v}) = V(\vec{x}_1,\dots,\vec{x}_N)$. In this case the linear momentum of the system is conserved. It is then standard to fix the total linear momentum and pass to a coordinate system located at the center of mass of particles.

Having passed to a center of mass coordinates we may discover that the potential is rotationally invariant, so that the angular momentum of the system is conserved. In this case we would fix the angular momentum and pass to a steadily rotating coordinate system.

We'll see that both of these procedures are instances of **symplectic reduction**, a procedure invented independently in early 1970's by K. Meyer and by J. Marsden and A. Weinstein.

Mathematically the procedure is the "correct" way to define quotients in the symplectic category. Suppose we have a Hamiltonian action of a Lie group G on a symplectic manifold (M,ω) with a corresponding moment map $\Phi: M\to\mathfrak{g}^*$. If we were to define the symplectic quotient of the manifold M by the action of G to be the ordinary quotient M/G, we would have two problems:

1. the quotient M/G need not be a manifold

2. even if the quotient were a manifold, there would be no reason for it to be even-dimensional, let alone admit a symplectic structure.

We get around these to problems by using the moment map. Suppose zero is a regular value of the moment map Φ. Then the zero level set $\Phi^{-1}(0)$ is a submanifold of M. Since Φ is equivariant and 0 is fixed by the coadjoint action, the action of G preserves $\Phi^{-1}(0)$. It turns out that the action of G on $\Phi^{-1}(0)$ has to be (almost) free. In the good case the action is actually free and the quotient $M_0 := \Phi^{-1}(0)/G$ is a manifold. Moreover, we'll see that M_0 has a natural symplectic form ω_0 such that $\pi^*\omega_0 = \omega|_{\Phi^{-1}(0)}$, where $\pi : \Phi^{-1}(0) \to M_0$ is the orbit map. The manifold M_0 is called the **reduced space** at zero or the **symplectic quotient** at zero.

The above construction and its various generalizations are now in wide use in different areas of mathematics. The original motivation however came from the study of symmetric Hamiltonian systems which we will now sketch. Let $(M, \omega, \Phi : M \to \mathfrak{g}^*)$ be as above, and let $h \in C^\infty(M)^G$ be a G-invariant function. We will refer to the quadruple $(M, \omega, \Phi : M \to \mathfrak{g}^*, h)$ as a **symmetric Hamiltonian system**. Recall that for any $X \in \mathfrak{g}$, the function $\langle \Phi, X\rangle$ is a constant of motion for an invariant function h. Therefore the flow ψ_t^h of the Hamiltonian vector field X_h of h preserves the level sets $\Phi^{-1}(\mu)$, $\mu \in \mathfrak{g}^*$.

Lemma 117. *Let $(M, \omega, \Phi : M \to \mathfrak{g}^*, h)$ be a symmetric Hamiltonian system. The flow $\psi_t = \psi_t^h$ of the Hamiltonian vector field of h is G-equivariant: $\psi_t(g \cdot m) = g \cdot \psi_t(m)$.*

Proof. Since the function h and the symplectic form ω are G-invariant, the Hamiltonian vector field X_h of h is G-invariant as well. Consequently its flow is G-equivariant (cf. Proposition 108). □

It follows that the flow $\psi_t^h : \Phi^{-1}(0) \to \Phi^{-1}(0)$ induces a flow $\bar{\psi}_t : \Phi^{-1}(0)/G \to \Phi^{-1}(0)/G$ on the symplectic quotient $M_0 = \Phi^{-1}(0)/G$. On the other hand, since h is G-invariant, the restriction $h|_{\Phi^{-1}(0)}$ descends to a function h_0 on M_0. The last piece of the reduction procedure is the claim that the flow of the Hamiltonian vector field $\psi_t^{h_0}$ of h_0 on (M_0, ω_0) is the induced flow $\bar{\psi}_t$.

Let us illustrate all of the above ideas in a simple example.

Example 68. Let $h(q, p) \in C^\infty(T^*\mathbb{R}^3)$ be a smooth function on the cotangent bundle of the three-space. Suppose $h(q_1, q_2, q_3 + a, p_1, p_2, p_3) = h(q_1, q_2, q_3, p_1, p_2, p_3)$ for all a, i.e., h does not depend on the third position coordinate q_3. The assumption is equivalent to: the function h is G-invariant where $G = \mathbb{R}$ acts on $T^*\mathbb{R}^3$ by translations in q_3. Note that this is a lift of the action of $\mathbb{R}$ on $\mathbb{R}^3$, hence is Hamiltonian. The moment map is easy to compute. For example we can do it from the first principles: the induced vector field is $\frac{\partial}{\partial q_3}$. Since $\iota(\frac{\partial}{\partial q_3}) \sum dq_i \wedge dp_1 = dp_3$, we may choose the moment map to be $\Phi(q, p) = p_3$. The map is conserved by the flow of h. This follows from the lemma above, but one can also see this directly: Hamilton's equations are

$$\begin{cases} \dot{q_i} = \frac{\partial h}{\partial p_i} \\ \dot{p_i} = -\frac{\partial h}{\partial q_i} \end{cases}, \tag{22}$$

hence $\dot{p_3} = -\frac{\partial h}{\partial q_3} = 0$ since h is independent of q_3. The zero level set of the moment map is $\Phi^{-1}(0) = \{(q,p) : p_3 = 0\}$. It is not hard to see that $\Phi^{-1}(0)/G$ is symplectomorphic to $T^*\mathbb{R}^2$. Indeed consider the embedding $j : T^*\mathbb{R}^2 \longrightarrow \Phi^{-1}(0)$ given by $j(x_1, x_2, \eta_1, \eta_2) = (x_1, x_2, 0, \eta_1, \eta_2, 0)$. The embedding clearly parameterizes the orbits of G in the level set $\Phi^{-1}(0)$, hence the quotient $\Phi^{-1}(0)/G$ is T^*R^2 as a differntial manifold. Note also that $j^*\omega = dx_1 \wedge d\eta_1 + dx_2 \wedge d\eta_2$. Hence $\omega\big|_{\Phi^{-1}(0)} = dq_1 \wedge dp_1 + dq_2 \wedge dp_2 + dq_3 \wedge d0 = dq_1 \wedge dp_1 + dq_2 \wedge dp_2 = \pi^*(j^*\omega)$ where $\pi : \Phi^{-1}(0) \to T^*\mathbb{R}^2$ is the projection $\pi(q_1, q_2, q_3, p_1, p_2, 0) = (q_1, q_2, p_1, p_2)$. Therefore, under the identification of $\Phi^{-1}(0)/G$ with T^*R^2, the reduced form ω_0 is $j^*\omega$.

The restriction $h|_{\Phi^{-1}(0)}$ descends to a function h_0 on the quotient $\Phi^{-1}(0)/G$. Under the identification of the quotient with $T^*\mathbb{R}^2$, $h_0(x_1, x_2, \eta_1, \eta_2) = h(x_1, x_2, 0, \eta_1, \eta_2, 0)$. The Hamilton's equation for the function h_0 on T^*R^2 are

$$\begin{cases} \dot{x_i} = \frac{\partial h}{\partial \eta_1} \\ \dot{\eta_i} = -\frac{\partial h}{\partial x_i} \end{cases}$$

These equations carry less information than the equations (22): we dropped the equation $\dot{q_3} = \frac{\partial h}{\partial p_3}$. They do, however, give us the flow of X_h modulo the action of the group G. Note also that the restriction of the Hamiltonian vector field of h to $\Phi^{-1}(0)$ does project down under π to the Hamiltonian vector field of h_0 on $T^*R^2 = \Phi^{-1}(0)/G$.

We now start developing the necessary mathematical background. Recall that a continuous map $f : X \to Y$ between two topological spaces is **proper** if the preimage under f of a compact set is compact.

Definition 118. An action of a Lie group G on a manifold M is **proper** if the map $G \times M \to M \times M$ defined by $(g, m) \mapsto (g \cdot m, m)$ is proper.

Note that an action of a compact group is automatically proper. Proper actions have a number of good properties: all the isotropy groups are compact, all orbits are closed, the orbit space is Hausdorff. We will prove the latter two properties later if we have time. The first one is easy to prove.

Definition 119. An action of a group G on a set X is **free** if for any $g \in G$, $x \in X$, the equation $g \cdot x = x$ implies that $g = 1$.

Example 69. Consider the flow of a vector field $\frac{\partial}{\partial\theta_1} + a\frac{\partial}{\partial\theta_2}$ on T^2 for a irrational. The flow is an action of the real line $\mathbb{R}$. The action is free but not proper. All the orbits are dense. The quotient is not Hausdorff.

Remark 120. When a group G acts on a set X, we have a natural map $\pi : X \to X/G$ which sends a point in X to its orbit in X/G. We will refer to π as the **orbit map**.

Definition 121. A **fiber bundle** a quadruple (N, B, F, π) where

1. N is a manifold called the total space,
2. B is a manifold called the base,
3. F is a manifold called a typical fiber,
4. $\pi : N \longrightarrow B$ is a smooth map, called the projection, such that for all $b \in B$ there exists a neighborhood U of b in B and a diffeomorphism $\psi : U \times F \longrightarrow \Pi^{-1}(U) \subseteq N$ with $\psi(u, f) \in \pi^{-1}(u)$ for all $u \in U$. In other words, the diagram

$$\begin{tikzcd} U \times F \arrow[rr, "\psi"] \arrow[dr, "pr_1"'] & & \pi^{-1}(U) \arrow[dl, "\pi"] \\ & U, & \end{tikzcd}$$

where $pr_1(u, f) = u$, commutes.

It follows from the definition that $\pi : N \longrightarrow B$ is a submersion and that for all $b \in B$ the fibers $\pi^{-1}(b)$ are diffeomorphic to the typical fiber F. We will often write $F \longrightarrow N \xrightarrow{\pi} B$ to denote the fiber bundle with total space N, base B, typical fiber F and projection π. Also one often refers to the total space N as a fiber bundle over B.

Example 70. Let B and F be any two manifolds, $N = B \times F$, $\pi : B \times F \longrightarrow B$ the projection onto the first factor. This is a fiber bundle called a **trivial bundle**.

Example 71. Any vector bundle is a fiber bundle.

Example 72. Let $S^{2n-1} = \{z \in \mathbb{C}^n \,|\, \|x\|^2 = 1\}$ be the standard odd-dimensional sphere. It is a fiber bundle over the complex projective space $\mathbb{C}P^{n-1} = \{\ell \,|\, \ell \text{ a line in } \mathbb{C}^n\} = (\mathbb{C}^n \smallsetminus \{0\})/\mathbb{C}^\times = \{z \,|\, z \in \mathbb{C}^n \smallsetminus \{0\}, z \neq 0\}/\sim$ where $z \sim \lambda z$, for any $\lambda \in \mathbb{C}$, $\lambda \neq 0$. The projection $\pi : S^{2n-1} \longrightarrow \mathbb{C}P^{n-1}$ is given by $\pi(z) = [z]$, where $[z]$ is the complex line through z.

To check that $S^1 \to S^{2n-1} \xrightarrow{\pi} \mathbb{C}P^{n-1}$ is a fiber bundle, we need to produce trivializations. Let $U_i = \{[z] \in \mathbb{C}P^{n-1} \,|\, z_i \neq 0\}$. Define $\psi_i : S^1 \times U_i \longrightarrow S^{2n-1}$ by $\psi(e^{i\theta}, [z]) = \frac{e^{i\theta}(z_1, \cdots, z_n)}{\|(z_1, \cdots, z_n)\|}$.

Definition 122. Let G be a Lie group. A fiber bundle of the form $G \longrightarrow P \xrightarrow{\pi} B$ is a **principal G-bundle** if G acts on P and for all $b \in B$, there exists a neighborhood U of B and a trivialization $\psi : U \times G \longrightarrow \pi^{-1}(U)$ which is equivariant, i.e. $g \cdot \psi(u, a) = \psi(u, ga)$.

Example 73. The bundle $S^1 \to S^{2n-1} \xrightarrow{\pi} \mathbb{C}P^{n-1}$ is a principal S^1-bundle.

Note that if $G \longrightarrow P \xrightarrow{\pi} B$ is a principal bundle, then the action of G on P is free, proper and transitive on the fibers. The converse is also true.

Theorem 123. *Suppose a Lie group G acts freely and properly on a manifold P. Then the quotient $B = P/G$ is a Hausdorff manifold and the orbit map $\pi : P \to B$ makes P into a principal G-bundle over B.*

Proof. Postponed indefinitely. See Appendix B of Cushman and Bates, *Global Aspects of Classical Integrable Systems* for details. □

We now make two blanket assumptions for the rest of this lecture:

1. All actions are proper.
2. If an isotropy group of a point is zero-dimensional then it is trivial.

The second assumption has content: consider the action of S^1 on $S^3 \subset \mathbb{C}^2$ given by $\lambda \cdot (z_1, z_2) = (\lambda^2 z_1, \lambda z_2)$. The isotropy group of $(1, 0)$ is $\{\pm 1\}$. All other isotropy groups are trivial.

We can now state the two main theorems of symplectic reduction (keep in mind the two blanket assumptions we have made!).

Theorem 124 (Marsden-Weinstein, Meyer)**.** *Consider a Hamiltonian action of a Lie group G on a symplectic manifold (M, ω) with a corresponding moment map $\Phi : M \to \mathfrak{g}^*$. Suppose 0 is a regular value of the moment map. Then $\Phi^{-1}(0)$ is a submanifold of M.*

Moreover, the action of G on $\Phi^{-1}(0)$ has zero dimensional isotropy groups. Hence $M_0 := \Phi^{-1}(0)/G$ is a smooth manifold and the orbit map $\pi : \Phi^{-1}(0) \to M_0$ makes $\Phi^{-1}(0)$ into a principal G bundle over M_0.

Finally, there exists a symplectic form ω_0 on M_0 such that $\pi^*\omega_0 = \omega|_{\Phi^{-1}(0)}$.

Definition 125. The symplectic manifold (M_0, ω_0) is called the **reduced symplectic space** at zero. It is also referred to as the **symplectic quotient** at zero and is denoted by $M//G$. The later notation is often used when more than one group is involved.

Suppose $G \longrightarrow P \xrightarrow{\pi} B$ is a principal G bundle and suppose $h \in C^\infty(P)$ is a G-invariant function. Then there exists a unique function $h_0 \in C^\infty(B)$ such that $\pi^* h_0 = h$: h_0 is defined by $h_0(\pi(p)) = h(p)$.

Theorem 126 (Marsden-Weinstein, Meyer)**.** *Let $(M, \omega, \Phi : M \to \mathfrak{g}^*, h)$ be a symmetric Hamiltonian system. Suppose that 0 is a regular value of the moment map Φ. Let $\pi : \Phi^{-1}(0) \to M_0$ denote the orbit map; let h_0 be the unique function on M_0 such that $\pi^* h_0 = h|_{\Phi^{-1}(0)}$.*

The Hamiltonian vector field X_h of h on M and the Hamiltonian vector field X_{h_0} of h_0 on M_0 are π-related:

$$d\pi(X_h) = X_{h_0} \circ \pi.$$

The rest of the lecture is devoted to the proof of the Marsden-Weinstein-Meyer reduction theorems. We start by recalling a few definition and facts.

- If V a vector space and $U \subset V$ a subspace, then the annihilator U° of U in V^* is the subspace $\{\ell \in V^* \mid \ell\big|_U = 0\}$.
- If (V, ω) is a symplectic vector space and $U \subset V$ is a subspace, then the symplectic perpendicular of U in V is the subspace $U^\omega = \{v \in V \mid \omega(v, u) = 0, \forall u \in U\}$. Recall also that $(U^\omega)^\omega = U$.
- Suppose a group G acts on a manifold M. Then the tangent space to the orbit $G \cdot x$ at $x \in M$ is

$$T_x(G \cdot x) = \{\xi_M(x) \mid \xi \in \mathfrak{g}\},$$

where $\xi_M(x) = \frac{d}{dt}\Big|_0 \exp(t\xi)\cdot x$. The Lie algebra of the isotropy group G_x of x is

$$\mathfrak{g}_x = \{\xi \in \mathfrak{g} \mid \xi_M(x) = 0\}.$$

The following easy observation is crucial in the proof of the reduction theorems.

Lemma 127. *Suppose a Lie group G acts in a Hamiltonian fashion on a symplectic manifold (M,ω) with a corresponding moment map $\Phi : M \longrightarrow \mathfrak{g}^*$. Then for $x \in M$, $v \in T_xM$, and $\xi \in \mathfrak{g}$,*

$$\langle d\Phi_x(v), \xi\rangle = \omega(x)(\xi_M(x), v). \tag{23}$$

Proof. Since $\langle\cdot,\xi\rangle : \mathfrak{g}^* \to \mathbb{R}$ is linear, $\langle d\Phi_x(v), \xi\rangle = d(\langle\Phi,\xi\rangle)_x(v) = \omega(x)(\xi_M(x), v)$ by the definition of the moment map. □

Corollary 128. *Suppose a Lie group G acts in a Hamiltonian fashion on a symplectic manifold (M,ω) with a corresponding moment map $\Phi : M \longrightarrow \mathfrak{g}^*$.*

1. *The annihilator of the image of the differential of the moment map at $x \in M$ is the isotropy Lie algebra of x:*
$$(\text{Image } d\Phi_x)^\circ = \mathfrak{g}_x.$$
2. *The kernel of the differential of the moment map at $x \in M$ is the symplectic perpendicular to the tangent space to the G-orbit through x:*
$$\ker d\Phi_x = (T_x(G\cdot x))^{\omega(x)}$$

Proof. Proof of (1): As we recalled above, a vector ξ is in the isotropy Lie algebra of x iff $\xi_M(x) = 0$. Since the form $\omega(x)$ is nondegenerate, the latter is true iff for all $v \in T_xM$ $0 = \omega(x)(\xi_M(x), v)$. But $\omega(x)(\xi_M(x), v) = \langle d\Phi_x(v), \xi\rangle$ by Lemma 127. Hence $\xi \in \mathfrak{g}_x$ iff ξ annihilates the image $d\Phi_x(T_xM)$.

Proof of (2): A vector v is in the kernel of the differential $d\Phi_x$ iff $0 = \langle d\Phi_x, \xi\rangle$ for all $\xi \in \mathfrak{g}$. Hence by Lemma 127, $v \in \ker d\Phi_x$ iff $0 = \omega(x)(\xi_M(x), v)$ for all $\xi \in \mathfrak{g}$. Since $T_x(G\cdot x) = \{\xi_M(x) \mid \xi \in \mathfrak{g}\}$, it follows that $v \in \ker d\Phi_x$ iff $v \in (T_x(G\cdot x))^{\omega(x)}$. □

Corollary 129. *Suppose a Lie group G acts in a Hamiltonian fashion on a symplectic manifold (M,ω) with a corresponding moment map $\Phi : M \longrightarrow \mathfrak{g}^*$.*

A point x is a regular point of Φ iff the isotropy group of x is zero dimensional.

Proof. A point x is a regular point of the moment map Φ iff the annihilator of the image of the moment map is zero. By part (1) of Corollary 128, the annihilator of the image is the Lie algebra of the isotropy group G_x. Hence x is a regular point of Φ iff G_x is zero dimensional. □

Combining the above corollary with our blanket assumptions we see that if zero is a regular value of the moment map then the action of G on $\Phi^{-1}(0)$ is free. Since we assumed all actions are proper, it follows from Theorem 123 that the zero level set $\Phi^{-1}(0)$ is a principal G bundle over the quotient $M_0 = \Phi^{-1}(0)/G$.

To finish the proof of Theorem 124, it is enough to show that there exists a two form ω_0 on M_0 with $\pi^*\omega_0 = \omega|_{\Phi^{-1}(0)}$ and that ω_0 is symplectic.

Notation 1. If a group G acts on a manifold M, we write τ_g for the image of $g \in G$ in the diffeomorphism group $\mathrm{Diff}(M)$.

Proposition 130. *Let $G \to P \xrightarrow{\pi} B$ be a principal G bundle. If a form $\nu \in \Omega^*(P)$ satisfies*

1. *ν is G invariant: $\tau_g^*\nu = \nu$ and*
2. *$\iota(\xi_M)\nu = 0$ for all $\xi \in \mathfrak{g}$*

then there exists a form $\nu_0 \in \Omega$ such that $\nu = \pi^*\nu_0$.

A differential form on P which is a pull-back by π of a form on the base B is called **basic**. Note that if $\nu_0 \in \Omega^*(B)$ is a form on the base then since π is constant on the orbits of G, we have that $\pi^*\nu_0$ is G invariant and that $\iota(\xi_M)\pi^*\nu_0 = 0$ for all $\xi \in \mathfrak{g}$, so the two conditions above are necessary for a form to be basic. The content of the proposition is that the two conditions are also sufficient.

Proof. Suppose $\nu \in \Omega^q(P)$ is a q-form satisfying the two conditions. We want to define a form $\nu_0 \in \Omega^q(B)$ such that $\pi^*\nu_0 = \nu$.

Fix $b \in B$ and let p be a point in $\pi^{-1}(b)$. The differential $d\pi_p : T_pP \to T_bB$ is surjective and its kernel is precisely the tangent space to the orbit $G \cdot p$. Let $v_1, \ldots, v_q \in T_bB$ be a collection of q vectors. Since $d\pi_p$ is onto there exist $\bar{v}_1, \ldots, \bar{v}_q \in T_pP$ such that $d\pi(\bar{v}_i) = v_i$, $1 \leq i \leq q$. We'd like to define $\nu_0(b)(v_1, \ldots, v_q) = \nu(p)(\bar{v}_1, \ldots, \bar{v}_q)$. For the definition to make sense we need to check that it doesn't depend on the choices made.

Say $d\pi(\bar{v}_1) = v_1 = d\pi(\bar{\bar{v}}_1)$. Then $\nu(p)(\bar{v}_1, \ldots, \bar{v}_q) = \nu(p)(\bar{v}_1 - \bar{\bar{v}}_1, \bar{v}_2, \cdots, \bar{v}_q) + \nu(p)(\bar{\bar{v}}_1, \bar{v}_2, \cdots, \bar{v}_q) = 0 + \nu(p)(\bar{\bar{v}}_1, \bar{v}_2, \cdots, \bar{v}_q)$ since $d\pi(\bar{v}_1 - \bar{\bar{v}}_1) = v_1 - v_1 = 0$ hence $\bar{v}_1 - \bar{\bar{v}}_1 = \xi_M(p)$ for some $\xi \in \mathfrak{g}$. Therefore $\nu_0(b)(\bar{v}_1, \ldots, \bar{v}_q)$ does not depend on the choice of $\bar{v}_1, \ldots, \bar{v}_q \in T_pP$.

Let us check that $\nu_0(b)(\bar{v}_1, \ldots, \bar{v}_q)$ does not depend on the choice of $p \in \pi^{-1}(b)$. If $p' \in \pi^{-1}(b)$, then $p' = g \cdot p$ for some $g \in G$. The independence follows from the invariance of ν and the fact that $\pi \circ \tau_g = \pi$ (for then $d\pi(d\tau_g(\bar{v})) = d\pi(\bar{v})$).

It remains to check that ν_0 is smooth. Since the question is local we may assume that $P = U \times G$ where U is an open subset of $\mathbb{R}^n$, $(n = \dim P - \dim G)$. Suppose $x_1, \ldots, x_n$ are coordinates on U. Since ν vanishes along the G-directions, it has to be of the form $\sum a_{i_1 \cdots i_q}(x, y) dx_{i_1} \wedge \ldots \wedge dx_{i_q}$ where $y \in G$. Since ν is G-invariant, the functions $a_{i_1 \cdots i_q}(x, y)$ do not depend on y. Therefore $\nu = \sum a_{i_1 \cdots i_q}(x) dx_{i_1} \wedge \ldots \wedge dx_{i_q} = \nu_0$. $\square$

By the Proposition above in order to show that there exists a two-form ω_0 on M_0 with $\pi^*\omega_0 = \omega|_{\Phi^{-1}(0)}$ it is enough to check that $\iota(\xi_M)(\omega\big|_{\Phi^{-1}(0)}) = 0$ for all $\xi \in \mathfrak{g}$ (since ω is G-invariant to begin with). Now for any $x \in \Phi^{-1}(0)$, the tangent space $T_x\Phi^{-1}(0)$ is the kernel of $d\Phi_x$ and for any $v \in \ker d\Phi_x$, we have by Lemma 127,

$$\omega(x)(\xi_M(x), v) = \langle d\Phi_x(v), \xi \rangle = \langle 0, \xi \rangle = 0.$$

Therefore $\omega|_{\Phi^{-1}(0)}$ is basic, i.e., $\omega|_{\Phi^{-1}(0)} = \pi^*\omega_0$ for a unique two-form ω_0 on M_0.

Since π^* is injective, in order to show that ω_0 is closed, it is enough to show that $\pi^*\omega_0$ is closed. Now $d(\pi^*\omega_0) = d(\omega|_{\Phi^{-1}(0)}) = (d\omega)|_{\Phi^{-1}(0)} = 0$.

It remains to prove that ω_0 is non-degenerate. If (V,ω) symplectic vector space and U is a subspace with the property that $U \subset U^\omega$ then the quotient U^ω/U is naturally a symplectic vector space. Therefore the form ω_0 is nondegenerate iff for any $x \in \Phi^{-1}(0)$, we have $T_x\Phi^{-1}(0)^{\omega(x)} = T_x(G\cdot x)$. On the other hand by Corollary 128, part (2), we have $\ker d\Phi_x = (T_x(G\cdot x))^{\omega(x)}$. Hence $T_x(G\cdot x) = ((T_x(G\cdot x))^{\omega(x)})^{\omega(x)} = \ker d\Phi_x = T_x\Phi^{-1}(0)$. This finishes the proof of Theorem 124.

Suppose now we have an invariant Hamiltonian $h \in C^\infty(M)^G$. Then $h|_{\Phi^{-1}(0)} = \pi^* h_0$ for some smooth function h_0 on M. We want to show that $d\pi(X_h) = X_{h_0}\circ\pi$. It is enough to show that $\iota(d\pi(X_h))\omega_0 = dh_0$. Now $\pi^* dh_0 = d\pi^* h_0 = di_0^* h = i_0^*(dh)$ where i_0 denotes the inclusion $\Phi^{-1}(0) \hookrightarrow M$. So it is enough to show that $\pi^*(\iota(d\pi(X_h))\omega_0) = i_0^*(dh)$. Fix $x \in \Phi^{-1}(0)$, $v \in T_x\Phi^{-1}(0)$. Then

$$\begin{aligned}\pi^*(\iota(d\pi(X_h))\omega_0)(x)(v) &= \iota(d\pi(X_h))\omega_0(\pi(x))(d\pi(v))\\ &= \omega_0(\pi(x))(d\pi(X_h), d\pi(v))\\ &= (\pi^*\omega_0)(x)(X_h, v)\\ &= (i_0^*\omega)(x)(X_h, v)\\ &= (i_0^* dh)(x)(v).\end{aligned}$$

This proves Theorem 126.

Homework Problem 21. Let (M,ω) be a symplectic manifold. Recall that a submanifold $Z \subseteq M$ is **coisotropic** if for all $z \in Z$,

$$(T_zZ)^\omega \subseteq T_zZ$$

where $(T_zZ)^\omega = \{v \in T_zM \mid \omega(z)(v,w) = 0,\ \forall w \in T_zZ\}$. Show that the null distribution

$$\mathcal{N} = \coprod_{z\in Z} (T_zZ)^\omega$$

is integrable in the sense of Frobenius, i.e. if X, Y are two vector fields defined on an open set $U \subset Z$ and satisfying $X(z), Y(z) \in (T_zZ)^\omega$, for all $z \in U$, then $[X,Y](z) \in (T_zZ^\omega$, for all $z \in U$.
Hint: For any vector fields X_1, X_2, X_3, $d\omega(X_1, X_2, X_3) = 0$.

Homework Problem 22. Consider $\mathbb{C} = \{x+iy \mid x, y \in \mathbb{R}\}$. Let $dz = dx + idy$, $d\bar{z} = dx - idy$. Show that $\omega = \frac{1}{i} dz \wedge d\bar{z}$ is a symplectic form preserved by the action of S^1 on $\mathbb{C}$:

$$e^{i\theta}\cdot z = e^{i\theta} z \qquad \text{(complex multiplication)}$$

Show that the action is Hamiltonian and that for any $c \in \mathbb{R}$,

$$\Phi(z) = |z|^2 + c$$

is a moment map (after an identification of $\mathrm{Lie}(S^1)^*$ with $\mathbb{R}$).

Homework Problem 23. (a) Suppose we have Hamiltonian actions of a Lie group G on symplectic manifolds (M_1,ω_1) and (M_2,ω_2). Suppose also $\Phi_1 : M_1 \longrightarrow \mathfrak{g}^*$, $\Phi_2 : M_2 \longrightarrow \mathfrak{g}^*$ are corresponding moment maps. Show that the action of G on $M_1 \times M_2$, $g\cdot(m_1,m_2) \stackrel{\text{def}}{=}$

$(g \cdot m_1, g \cdot m_2)$ is Hamiltonian and that $\Phi(m_1, m_2) = \Phi_1(m_1) + \Phi_2(m_2)$ is a corresponding moment map. (The symplectic form on $M_1 \times M_2$ is $\omega_1 + \omega_2$).

(b) Consider $\mathbb{C}^n = \{(z_1, \dots, z_n) \mid z_j \in \mathbb{C}\}$ with a form $\omega = \frac{1}{i} \sum_j dz_j \wedge d\bar{z}_j$. Show that ω is symplectic and that the action of S^1 on $\mathbb{C}^n$,

$$S^1 \times \mathbb{C}^n \longrightarrow \mathbb{C}^n, \quad (e^{i\theta}, (z_1, \dots, z_n)) \mapsto (e^{i\theta} z_1, \dots, e^{i\theta} z_n)$$

is Hamiltonian.
Moreover, show $\Phi(z) = \sum |z_i|^2 = \|z\|^2$ is a moment map. Compute the manifolds $\Phi^{-1}(\mu)/S^1$ for regular values of Φ.

Homework Problem 24. Let $\Phi : M \longrightarrow \mathfrak{g}^*$ be a moment map for an action of G on M. Show that for all $\mu \in \mathfrak{g}^*$,

$$\Phi^{-1}(Ad^\dagger(G)\mu)/G \simeq \Phi^{-1}(\mu)/G_\mu$$

as **sets**, where $G_\mu = \{g \in G \mid Ad^\dagger(g)\mu = \mu\}$ is the isotropy group of μ.

19. Lecture 19. Reduction at nonzero values of the moment map

We continue to assume that all actions are proper and that all zero dimensional isotropy groups are trivial.

The only property of zero that was used in the proof of Theorems 124 and 126 is that zero is fixed by coadjoint action. Therefore we can restate them as follows.

Theorem 131. *Let $(M, \omega, \Phi : M \to \mathfrak{g}^*, h)$ be a symmetric Hamiltonian system. Assume that the action of G on M is proper and that any zero dimensional isotropy group is trivial. Let $\mu \in \mathfrak{g}^*$ be a regular value of the moment map Φ.*

If μ is fixed by the coadjoint action then the action of G preserves $\Phi^{-1}(\mu)$. For any $z \in \Phi^{-1}(\mu)$ the isotropy group G_z is trivial. Hence $M_\mu := \Phi^{-1}(\mu)/G$ is a smooth manifold and the orbit map $\pi_\mu : \Phi^{-1}(\mu) \to M_\mu$ makes $\Phi^{-1}(\mu)$ into a principal G bundle over M_μ.

Moreover, there exists a symplectic form ω_μ on M_μ such that $(\pi_\mu)^*\omega_\mu = \omega|_{\Phi^{-1}(\mu)}$.

Finally, let h_μ be the unique function on M_μ such that $(\pi_\mu)^* h_\mu = h|_{\Phi^{-1}(\mu)}$. Then the Hamiltonian vector field X_h of h on M and the Hamiltonian vector field X_{h_μ} of h_μ on M_μ are π_μ-related:

$$d\pi_\mu(X_h) = X_{h_\mu} \circ \pi_\mu.$$

This begs the question of what to do when $\mu \in \mathfrak{g}^*$ is not fixed by the coadjoint action. There are two ways to proceed: the so called "point reduction" and the so called "orbital reduction." In the case of the former we observe that in general the level set $\Phi^{-1}(\mu)$ need not be preserved by the action of the full group G. However, since the moment map is equivariant, it is preserved by the action of the isotropy group G_μ of μ. The reduction theorem can then be modified as follows.

Theorem 132. *Let $(M, \omega, \Phi : M \to \mathfrak{g}^*, h)$ be a symmetric Hamiltonian system. Assume that the action of G on M is proper and that any zero dimensional isotropy group is trivial. Suppose that $\mu \in \mathfrak{g}^*$ is a regular value of the moment map Φ.*

Then the action of the isotropy group G_μ preserves $\Phi^{-1}(\mu)$. For any $z \in \Phi^{-1}(\mu)$ the isotropy group G_z is trivial. Hence $M_\mu := \Phi^{-1}(\mu)/G_\mu$ is a smooth manifold and the orbit map $\pi_\mu : \Phi^{-1}(\mu) \to M_\mu$ makes $\Phi^{-1}(\mu)$ into a principal G_μ bundle over M_μ.

Moreover, there exists a symplectic form ω_μ on M_μ such that $(\pi_\mu)^*\omega_\mu = \omega|_{\Phi^{-1}(\mu)}$.

Finally, let h_μ be the unique function on M_μ such that $(\pi_\mu)^* h_\mu = h|_{\Phi^{-1}(\mu)}$. Then the Hamiltonian vector field X_h of h on M and the Hamiltonian vector field X_{h_μ} of h_μ on M_μ are π_μ-related:

$$d\pi_\mu(X_h) = X_{h_\mu} \circ \pi_\mu.$$

We leave a direct proof of the above theorem as an exercise and now describe an alternative approach. Observe that if a group G acts on the symplectic manifolds (M_1, ω_1) and (M_2, ω_2) in a Hamiltonian manner with corresponding moment maps $\Phi_i : M_i \to \mathfrak{g}^*$, $i = 1, 2$, then the diagonal action of G on $(M_1 \times M_2, \omega_1 + \omega_2)$, $g \cdot (m_1, m_2) := (g \cdot m_1, g.m_2)$ is also Hamiltonian. Moreover, $\Psi(m_1, m_2) = \Phi_1(m_1) + \Phi_2(m_2)$ is a corresponding moment map [check the assertions].

Lemma 133. *Consider a Hamiltonian action of a Lie group G on a symplectic manifold (M, ω) with a corresponding moment map $\Phi : M \to \mathfrak{g}^*$. Suppose μ is a regular value of the moment map.*

Then 0 is a regular value of the moment map $\Psi : M \times \mathcal{O}_\mu^- \to \mathfrak{g}^*$ for the diagonal action of G on the product of (M, ω) with the coadjoint orbit $\mathcal{O}_\mu$ through μ, where $\mathcal{O}_\mu^-$ means that we consider the orbit with minus the standard symplectic form.

Conversely, if 0 is a regular value of Ψ, then μ is a regular value of Φ.

Proof. Suppose $\Phi(m) = f \in \mathfrak{g}^*$. Since the moment map Φ is equivariant, $\Phi(G \cdot m) = Ad^\dagger(G)\Phi(m) = Ad^\dagger(G)f$. It follows that the image of the differential $d\Phi_m : T_m M \to \mathfrak{g}^*$ contains the tangent space at f to the orbit coadjoint $Ad^\dagger(G)f$:

$$T_f(Ad^\dagger(G)f) \subset d\Phi_m(T_m M).$$

Next note that $\Psi^{-1}(0) = \{(m, f) \in M \times \mathcal{O}_\mu \mid \Phi(m) - f = 0\}$. Therefore for $(m, f) \in \Psi^{-1}(0)$ we have $d\Psi_{(m,f)}(T_m M \times T_f \mathcal{O}_\mu) = d\Phi_m(T_m M) + T_f \mathcal{O}_\mu = d\Phi_m(T_m M)$. □

Note that we can define a G-equivariant diffeomorphism $\sigma : \Phi^{-1}(\mathcal{O}) \to \Psi^{-1}(0)$ by $\sigma(m) = (m, \Phi(m)$. Note also that $\Phi^{-1}(\mu)$ embeds into $\Phi^{-1}(\mathcal{O})$ and that a G-orbit in $\Phi^{-1}(\mathcal{O})$ intersects $\Phi^{-1}(\mu)$ in a G_μ-orbit. Consequently $\Psi^{-1}(0)/G \simeq \Phi^{-1}(\mu)/G_\mu$ as manifolds. In fact more is true.

Lemma 134 (The shifting trick). *Consider a proper Hamiltonian action of a Lie group G on a symplectic manifold (M, ω) with a corresponding moment map $\Phi : M \to \mathfrak{g}^*$. Suppose μ is a regular value of the moment map. Let $\mathcal{O}_\mu^-$ denote the coadjoint orbit through μ with minus the standard symplectic structure. Let $\Psi : M \times \mathcal{O}_\mu^- \to \mathfrak{g}^*$ be the moment map for the diagonal action of G on $M \times \mathcal{O}_\mu^-$. Then*

$$\Psi^{-1}(0)/G \simeq \Phi^{-1}(\mu)/G_\mu$$

as symplectic manifolds.

Proof. Consider the map $\tau : \Phi^{-1}(\mu) \to \Psi^{-1}(0) \hookrightarrow M \times \mathcal{O}^-$ given by $\tau(m) = (m, \Phi(m))$. We are done if we can show that $\tau^*(\omega - \omega_{\mathcal{O}}) = \omega|_{\Phi^{-1}(\mu)}$. But $\tau^*\omega_{\mathcal{O}} = 0$ since Φ is constant on $\Phi^{-1}(\mu)$, and $\tau^*\omega = \omega|_{\Phi^{-1}(\mu)}$ by definition. □

The following theorem and its corollaries are very useful in computing the reduced spaces.

Theorem 135 (reduction in stages). *Let (M, ω) be a symplectic manifold with a proper Hamiltonian action of a product $G \times H$. Let $\Phi : M \to \mathfrak{g}^* \times \mathfrak{h}^*$, $\Phi(m) = (\Phi_1(m), \Phi_2(m)) \in \mathfrak{g}^* \times \mathfrak{h}^*$ be a corresponding moment map. Suppose $(\alpha, \beta) \in \mathfrak{g}^* \times \mathfrak{h}^*$ is a regular value of Φ.*

Then $M_\alpha := \Phi^{-1}(\alpha)/G_\alpha$ is a Hamiltonian H-space and $\Phi_2|_{\Phi^{-1}(\alpha)}$ descends to a moment map $\Psi : M_\alpha \to \mathfrak{h}^*$ for the action of H on M_α.

Moreover, β is a regular value of Ψ and

$$\Psi^{-1}(\beta)/H_\beta \simeq \Phi^{-1}(\alpha, \beta)/(G \times H)_{(\alpha,\beta)}$$

as symplectic manifolds.

Note that the hypothesis that the product $G \times H$ acts really amounts to saying that both G and H act and that the actions of G and H commute.

Proof. Since $G \times H$ is a product, the coadjoint actions of G on $\mathfrak{h}^*$ and of H on $\mathfrak{g}^*$ are trivial. Hence Φ_1 is H-invariant and Φ_2 is G-invariant.

It follows that the action of H on M preserves the level set $\Phi_1^{-1}(\alpha)$. Moreover, since it commutes with the action of G, it descends to an action on $M_\alpha := \Phi_1^{-1}(\alpha)/G_\alpha$. Also, the projection $\pi_\alpha : \Phi_1^{-1}(\alpha) \to M_\alpha$ is H-equivariant. It follows that for any $\xi \in \mathfrak{h}$, we have the following relation between the induced vector fields on M and on M_α: $d\pi_\alpha(\xi_M|_{\Phi_1^{-1}(\alpha)}) = \xi_{M_\alpha} \circ \pi_\alpha$. Also, since Φ_2 is G-invariant, the restriction $\Phi_2|_{\Phi_1^{-1}(\alpha)}$ descends to a map $\Psi : M_\alpha \to \mathfrak{h}^*$, which is H-equivariant.

To show that Ψ is a moment map for the action of H on M_α we need to show that $\iota(\xi_{M_\alpha})\omega_\alpha = d\langle \Psi, \xi \rangle$ where ω_α is the reduced symplectic form on M_α. We know that $d\pi_\alpha(X_{\langle \Phi_2, \xi \rangle}) = X_{\langle \Psi, \xi \rangle} \circ \pi_\alpha$ where $X_{\langle \Phi_2, \xi \rangle}$ and $X_{\langle \Psi, \xi \rangle}$ denote the Hamiltonian vector fields of the appropriate functions. Since $X_{\langle \Phi_2, \xi \rangle} = \xi_M$ we have

$$X_{\langle \Psi, \xi \rangle} \circ \pi_\alpha = d\pi_\alpha(X_{\langle \Phi_2, \xi \rangle}) = d\pi_\alpha(\xi_M) = \xi_{M_\alpha} \circ \pi_\alpha.$$

Therefore $X_{\langle \Psi, \xi \rangle} = \xi_{M_\alpha}$. Thus the action of H on M_α is Hamiltonian and $\Psi : M_\alpha \to \mathfrak{h}^*$ is a corresponding moment map (we have already checked that Ψ is equivariant).

Next we argue that if (α, β) is a regular value of Φ then β is a regular value of Ψ. Note first that by definition of Ψ, we have $\Psi^{-1}(\beta) = (\Phi_1^{-1}(\alpha) \cap \Phi_2^{-1}(\beta))/G_\alpha = \Phi^{-1}(\alpha, \beta)/G_\alpha$.

Thus to show that β is a regular value of Ψ it is enough to show that for any $x \in \Phi_1^{-1}(\alpha) \cap \Phi_2^{-1}(\beta)$ the level sets $\Phi_1^{-1}(\alpha)$ and $\Phi_2^{-1}(\beta)$ intersect transversely at x. Since $T_x\Phi_1^{-1}(\alpha) = \ker d(\Phi_1)_x$ it follows from part (2) of Corollary 128 that the tangent space $T_x\Phi_1^{-1}(\alpha)$ is the symplectic perpendicular to the tangent space $T_x(G \cdot x)$ to the G-orbit through x: $T_x\Phi_1^{-1}(\alpha) = (T_x(G \cdot x))^\omega$. A similar statement holds for the level set of Φ_2. Thus in order to prove that

$T_x\Phi_1^{-1}(\alpha) + T_x\Phi_2^{-1}(\beta) = T_xM$, it is enough to prove that $0 = \left(T_x\Phi_1^{-1}(\alpha) + T_x\Phi_2^{-1}(\beta)\right)^\omega = T_x(G\cdot x) \cap T_x(H \cdot x)$. Since (α, β) is a regular value of Φ, it follows from part (1) of Corollary 128 that $T_x((G\times H)\cdot x) \simeq \mathfrak{g}\times\mathfrak{h} \simeq (T_x(G\cdot x))\times(T_x(H\cdot x))$. Thus $T_x(G\cdot x)\cap T_x(H\cdot x) = \{0\}$. This proves that β is a regular value of Ψ.

Since $\Psi^{-1}(\beta) = \Phi^{-1}(\alpha,\beta)/G_\alpha$, the reduced spaces $(M_\alpha)_\beta := \Psi^{-1}(\beta)/H_\beta$ and $M_{(\alpha,\beta)} := \Phi^{-1}(\alpha,\beta)/(G_\alpha \times H_\beta)$ are equal as manifolds. It remains to show that they are symplecticly the same. Let $\pi_{(\alpha,\beta)} : \Phi^{-1}(\alpha,\beta) \to M_{(\alpha,\beta)}$ and $(\pi_\alpha)_\beta : \Psi^{-1}(\beta) \to (M_\alpha)_\beta$ denote the orbit maps. By definition, the reduced forms $\omega_\alpha \in \Omega^2(M_\alpha)$, $\omega_{(\alpha,\beta)} \in \Omega^2(M_{(\alpha,\beta)}$ and $(\omega_\alpha)_\beta \in \Omega^2((M_\alpha)_\beta)$ satisfy

$$\begin{aligned}(\pi_\alpha)^*\omega_\alpha =& \ \omega|_{\Phi_1^{-1}(\alpha)}\\ (\pi_{(\alpha,\beta)})^*\omega_{(\alpha,\beta)} =& \omega|_{\Phi^{-1}(\alpha,\beta)}\\ (\pi_\alpha)_\beta)^*(\omega_\alpha)_\beta =& \omega_\alpha|_{\Psi^{-1}(\beta)}\end{aligned}$$

Therefore $(\pi_\alpha)^*((\pi_\alpha)_\beta)^*((\omega_\alpha)_\beta) = (\pi_\alpha)^*(\omega_\alpha|_{\Psi^{-1}(\beta)}) = \omega|_{\Phi_1^{-1}(\alpha)\cap\Phi_2^{-1}(\beta)} = (\pi_{(\alpha,\beta)})^*\omega_{(\alpha,\beta)}$. Since pull-backs by submersions are injective, it follows that $\omega_{(\alpha,\beta)} = (\omega_\alpha)_\beta$. □

Corollary 136. *Let (M,ω) be a symplectic manifold with a Hamiltonian action of a product $G \times H$ (so that the actions of G and H on M commute). Let $\Phi : M \to \mathfrak{g}^* \times \mathfrak{h}^*$, $\Phi(m) = (\Phi_1(m), \Phi_2(m)) \in \mathfrak{g}^* \times \mathfrak{h}^*$ be a corresponding moment map. Suppose $(\alpha,\beta) \in \mathfrak{g}^* \times \mathfrak{h}^*$ is a regular value of Φ.*

The the symplectic quotient obtained by first reducing M by the action of G at α and then by the induced action of H at β is isomorphic to the symplectic quotient obtained by by first reducing M by the action of H at β and then by the induced action of G at α:

$$(M_\alpha)_\beta = (M_\beta)_\alpha.$$

Example 74. The standard action of the unitary group $U(n)$ on $(\mathbb{C}^n, \frac{i}{2}\sum dz_j \wedge d\bar{z}_j)$ is Hamiltonian. It commutes with the action of S^1 given by $e^{i\theta}\cdot z = e^{i\theta}z$, which is also Hamiltonian. A moment map for the action of S^1 is $\Psi(z) = \frac{1}{2}||z||^2$. Hence the reduced spaces for the action of S^1 are a point $\Psi^{-1}(0)/S^1$ and complex projective spaces $\mathbb{C}P^{n-1} = \{||z||^2 = c\}/S^1$ (if $c \neq 0$). It follows, in particular, that complex projective spaces are symplectic manifolds.

If K is a subgroup of $U(n)$ then its action on $\mathbb{C}^n$ is also Hamiltonian and it also commutes with the action of S^1. Hence it descends to a Hamiltonian action on $\mathbb{C}P^{n-1}$.

20. Lecture 20. Rigid body dynamics

We now consider examples of symmetric Hamiltonian systems. We start with a free rigid body in $\mathbb{R}^3$ constrained to pivot freely about a fixed point, say 0. We can think of a body as n particles with masses $m_1,\ldots,m_n$ and positions $x_1,\ldots,x_n$, $x_i \in \mathbb{R}^3$. Since the body is rigid the distances $||x_i - x_j||$ and $||x_i||$ are constant. These conditions give us constraints. We claim that the resulting configuration space is the Lie group $\mathrm{SO}(3) = \{A \in \mathbb{R}^{3^2} \mid AA^T = I, \det A > 0\}$ of orthogonal orientation preserving matrices. Indeed let $x_1^0,\ldots,x_n^0$ be the positions of the particles at time 0, the reference configuration. Then at another time $x_i = Ax_i^0$ for some matrix $A \in \mathrm{SO}(3)$.

To understand the dynamics, we need to determine the Lagrangian. Then using the Legendre transform we will compute the corresponding Hamiltonian. Since the body is free and rigid, there is no potential energy to worry about, it is all taken care of by the constraints. Thus we need to compute the restriction of the kinetic energy of the system of free particles to our constraint set. Let $A(t)$ be a path in $\mathrm{SO}(3)$. Then $x_i(t) = A(t)x_i^0$ and $\dot{x}_i(t) = \dot{A}(t)x_i^0$. Consequently the kinetic energy of one particle is

$$\frac{1}{2}m_i\|\dot{x}_i\|^2 = \frac{1}{2}m_i(\dot{A}(t)x_i^0, \dot{A}(t)x_i^0).$$

Therefore the kinetic energy of the whole body is $KE(A, \dot{A}) = \frac{1}{2}\sum_i m_i(\dot{A}x_i^0, \dot{A}x_i^0)$, where $\dot{A} \in T_A\,\mathrm{SO}(3)$ is a tangent vector.

Note that $T_A\,\mathrm{SO}(3) = \{X \in \mathbb{R}^{3^2} \mid XA^T + AX = 0\}$. This is because if $A(t)$ is a curve in $\mathrm{SO}(3)$ with $A(0) = A$ then $I = A(t)A(t)^T$. Therefore $0 = \frac{d}{dt}\big|_0(A(t)A(t)^T) = \dot{A}(0)A(0)^T + A(0)\dot{A}(0)^T$.

The quadratic form $KE(A, X) := \frac{1}{2}\sum_i m_i(Xx_i^0, Xx_i^0)$ for $X \in T_A\,\mathrm{SO}(3)$ defines a metric $\mathbb{I}$ in $\mathrm{SO}(3)$:

$$\mathbb{I}(A)(X, Y) = \sum m_i(Xx_i^0, Yx_i^0) \quad \text{for } X, Y \in T_A\,\mathrm{SO}(3).$$

We claim that this metric is left invariant: $(L_A^*\mathbb{I})(B) = \mathbb{I}(B)$ for all $A, B \in \mathrm{SO}(3)$. To check the claim we need to compute the differential of left multiplication dL_A. Let $\gamma(t)$ be a curve in $\mathrm{SO}(3)$ with $\gamma(0) = A$ and $\dot{\gamma}(0) = X$. Then $dL_A(X) = \frac{d}{dt}\big|_0 L_A(\gamma(t)) = \frac{d}{dt}\big|_0 A\gamma(t) = A\dot{\gamma(0)} = AX$. Therefore $dL_A(X) = AX$.

We now compute: for any $X, Y \in T_B\,\mathrm{SO}(3)$ we have $(L_A^*\mathbb{I})(B)(X, Y) = \mathbb{I}(L_AB)(dL_AX, dL_AY) = \mathbb{I}(AB)(AX, AY) = \sum_i m_i(AXx_i^0, AYx_i^0) = \sum_i m_i(Xx_i^0, Yx_i^0) = \mathbb{I}(B)(X, Y)$ since $A \in \mathrm{SO}(3)$.

We conclude that if we have a rigid body, which is free to rotate about a fixed point, then the corresponding Hamiltonian system is $T^*\,\mathrm{SO}(3)$ with a Hamiltonian $h \in C^\infty(T^*\,\mathrm{SO}(3))$ of the form $h(A, \eta) = \frac{1}{2}\mathbb{I}^*(A)(\eta, \eta)$ where $\mathbb{I}^*$ is a metric on $T^*\,\mathrm{SO}(3)$ dual to a left-invariant metric on $\mathrm{SO}(3)$.

Homework Problem 25. Suppose we have an action $\tau : G \longrightarrow \mathrm{Diff}(N)$ of a Lie group G on a manifold N. Let $\widetilde{\tau} : G \longrightarrow \mathrm{Diff}(T^*N)$ be the lifted action. Let g be a metric on N such that $\tau_a^*g = g$, for all $a \in G$, i.e., let g be a G-invariant metric. Let $h(a, \eta) = \frac{1}{2}g^*(a)(\eta, \eta)$ be the Hamiltonian defined by the dual metric g^*. Then h is G-invariant, i.e. $(\widetilde{\tau_a})^*h = h$, for all $a \in G$.

All left-invariant metrics on a Lie group G are determined by their value on the tangent space at the identity: if g is left-invariant, $g(a)(v, w) = g(1)(dL_{a^{-1}}v, dL_{a^{-1}}w)$, where $d(L_{a^{-1}}) : T_aG \longrightarrow T_1G$. In other words, left invariant metrics are classified by the inner products on the Lie algebra $\mathfrak{g}$.

In the case of $\mathrm{SO}(n)$, the special orthogonal group, the Lie algebra $\mathfrak{so}(n) = \{X \in \mathbb{R}^{n^2} \mid X + X^T = 0\}$. There is a standard inner product on $\mathfrak{so}(n)$:

$$(X, Y)_{\mathrm{st}} = -\frac{1}{2}\mathrm{tr}XY.$$

Clearly it is symmetric. To check positivity we compute: $(X, X)_{\mathrm{st}} = -\frac{1}{2}\mathrm{tr}X^2 = -\frac{1}{2}\sum_i \sum_j X_{ij}X_{ji} = \frac{1}{2}\sum_i \sum_j X_{ij}X_{ij}$ (since $X = -X^T$). So $(X, X)_{\mathrm{st}} = \frac{1}{2}\sum X_{ij}^2 \geq 0$.

Note that this standard inner product is invariant under the adjoint action. Indeed $Ad(A)X = AXA^{-1}$ and $(Ad(A)X, Ad(A)Y)_{\mathrm{standard}} = -\frac{1}{2}\mathrm{tr}AXA^{-1}AYA^{-1} = -\frac{1}{2}\mathrm{tr}XYA^{-1}A = -\frac{1}{2}\mathrm{tr}XY = (X, Y)_{\mathrm{standard}}$. So the action of $\mathrm{SO}(n)$ on $\mathfrak{so}(n)$ is orthogonal.

With the standard inner product, the matrices

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

form an orthonormal basis of $\mathfrak{so}(3)$. This choice of a basis identifies $(\mathfrak{so}(3), (\cdot,\cdot)_{\mathrm{st}})$ with $(\mathbb{R}^3, (\cdot,\cdot))$. Under the identification, the adjoint action becomes the standard action of $\mathrm{SO}(3)$ on $\mathbb{R}^3$. Explicitly the identification is given by

$$\begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix} \mapsto (w_1, w_2, w_3)$$

Note that

$$\begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{w} \times \vec{x}$$

Lemma 137. *Let G be a Lie group. Let σ be a left-invariant metric on G. Then for $b \in G$ we have $(R_{b^{-1}})^*\sigma = \sigma$ if and only if $(Ad(b))^*\sigma(1) = \sigma(1)$. Here $R_{b^{-1}}g := gb^{-1}$.*

Proof. The proof is a computation. By definition $((R^*_{b^{-1}}\sigma)(g)(v, w) = \sigma(R_{b^{-1}}g)(dR_{b^{-1}}v, dR_{b^{-1}}w)$. The metric σ is left invariant if and only if $\sigma(g)(v, w) = \sigma(1)(dL_{g^{-1}}v, dL_{g^{-1}}w)$ for all $g \in G$. Hence $((R^*_{b^{-1}}\sigma)(g)(v, w) = \sigma(1)(dL_{(gb^{-1})^{-1}}dR_{b^{-1}}v, dL_{(gb^{-1})^{-1}}dR_{b^{-1}}w)$. Now $dL_{(gb^{-1})^{-1}}dR_{b^{-1}} = d(L_{bg^{-1}}R_{b^{-1}}) = d(L_bL_{g^{-1}}R_{b^{-1}}) = d(L_bR_{b^{-1}}L_{g^{-1}}) = dL_bR_{b^{-1}}dL_{g^{-1}}$. Therefore $((R_{b^{-1}})^*\sigma)(g)(v, w) = \sigma(1)(Ad(b)dL_{g^{-1}}v, Ad(b)dL_{g^{-1}}w)$. So if $Ad(b)^*(\sigma(1)) = \sigma(1)$, we get $((R^*_{b^{-1}}\sigma)(g)(v, w) = \sigma(1)(DL_{g^{-1}}v, DL_{g^{-1}}w) = \sigma(g)(v, w)$.

Conversely, if $(R_{b^{-1}})^*\sigma = \sigma$, then

$$\sigma(1)(Ad(b)dL_{g^{-1}}v, Ad(b)dL_{g^{-1}}w) = \sigma(1)(dL_{g^{-1}}v, dL_{g^{-1}}w)$$

for all g, for all $v, w \in T_gG$. So $\sigma(1)(Ad(b)\cdot, Ad(b)\cdot) = \sigma(1)(\cdot,\cdot)$ or $(Ad(b))^*\sigma(1) = \sigma(1)$. $\square$

It follows that if g is a left-invariant metric on $\mathrm{SO}(3)$ and the inner product $g(1)$ on $T_1\,\mathrm{SO}(3)$ is preserved by the adjoint action of a subgroup K of $\mathrm{SO}(3)$, then the Hamiltonian system $(T^*\,\mathrm{SO}(3), h(q, p) = \frac{1}{2}g^*(q)(p, p))$ has $\mathrm{SO}(3) \times K$ as a symmetry group.

Let us now take a close look at the inner products on $\mathfrak{so}(3)$. Let $\mathbb{I}$ be such an inner product. The map $\mathfrak{I} : \mathfrak{so}(3) \to \mathfrak{so}(3)$ defined by $(X, \mathfrak{I}Y)_{\mathrm{st}} = \mathbb{I}(X, Y)$ is uniquely determined by the inner product $\mathbb{I}$. It is symmetric with respect to $(\cdot,\cdot)_{\mathrm{st}}$: $(X, \mathfrak{I}Y)_{\mathrm{st}} = (\mathfrak{I}X, Y)_{\mathrm{st}}$. Therefore $\mathfrak{I}$ has an orthonormal basis of eigenvectors with positive eigenvalues I_1, I_2 and I_3. These eigenvalues are called **moments of inertia** of a rigid body.

Exercise 12. $Ad(A)^*\mathbb{I} = \mathbb{I}$ if and only if $\mathfrak{I}Ad(A) = Ad(A)\mathfrak{I}$.

Therefore, if all moments of inertia are equal, the inner product $\mathbb{I}$ is SO(3)-invariant. Hence the corresponding Hamiltonian system on $T^*\,\mathrm{SO}(3)$ has $\mathrm{SO}(3) \times \mathrm{SO}(3)$ as a symmetry group. If only two moments of inertia are equal, then the inner product is invariant under the action of SO(2). Therefore the corresponding Hamiltonian system on $T^*\,\mathrm{SO}(3)$ has $\mathrm{SO}(3) \times \mathrm{SO}(2)$ as a symmetry group.

Let us now suppose that there is gravity. Such a system is called a heavy top. The Lagrangian (hence the Hamiltonian) then has a potential term. Let us compute it. Let e_3 denote the unit vertical vector. Let $m = \sum m_i$. Let $x^0 = \frac{1}{m}\sum m_i x_i^0$; it is the center of mass of the rigid body. Generically the center of mass is not at the origin. This is going to be our blanket assumption. The potential $V(A) = \sum g m_i (A x_i^0, e_3) = mg(A(\frac{\sum m_i x_i^0}{m}), e_3) = mg(Ax^0, e_3)$, where g denotes the gravitational acceleration 9.8 m/s^2 (and not a metric!). The corresponding Hamiltonian is then of the form $h(A, p_A) = \frac{1}{2}\mathbb{I}^*(A)(p_A, p_A) + V(A)$ where $p_A \in T^*_A\,\mathrm{SO}(3)$, $\mathbb{I}^*$ is the metric on $T^*\,\mathrm{SO}(3)$ dual to a left invariant metric $\mathbb{I}$ on SO(3) and the potential $V(A)$ is as above.

Let us now examine the symmetries of a heavy top. We know that the metric $\mathbb{I}$ is invariant under the left multiplication by SO(3). What about the potential? For $B \in \mathrm{SO}(3)$ we have $V(BA) = mg(BAx^0, e_3) = mg(Ax^0, B^{-1}e_3)$. Thus the isotropy group H of e_3:

$$H = \{B \in \mathrm{SO}(3) \,|\, Be_3 = e_3\} = \{\begin{pmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & 1 \end{pmatrix}\}$$

preserves the potential V. The group H is (isomorphic to) SO(2).

Let $K = \{B \in \mathrm{SO}(3) \,|\, Bx^0 = x^0\}$. The group K is also isomorphic to SO(2). For any $B \in K$ we have $V(AB) = mg(ABx^0, e_3) = mg(Ax^0, e_3)$. So potential energy is also right-SO(2) invariant as well. Thus generically the potential has $\mathrm{SO}(2) \times \mathrm{SO}(2)$ worth of symmetries.

We now consider several cases. If all moments of inertia are equal, then the kinetic energy is left and right SO(3) invariant. Hence the symmetries of the Lagrangian are exactly the symmetries of the potential V. Therefore the system has $\mathrm{SO}(2) \times \mathrm{SO}(2)$ worth of symmetries.

Suppose now that only two moments of inertia are equal. Then the metric on SO(3) is invariant under the group $\mathrm{SO}(3) \times L$, where L is isomorphic to SO(2). Recall that the first factor acts by left multiplication and the second by right multiplication. If the group L also fixes the center of mass x^0, then the whole system has $\mathrm{SO}(2) \times L = \mathrm{SO}(2) \times \mathrm{SO}(2)$ worth of symmetries. This kind of heavy top is known as the **Lagrange top**.

Finally, if all moments of inertia are different, the system only has SO(2) as a connected symmetry group (there may also be some discrete symmetries).

21. Lecture 21. Relative equilibria

Suppose we want to understand a dynamical system, which for the purposes of the present discussion we can think of as the flow $\{\varphi_t\}$ of a vector field X on a manifold M. We would then start by trying to find two types of orbits of the flow:

1. look for **fixed points** of the flow φ_t, i.e., look for points $m \in M$ such that $\varphi_t(m) = m$ for all t;
2. look for **periodic orbits** of the flow φ_t, i.e., look for $m \in M$ such that $\varphi_T(m) = m$ for some $T > 0$ (so that $\varphi_{T+t}(m) = \varphi_t(m)$ for all t).

Note that m is a fixed point of the flow if and only if $X(m) = 0$ (so fixed points are zero of the vector field X). Having found fixed points and periodic orbits, we would then study the flow in the neighborhood of fixed points and of periodic orbits.

Suppose now we have a symmetric Hamiltonian system $(M, \omega, G, \Phi : M \longrightarrow \mathfrak{g}^*, h \in C^\infty(M)^G)$, where M is a manifold, ω is a symplectic form on M, G is a Lie group acting on M in a Hamiltonian manner, $\Phi : M \longrightarrow \mathfrak{g}^*$ is a corresponding moment map and $h \in C^\infty(M)^G$ is a G-invariant Hamiltonian on M. If $(M_\mu, \omega_\mu, h_\mu)$ is a reduced Hamiltonian system, then $\dim M_\mu = \dim \Phi^{-1}(\mu) - \dim(G_\mu) = \dim M - \dim G - \dim G_\mu$ (because μ is a regular value). Therefore the reduced system should be easier to understand than the original system on M.

Suppose a point $\bar{m} \in M_\mu$ is a zero of of the Hamiltonian vector field X_{h_μ} of h_μ. If $m \in \Phi^{-1}(\mu)$ is a point such that $\pi(m) = \bar{m}$ (where $\pi : \Phi^{-1}(\mu) \to M_\mu$ is the orbit map), then $d\pi(X_h(m)) = X_{h_\mu}(\bar{m}) = 0$ since the Hamiltonian vector fields of h and of h_μ are π-related. Consequently the vector $X_h(m)$ is tangent to the orbit $G_\mu \cdot m$. In other words, $X_h(m) = \xi_M(m)$ for some vector ξ in the isotropy Lie algebra $\mathfrak{g}_\mu$.

We claim that then $\varphi_t(m) = (\exp t\xi) \cdot m$ for all t, where $\{\varphi_t\}$ denotes the flow of the Hamiltonian vector field X_h of h. Indeed since the Hamiltonian h is G-invariant, its Hamiltonian vector field is G-invariant as well, so the flow φ_t is G-equivariant. Consequently $d\varphi_t(\xi_M(m)) = \xi_M(\varphi_t(m))$. On the other hand, $d\varphi_t(X_h(m)) = X_h(\varphi_t(m))$ by definition, since φ_t is the flow of X_h. Since $d\varphi_t$ is injective,

$$\xi_M(\varphi_t(m)) = X_h(\varphi_t(m)). \tag{24}$$

By equation (24), the curve $t \mapsto \varphi_t(m)$ is an integral curve of the induced vector field ξ_M. Since $t \mapsto (\exp t\xi) \cdot m$ is also an integral curve of ξ_M, it follows that

$$\varphi_t(m) = (\exp t\xi) \cdot m,$$

which proves the claim.

Definition 138. Let (M, ω) be a symplectic manifold with a symplectic action of a Lie group G. A point $m \in M$ is a **relative equilibrium** of an invariant Hamiltonian $h \in C^\infty(M)^G$ if the vector $X_h(m)$ is tangent to the G_μ orbit of m, where G_μ is the isotropy group of $\mu = \Phi(m)$.

Thus if $(M, \omega, G, \Phi : M \longrightarrow \mathfrak{g}^*, h \in C^\infty(M)^G)$ is a symmetric Hamiltonian system, a point $m \in M$ is a relative equilibrium if and only if any one of the following conditions holds.

1. $X_h(m) = \xi_M(m)$ for some $\xi \in \mathfrak{g}$.

2. $d(\langle\Phi, \xi\rangle - h\rangle)(m) = 0$ for some $\xi \in \mathfrak{g}$.
3. $\bar{m} \in M_\mu$ is an equilibrium of the reduced Hamiltonian system $(M_\mu, \omega_\mu, h_\mu)$ where $\bar{m}$ is the image of m under the projection $\pi : \Phi^{-1}(\mu) \to M_\mu$ and $\mu = \Phi(m)$.
4. $t \mapsto (\exp t\xi) \cdot m$ is an integral curve of X_h.
5. $X_h(m) \in T_m(G \cdot m)$.

Rigid body rotating freely about a fixed point. Recall that the phase space for a rigid body rotating freely about a fixed point is $M = T^* \mathrm{SO}(3)$ with the standard symplectic structure. The motion is governed by a Hamiltonian of the form $h(q,p) = \frac{1}{2}\|dL_q^T p\|^2$, $p \in T_q^* \mathrm{SO}(3)$, where $\| \;\|$ is an inner product on the Lie algebra $\mathfrak{so}(3)^* \simeq \mathbb{R}^3$. The symmetry group is $G = \mathrm{SO}(3)$, and the action of G on M is the lift of left multiplication. Recall, that left multiplication induces right invariant vector fields: for every $\xi \in \mathfrak{so}(3)$ the corresponding vector field is given by $\xi_{\mathrm{SO}(3)}(q) = \frac{d}{dt}\big|_0((\exp t\xi)q) = dR_q(\xi)$. Consequently the moment map $\Phi : T^* \mathrm{SO}(3) \longrightarrow \mathfrak{so}(3)^*$ is determined by $\langle\Phi(q,p), \xi\rangle = \langle p, \xi_{\mathrm{SO}(3)}(q)\rangle = \langle p, dR_q(\xi)\rangle = \langle dR_q^T p, \xi\rangle$. Hence $\Phi(q,p) = dR_q^T p$. Note that the restriction of the moment map to each fiber of the cotangent bundle $\Phi : T_q^* \mathrm{SO}(3) \longrightarrow \mathfrak{so}(3)^*$ is an isomorphism, so all values of Φ are regular. Note also that $\mathrm{SO}(3)$ acts freely on itself, hence on $T^* \mathrm{SO}(3)$.

Let us compute the reduced system at $\mu \in \mathfrak{so}(3)^*$. We have $\Phi^{-1}(\mu) = \{(q,p) : dR_q^T p = \mu\} = \{(q, (dR_q^T)^{-1}\mu) : q \in \mathrm{SO}(3)\} \simeq \mathrm{SO}(3)$. The reduced space $\Phi^{-1}(\mu)/G_\mu$ is diffeomorphic to the coadjoint orbit $Ad^\dagger(G)\mu$: the diffeomorphism is given by

$$[(q, dR_{q^{-1}}^T \mu)] \mapsto Ad^\dagger(q^{-1})\mu. \tag{25}$$

Note that this argument works for any Lie group G and not just for $\mathrm{SO}(3)$. Note also that we haven't computed the reduced symplectic form. What is it? Hint: reduction in stages — the left and right multiplications commute.

To compute the reduced Hamiltonian we need to compute the restriction $h\big|_{\Phi^{-1}(\mu)}$. If $(q,p) \in \Phi^{-1}(\mu)$ then $h(q,p) = \frac{1}{2}\|dL_q^T p\|^2 = \frac{1}{2}\|dL_q^T (dR_{q^{-1}}^T)^{-1}\mu\|^2 = \frac{1}{2}\|d(L_q R_{q^{-1}})^T \mu\|^2 = \frac{1}{2}\|Ad^\dagger(q^{-1})\mu\|^2$. Hence under the identification (25) of the reduced space with the coadjoint orbit, $h_\mu : Ad^\dagger(G)\mu \longrightarrow \mathbb{R}$ is simply $h_\mu(\eta) = \frac{1}{2}\|\eta\|^2$.

Recall that the coadjoint orbits of $\mathrm{SO}(3)$ are two-spheres and the origin. Recall also that there exists an identification of $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ such that under this identification $\|\eta\|^2 = I_1\eta_1^2 + I_2\eta_2^2 + I_3\eta_3^2$ where I_1, I_2, I_3 are moments of inertia.

Thus to understand the reduced dynamics of a free rigid body it is enough to understand a Hamiltonian system on $S^2 = \{\eta \in \mathbb{R}^3 \mid \eta_1^2 + \eta_2^2 + \eta_3^2 = c^2\}$ with a Hamiltonian of the form $h_\mu(\eta) = \frac{1}{2}(I_1\eta_1^2 + I_2\eta_2^2 + I_3\eta_3^2)$, $I_j > 0$. We know that the flow of the Hamiltonian vector field X_{h_μ} preserves h_μ. Thus to understand the dynamics it is enough to understand level sets of h_μ, which are the intersections of a sphere with ellipsoids. In particular, if all moments of inertia are different, we get six equilibria corresponding to the intersections of the axes of the ellipse with the sphere. Two equilibria are the maximal points of h_μ, two are the minimal points and two are the saddle points.

Definition 139. Let $\varphi_t : M \longrightarrow M$ be a flow on a manifold M. Suppose a point $m_0 \in M$ a fixed point: $\varphi_t(m_0) = m_0$for all t. The point m_0 is **Liapunov stable** if for any neighborhood U of m_0, there exists a neighborhood V of m_0, $V \subset U$, such that for all$m \in V$ we have $varphi_t(m) \in U$ for all t.

Lemma 140. *Suppose $\varphi_t : M \longrightarrow M$ is a flow on a manifold M. Suppose there exists a smooth function f defined on M near m_0 such that*

1. *$f(\varphi_t(m)) = f(m)$,*
2. *$df_{m_0} = 0$,*
3. *The quadratic form $d^2 f_{m_0}$ is positive definite.*

Then m_0 is Liapunov stable.

Proof. Since the claim is local, there is no loss of generality in assuming that $M = \mathbb{R}^n$ and that $m_0 = 0$. By Morse lemma there exist coordinates$x_1, \dots, x_k$ such that $f(x) = f(0) + \sum a_{ij} x_i^2$ where $a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(0)$. Since the Hessian $d^2 f(0) = (a_{ij})$ is positive definite by assumption, it follows that the level sets of f near $f(0)$ are compact. Since f is constant along the flow, the stability follows. □

It follows from the lemma above that maxima and minima of h_μ are stable. It is also easy to see that the saddle points corresponding to the intermediate axis are unstable.

We will see later that in good cases stability of an equilibrium in the reduced space implies **relative stability** of a corresponding relative equilibrium in the original system, where relative stability is defined as follows:

Definition 141. Let M be a manifold with an action of a Lie group G, and let $\varphi_t : M \longrightarrow M$ be a G-equivariant flow on M. Then φ_t descends to a flow $\bar{\varphi}_t$ on the quotient M/G. A **relative equilibrium**j of φ_t is a x such that the curve $x \mapsto \varphi_t(x)$ projects down to a constant curve in M/G. A relative equilibrium $t \mapsto \varphi_t(x)$ is **G-stable** if for every G-invariant neighborhood V in M of the relative equilibrium, there exists a G-invariant neighborhood U such that for every $y \in V$ we have $\varphi_t(y) \in U$ for all t.

Spherical pendulum. A spherical pendulum is a Hamiltonian system on the cotangent bundle of the two-sphere T^*S^2 where T^*S^2 is given the standard symplectic form. Recall that the standard metric on $\mathbb{R}^3$ allows us to embed T^*S^2 symplecticly into $T^*\mathbb{R}^3$: $T^*S^2 \simeq TS^2 \hookrightarrow T\mathbb{R}^3 \simeq T^*\mathbb{R}^3$. Hence

$$T^*S^2 = \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|x\|^2 = 1, x \cdot y = 0\}.$$

With this identification the Hamiltonian h is given by

$$h(x, y) = \frac{1}{2}\|y\|^2 + x_3,$$

where x_3 is the third coordinate of $x \in \mathbb{R}^3$ (we have assumed that all relevant physical constants, such as mass, length of the pendulum and gravitational acceleration, are equal to 1).

The group S^1 acts on S^2 be rotations about the x_3 axis, preserving the metric and the potential $V(x) = x_3$. This lifts to an action on T^*S^2 preserving the Hamiltonian h.

Let us start the study of the spherical pendulum by looking for equilibria of the Hamiltonian flow of h on T^*S^2. Let us first consider a more general situation: a classical mechanical system on a manifold B (cf. Definition 61. That is suppose (B, g) is a Riemannian manifold. Let $M = T^*B$ be the cotangent bundle of B with the standard symplectic form and let $h(q, p) = \frac{1}{2}g^*(q)(p, p) + V(q)$ be a Hamiltonian where $p \in T_q^*B$, the potential V is a smooth function on B and g^* is the dual metric on T^*B. To find equilibria of the Hamiltonian system $(T^*B, \omega = \omega_{T^*B}, h)$ we need to find the zeros of the Hamiltonian vector field X_h of h. Since the symplectic form is nondegenerate, $X_h(q, p) = 0$ if and only if $dh(q, p) = 0$. Now if $(q, p) \in T^*B$ is a critical point of h, it is a critical point of the restriction of h to the fiber T_q^*B. But the restriction $h|_{T_q^*B}$ is the quadratic form $\frac{1}{2}g^*(q)(\cdot, \cdot)$ plus the constant $V(q)$. So if $(q, p) \in T^*B$ is a critical point of h, then p must be zero. And conversely, the points of the zero section are critical for the kinetic part of the Hamiltonian : $(q, p) \mapsto \frac{1}{2}g^*(q)(p, p)$. Therefore the critical points of h are the points of the form $(q, 0)$ where $dV(q) = 0$. In the case of spherical pendulum $V(x) = x_3 : S^2 \longrightarrow \mathbb{R}$. It follows that the critical points of V are north and south pole $(0, 0, 1)$ and $(0, 0, -1)$ respectively. We conclude that the spherical pendulum Hamiltonian has exactly two equilibria corresponding to the pendulum hanging straight down and the pendulum being straight up. It is not hard to check using Lemma 140 that the straight down position is stable. The straight up position is unstable (see if you can produce a trajectory that leaves any sufficiently small neighborhood of the north pole in T^*S^2).

Note that the two equilibria are fixed points of the S^1 action on T^*S^2. In fact they are the only fixed points. Indeed if a group G acts on a manifold B and by a lifted action on T^*B, then $(q, p) \in T^*B$ is fixed by the action of G only if $q \in B$ is fixed. If q is fixed by G, then the lifted action of G sends the fiber T_q^*B to itself. In fact we get a representation of G on the vector space T_q^*B. Thus a lifted action of G on T^*B has fixed points only if the action of G on B has fixed points and the corresponding representation of G on the fibers above the fixed points in B has fixed vectors. In the case of S^1 acting on S^2 by rotations about the x_3 axis, the (co)tangent spaces at the north and south poles are planes perpendicular to the vertical axis, and S^1 acts on these planes by rotations. It follows that the north and south pole are the only fixed points.

By Corollary 128 the north and south poles are the only critical points of the moment map J for the lifted action of S^1 on S^2. We can check this by computing the moment map $J : T^*S^2 \to \mathrm{Lie}(S^1)^* \simeq \mathbb{R}$ explicitly (we identify the dual of the Lie algebra with $\mathbb{R}$ by choosing a basis of the Lie algebra of S^1, say $\frac{\partial}{\partial \theta}$). Recall that we identified T^*S^2 with a subset of $\mathbb{R}^3 \times \mathbb{R}^3$. Since the action is a lifted action, $J(x, y) = \langle y, \xi(x)\rangle = y \cdot \xi(x)$ where ξ is the vector field induced by the action of S^1 on S^2. Now $\xi(x) = -x_2 \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2}$. Hence $J(x, y) = -x_2 y_1 + x_1 y_2$. One can check that the points $(0, 0, \pm 1, 0, 0, 0) \in \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|x\|^2 = 1, x \cdot y = 0\}$ are indeed the critical points of J.

Let us compute the reduced spaces at regular values of J. Note that $J(0, 0, \pm 1, 0, 0, 0) = 0$, so zero is the only singular value of J. Therefore we should be able to carry out reduction at any $c \neq 0$. Since $(0,0,\pm 1) \notin J^{-1}(c)$ for $c \neq 0$ we may use cylindrical coordinates (without losing any points):

$$\begin{aligned} x_1 &= \sqrt{1-z^2}\cos\theta, \\ x_2 &= \sqrt{1-z^2}\sin\theta, \\ x_3 &= z \end{aligned}$$

The action of S^1 in these coordinates is given by $e^{i\tau} \cdot (z,\theta) = (z, \tau + \theta)$. Hence the induced vector field $\xi(\theta, z)$ is $\frac{\partial}{\partial\theta}$. The canonical symplectic form ω is $dz \wedge dp_z + d\theta \wedge dp_\theta$ where p_z, p_θ denote the coordinates on the cotangent bundle corresponding to (z,θ). The moment map in these coordinates becomes $J(\theta, z, p_\theta, p_z) = \langle p_\theta d\theta + p_z dz, \frac{\partial}{\partial\theta}\rangle = p_\theta$. Consequently the level set of the moment map $J^{-1}(c) = \{(\theta, z, c, p_z)\}$. Hence the reduced space $(T^*S^2)_c = J^{-1}(c)/S^1 \simeq \{(z, p_z)\} \simeq T^*(-1,1)$. Note that $J^{-1}(c) = S^1 \times T^*(-1,1)$.

What is the reduced form ω_c? We have $\omega\big|_{J^{-1}(c)} = dz \wedge dp_z$. Therefore $\omega_c = dz \wedge dp_z$, which is the standard symplectic form on $T^*(-1,1)$. We conclude that

$$(J^{-1}(c)/S^1, \omega_c) = (T^*(-1,1), dz \wedge dp_z).$$

Let us now compute the reduced Hamiltonian $h_c(z,p_z)$. The metric g in cylindrical coordinates is $(dx_1)^2 + (dx_2)^2 + (dx_3)^2 = [d(\sqrt{1-z^2}\cos\theta)]^2 + [d(\sqrt{1-z^2}\sin\theta)]^2 + dz^2 = \frac{dz^2}{1-z^2} + (1-z^2)d\theta^2$. Hence $g^*(z,\theta)((p_z,p_\theta),(p_z,p_\theta)) = (\frac{1}{1-z^2}p_\theta^2 + (1-z^2)p_z^2)$ and the full Hamiltonian is give by $h(\theta, z, p_\theta, p_z) = \frac{1}{2}(\frac{1}{1-z^2}p_\theta^2 + (1-z^2)p_z^2) + z$. Consequently $h\big|_{J^{-1}(c)} = \frac{1}{2}(1-z^2)p_z^2 + (\frac{c^2}{2(1-z^2)} + z)$. We conclude that the reduced Hamiltonian h_c is given by

$$h_c(z,p_z) = \frac{1}{2}(1-z^2)p_z^2 + (\frac{c^2}{2(1-z^2)} + z).$$

Note that h_c is again of the form kinetic + potential where the potential term is $V_{\text{eff}}(z) = \frac{c^2}{2(1-z^2)} + z$. It is called the **effective** potential.

It follows from the discussion above that the critical points of the reduced Hamiltonian h_c are points of the form (z, p_z) where z is a critical point of the effective potential. Now

$$V'_{\text{eff}}(z) = -\frac{c^2}{2}\frac{-2z}{(1-z^2)^2} + 1 = \frac{c^2 z + (1-z^2)^2}{(1-z^2)^2}.$$

It follows that for all $c \neq 0$, there exists a unique critical point z_{crit} of V'_{eff}. Note that $-1 < z_{\text{crit}} < 0$. Hence for all $c \neq 0$, we have relative equilibria which, when projected to S^2 from T^*S^2 are S^1 orbits, which are horizontal circles.

Since the reduced spaces are two dimensional and since the Hamiltonian itself is a constant of motion, the trajectories of the reduced systems lie on the level curves of the reduced Hamiltonians. If $h_c(z,p_z) = \text{const}$ then $p_z^2 = \frac{2(\text{const} - V_{\text{eff}})}{1-z^2}$. This is a simple closed curve in the z-p_z plane. Since $J^{-1}(c) = S^1 \times T^*(-1,1)$, it follows that $(h,J)^{-1}(\text{const}, c)$ is a 2-torus. We will see that the reason for the motion of the spherical pendulum to be confined to tori and circles is the fact that the spherical pendulum is a completely integrable system.

It remains to analyze the motion on the level set $J^{-1}(0)$. We leave it as an exercise. Hint: delete the north and south pole. The action of S^1 is free on the remaining set.

22. Lecture 22. Lagrange top

Recall from Lecture 20 that a Lagrange top is a classical Hamiltonian system on $T^*\,\mathrm{SO}(3)$ with the standard symplectic form ω and with a Hamiltonian of the form $h(A, p_A) = \frac{1}{2}\mathbb{I}^*(p_A, p_A) + mg(Ax^0, e_3)$ where x^0 is a vector in $\mathbb{R}^3$, e_3 is the unit vector along the vertical axis, m and g are constants, and $\mathbb{I}^*$ is a left invariant metric on $T^*\,\mathrm{SO}(3)$ which is also preserved by the right action of $K = \{B \in \mathrm{SO}(3) \mid Bx^0 = x^0\}$. Without loss of generality we may assume that $x^0 = e_3$ and that $mg = 1$. Thus the system $(T^*\,\mathrm{SO}(3), \omega, h)$ has $\mathrm{SO}(2) \times \mathrm{SO}(2)$ worth of symmetries (cf. discussion at the end of Lecture 20).

Our strategy is to reduce by one of the two $\mathrm{SO}(2)$ and then study the resulting system with a remaining $\mathrm{SO}(2)$ symmetry. Note that the action $\mathrm{SO}(2)$ on $\mathrm{SO}(3)$ is free and that the quotient space is S^2. In other words, $\mathrm{SO}(2) \to \mathrm{SO}(3) \to S^2$ is a principal $\mathrm{SO}(2) = S^1$ bundle. For this reason we wish to study the reductions of a cotangent bundle of principal S^1-bundle.

Connections and curvature for principal S^1-bundles.

Definition 142. Let $S^1 \longrightarrow P \xrightarrow{\pi} B$ be a principal S^1-bundle. Let ξ denote the induced vector field: $\xi(p) = \frac{d}{d\theta}\big|_0(e^{i\theta} \cdot p)$. A **connection 1-form** A on P is an S^1-invariant 1-form such that $A(\xi) = 1$.

It is not hard to show that connection 1-forms always exist. For example a connection 1-form A can be constructed out of an S^1-invariant metric $\bar{g}$:

$$A(p)(v) = \frac{\bar{g}(p)(\xi_P(p), v)}{\bar{g}(p)(\xi_P(p), \xi_P(p))} \tag{26}$$

for all $v \in T_pP$.

In turn, an S^1 invariant metric $\bar{g}$ on a principal bundle P can be manufactured out of an arbitrary metric g on P by averaging:

$$\bar{g}(p)(v, w) := \frac{1}{2\pi}\int_0^{2\pi} ((e^{i\theta})^* g)(p)(v, w)\, d\theta$$

where $(e^{i\theta})^*$ denotes, by abuse of notation, the pull-back by the diffeomorphism of P defined by $e^{i\theta} \in S^1$.

Proposition 143. *Let A be a connection 1-form on a principal S^1 bundle $\pi : P \to B$. The two-form dA is basic.*

Proof. Since A is S^1-invariant, dA is also invariant. Moreover, since A is S^1 invariant, the Lie derivative of A with respect to the induced vector field ξ is zero. But $L_\xi A = d\iota(\xi)A + \iota(\xi)dA$ and $\iota(\xi)A = 1$. Hence $\iota(\xi)dA = 0$. Therefore dA is basic by Proposition 130. □

Definition 144. Let A be a connection 1-form on a principal S^1 bundle $\pi : P \to B$. The unique two-form F on B such that $\pi^* F = dA$ is called the **curvature** two-form of the connection A.

Lemma 145 (Special case of "cotangent bundle reduction theorem" of Abraham and Marsden, Kummer). *Let $S^1 \longrightarrow P \xrightarrow{\pi} B$ be a principal S^1-bundle. Let $\Phi : T^*P \longrightarrow \mathbb{R}$ denote the moment map for the lifted action of S^1 on T^*P. Then all points in $\mathbb{R}$ are regular values of Φ. Moreover, a choice of a connection 1-form A on P allows one to identify the reduced space $\Phi^{-1}(\lambda)/S^1$ at μ with the cotangent bundle T^*B of the base. Under this identification, the reduced form ω_λ satisfies $\omega_\lambda = \omega_{T^*B} + \lambda \pi_B^* F$ where $\pi_B : T^*B \longrightarrow B$ is the projection and F is the curvature of A.*

Proof. Since the action of S^1 on T^*P is a lifted action the moment map Φ is given by $\Phi(p, eta) = \langle \eta, \xi(p) \rangle$ for $\eta \in T_p^*P$. It follows that $\Phi^{-1}(\lambda) = \{(p, \eta) \in T^*P \mid \langle \eta, \xi(p) \rangle = \lambda\}$.

We claim that (1) all level sets $\Phi^{-1}(\lambda)$ are S^1-equivariantly diffeomorphic to the zero level set $\Phi^{-1}(0)$ and (2) that the zero level set is diffeomorphic to the pull-back of the vector bundle $\pi_B : T^*B \to B$ along the projection map $\pi : P \to B$ (see equation 27).

Indeed the map $s_\lambda(p, \eta) = (p, \eta + \lambda A(p))$ sends $\Phi^{-1}(0)$ to $\Phi^{-1}(\lambda)$: $\Phi(p, \eta + \lambda A) = \langle \eta + \lambda A(p), \xi(p) \rangle = \langle \eta, \xi(p) \rangle + \lambda A(p)(\xi(p)) = 0 + \lambda \cdot 1$. The map s_λ is an S^1-equivariant diffeomorphism of $\Phi^{-1}(0)$ and $\Phi^{-1}(\lambda)$.

To prove the second claim note first that $T_p^*P \cap \Phi^{-1}(0) = \{\eta \in T_p^*P, \langle \eta, \xi(p) \rangle = 0\}$, i.e., it is the annihilator of the kernel of the projection $d\pi_p : T_pP \to T_{\pi(p)}B$. It follows that the transpose $(d\pi_p)^T : T^*_{\pi(p)}B \to T_p^*P$, which is injective, is an isomorphism between $T^*_{\pi(p)}B$ and $T_p^*P \cap \Phi^{-1}(0)$. Define $\varpi_p : T_p^*P \cap \Phi^{-1}(0) \to T^*_{\pi(p)}B$ to be $((d\pi_p)^T)^{-1}$. This gives us a map $\varpi : \Phi^{-1}(0) \to T^*B$ making the diagram

$$
\begin{tikzcd}
\Phi^{-1}(0) \arrow[r, "\pi_P"] \arrow[d, "\varpi"'] & P \arrow[d, "\pi"] \\
T^*B \arrow[r, "\pi_B"] & B
\end{tikzcd}
\tag{27}
$$

commute. Consequently the quotient $\Phi^{-1}(0)/S^1$ is diffeomorphic to T^*B. Since map $s_\lambda : \Phi^{-1}(0) \to \Phi^{-1}(\lambda)$ is a S^1-equivariant, it descends to a diffeomorphism $\bar{s}_\lambda : \Phi^{-1}(0)/S^1 \to \Phi^{-1}(\lambda)/S^1$. Composing the $\bar{s}_\lambda$ with the diffeomorphism $T^*B \simeq \Phi^{-1}(0)/S^1$, we get the desired identification of the reduced space at λ with the cotangent bundle of the base.

It remains to compute the reduced symplectic form ω_λ. Let α_P denote the tautological 1-form on T^*P, and let α_B denote the tautological 1-form on T^*B. We claim that

$$s_\lambda^* \alpha_P = \varpi^* \alpha_B + \lambda \pi_P^* A.$$

To prove the claim we compute in coordinates. Let $U \subseteq B$ be a sufficiently small open set so that $\pi^{-1}(U)$ is diffeomorphic to $U \times S^1$ and so that U has coordinates $(q_1, \ldots, q_n)$. Let $(q_1, \ldots, q_n, \theta, \eta_1, \ldots, \eta_n, \eta_\theta)$ be the corresponding coordinates on $T^*(\pi^{-1}(U)) \simeq T^*U \times T^*S^1$.

In these coordinates the induced vector field $\xi = \frac{\partial}{\partial\theta}$. Since $A(\xi) = 1$, the connection one-form A has to be of the form $A = \sum a_i dq_i + d\theta$ where $a_i(q, \theta) \in C^\infty(U \times S^1)$ are S^1-invariant functions (since A is S^1 invariant). Therefore $a_i(q, \theta) = a_i(q)$, i.e., a_i's do not depend on θ.

In coordinates $(q_1, \ldots, q_n, \theta, \eta_1, \ldots, \eta_n, \eta_\theta)$ the moment map is given by

$$\Phi(q_1, \ldots, q_n, \theta, \eta_1, \ldots, \eta_n, \eta_\theta) = \eta_\theta.$$

Therefore $\Phi^{-1}(0) = \{(q_1, \ldots, q_n, \theta, \eta_1, \ldots, \eta_n, 0)\}$. Since $s_\lambda(\sum \eta_i dq_i) = \sum \eta_i dq_i + \lambda(\sum a_i dq_i + d\theta)$, it follows that in coordinates the map s_λ is given by the formula:

$$s_\lambda(q_1, \ldots, q_n, \theta, \eta_1, \ldots, \eta_n, 0) = (q_1, \ldots, q_n, \theta, \eta_1 + \lambda a_1, \ldots, \eta_n + \lambda a_n, \lambda).$$

Since $\alpha_P = \sum \eta_i dq_i + \eta_\theta d_\theta$,

$$s_\lambda^* \alpha_P = \sum (\eta_i + \lambda a_i) dq_i + \lambda d\theta = \underbrace{\sum \eta_i dq_i}_{\alpha_B} + \lambda \underbrace{\left(\sum a_i dq_i + d\theta\right)}_{A}.$$

Since $\omega_P = d\alpha_P$ and $\omega_B = d\alpha_B$,

$$\begin{aligned} s_\lambda^* \omega_P &= \varpi^* \omega_B - \lambda \pi_P^* dA \\ &= \varpi^* \omega_B - \lambda \pi_P^* \pi^* F \quad \text{by definition of } F \\ &= \varpi^* \omega_B - \lambda \varpi^* \pi_B^* F \quad \text{since the diagram (27) commutes} \\ &= \varpi^* (\omega_B + \lambda \pi_B^* F). \end{aligned}$$

We conclude that the reduced form ω_λ under the identification of $\Phi^{-1}(\lambda)/S^1$ with T^*B is $\omega_\lambda = \omega_B + \lambda \pi_B^* F$. □

We now consider a classical Hamiltonian system on the cotangent bundle of the principal bundle P which is S^1 invariant and compute the corresponding reduced Hamiltonian systems. Let g be an S^1 invariant metric on the principal S^1 bundle $\pi : P \to B$. Let V be a smooth S^1-invariant function on P. Since $B = P/S^1$, there exists a smooth function V_B on B with $V = \pi^* V_B$. The corresponding classical Hamiltonian h is of the form

$$h(p, \eta) = \frac{1}{2} g^*(p)(\eta, \eta) + (\pi^* V_B)(p)$$

where g^* is the dual metric on T^*P.

Since the metric g is S^1 invariant, it defines by equation (26) a connection one-form A on P . Unwinding the definitions we discover that $g^*(p)(A(p), (d\pi)^T \nu) = 0$ for all covectors $\nu \in T^*_{\pi(p)}B$. Since $\Phi^{-1}(0) \cap T_p^* P = (d\pi)^T (T^*_{\pi(p)} B)$, we see that $g^*(p)(\eta, A(p)) = 0$ for all $(p, \eta) \in \Phi^{-1}(0)$. Therefore

$$\begin{aligned} (s_\lambda^* h)(p, \eta) &= \frac{1}{2} g^*(p)(\eta + \lambda A(p), \eta + \lambda A(p)) + \pi^* V_B(p) \\ &= \frac{1}{2} g^*(p)(\eta, \eta) + \frac{\lambda^2}{2} g^*(p)(A(p), A(p)) + \pi^* V_B(p). \end{aligned}$$

Since $\frac{\lambda^2}{2}g^*(p)(A(p), A(p))$ is S^1-invariant, the function $\frac{\lambda^2}{2}g^*(p)(A(p), A(p)) + \pi^* V_B(p)$ is a pull-back of a smooth function V_{eff} on B. Since the metric g is S^1-invariant and since $d\pi_p : (\mathbb{R}\xi(p))^g \to T_{\pi(p)}B$ is an isomorphism, g defines a metric g_B on B making $d\pi_p : (\mathbb{R}\xi(p))^g \to T_{\pi(p)}B$ an isometry.

We conclude that under the identification of the reduced space at λ with T^*B induced by s_λ the reduced Hamiltonian h_λ is given by

$$h_\lambda(b, \eta) = \frac{1}{2}g_B^*(b)(\eta, \eta) + V_{\text{eff}}(b)$$

for all $\eta \in T_b^*B$.

We are now in position to apply our computations to the Lagrange top. Let us reduce by the right action of S^1. Since the metric $\mathbb{I}$ on SO(3) is left SO(3)-invariant, the reduced metric g on $S^2 = \mathrm{SO}(3)/S^1$ is also SO(3)-invariant. It follows that the kinetic energy term of the reduced Hamiltonian is (up to a constant multiple) determined by the standard round metric on S^2. Since the connection A is SO(3)-invariant (since A is defined by an SO(3)-invariant metric), the curvature F of A is also an SO(3) invariant two-form. Hence, up to a constant multiple, F is the standard area 2-form. Also, since A is SO(3) invariant, the function on S^2 defined by $\frac{\lambda^2}{2}g^*(p)(A(p), A(p))$ is SO(3)-invariant, hence is constant. It is no loss of generality to assume that it is zero (since constant terms do not affect the Hamiltonian vector field).

We conclude that the reduced Hamiltonian h_λ on T^*S^2 is given by

$$h_\lambda(q, p) = \frac{1}{2}g^*(q)(p, p) + q_3$$

where g is the round metric on S^2. Note that this is exactly the Hamiltonian of the spherical pendulum. The symplectic form on T^*S^2, however, is the sum of the standard symplectic form and of the pull-back of a multiple of an area form on S^2.

23. Lecture 23. Extremal equilibria and stability

In this section we apply a stability result of James Montaldi [Persistence and stability of relative equilibria, *Nonlinearity* **10** (1997), no. 2, 449–466] to show that a Lagrange top spinning in a upright position is stable, provided the spin is sufficiently large.

We will need the following topological result on group actions that we won't have time to prove.

Proposition 146. *Suppose a compact Lie group G acts on a (Hausdorff, second countable) manifold M. The the quotient space M/G is Hausdorff and locally compact.*

The next few paragraphs set up notation. Consider now a symmetric Hamiltonian system $(M, \omega, G, \Phi : M \to \mathfrak{g}^*, h \in C^\infty(M)^G)$. The flow ψ_t of the Hamiltonian vector field of h is G-equivariant. Therefore it descends to a flow $\bar{\psi}_t$ on the quotient M/G.

Notation 2. For a point $m \in M$ the orbit $G \cdot m$ is a point in M/G. We denote it by $\bar{m}$.

Note that a point $m \in M$ is a relative equilibrium (of our symmetric Hamiltonian system) if and only if $\bar{m}$ is fixed by $\bar{\psi}_t$.

Definition 147. A relative equilibrium $m \in M$ of a symmetric Hamiltonian system is **G-stable** if for any neighborhood U of $\bar{m}$ in M/G there is a neighborhood V of $\bar{m}$ such that $\bar{\psi}_t(V) \subset U$ for all t, i.e., the point $\bar{m}$ is Liapunov stable in M/G.

Since the moment map is G-equivariant, it descends to a map $\bar{\Phi} : M/G \to \mathfrak{g}^*/G$. Note that the level sets of $\bar{\Phi}$ are the quotients of the form $\Phi^{-1}(G \cdot \mu)/G$, $\mu \in \mathfrak{g}^*$, that is, they are the reduced spaces.

Since the Hamiltonian h is G-invariant, it descends to a map $\bar{h} : M/G \to \mathbb{R}$. Observe that the restriction of $\bar{h}$ to $\bar{\Phi}^{-1}(G \cdot \mu)$ is the reduced Hamiltonian h_μ on the reduced space $M_\mu = \Phi^{-1}(G \cdot \mu)/G$. Note that we view the reduced spaces M_μ simply as topological spaces.

Definition 148. A relative equilibrium $m \in M$ is **extremal** if $\bar{m} \in M/G$ is a local extremum for the reduced Hamiltonian h_μ on the reduced space $M_\mu = \bar{\Phi}^{-1}(\mu)$, where $\mu = \bar{\Phi}(\bar{m})$.

Theorem 149 (Montaldi)**.** *Let $(M, \omega, G, \Phi : M \to \mathfrak{g}^*, h \in C^\infty(M)^G)$ be a symmetric Hamiltonian system. If a point $m \in M$ is an extremal relative equilibrium than it is G-stable.*

The proof of the theorem relies on the following topological lemma. Recall that a precompact set is a set with a compact closure and that a precompact neighborhood is open.

Lemma 150. *Let $f : X \to Y$ be a continuous map between two topological spaces with X locally compact and Y Hausdorff. Suppose a point $y \in Y$ is such that $S := f^{-1}(y)$ is compact. Then for any precompact neighborhood U of S there exists a neighborhood V of y such that $f^{-1}(V)$ and the boundary ∂U are disjoint.*

Proof. First note that since X is locally compact, the set S does indeed have a precompact neighborhood: each point of S has a precompact neighborhood, so extracting a finite subcover we obtain the desired precompact neighborhood of S. Let $\{V_\alpha\}$ be the collection of all closed neighborhoods of $y \in Y$ (so that $\cap_\beta V_\beta = \{y\}$ since Y is Hausdorff). Let $Z_\beta = f^{-1}(V_\beta)$; the sets Z_β are closed. Then $\cap_\beta Z_\beta = S$, and since $S \cap \partial U = \emptyset$, we have $\partial U \subset X \smallsetminus \cap_\beta Z_\beta = \cup_\beta (X \smallsetminus Z_\beta)$. Since ∂U is compact there is a finite subcover $\{X \smallsetminus Z_{\beta_i}\}_{i=1}^n$ of ∂U. Then

$$V = \bigcap_{i=1}^{n} \operatorname{int}(V_{\beta_i})$$

is a neighborhood of y with the desired property (here $\operatorname{int}(V_{\beta_i})$ denotes the interior of V_{β_i}). $\square$

Proof of Theorem 149. It is no loss of generality to assume that $\bar{m}$ is a local maximum for $\bar{h}|_{\bar{\Phi}^{-1}(\mu)}$ where $\mu = \bar{\Phi}(\bar{m})$. Let X be a small enough neighborhood of $\bar{m}$ in M/G such that $\bar{h}(\bar{m})$ is the maximal value of $\bar{h}$ on $X \cap \bar{\Phi}^{-1}(\mu)$. Let $Y = \mathfrak{g}^*/G \times \mathbb{R}$, and let $f : X \to Y$ be given by $f(x) = (\bar{\Phi}(x), \bar{h}(x))$. Then for $y = (\mu, \bar{h}(\bar{m})$ the set $f^{-1}(y)$ is one point $\{\bar{m}\}$, hence compact. Let U be a precompact neighborhood of $f^{-1}(y)$. By Lemma 150 there is a neighborhood V of y in Y such that $f^{-1}(V) \cap \partial U = \emptyset$. Then since the fibers of f are preserved by the flow

$\bar{\psi}_t$ and since the flow preserves connected components, the set $f^{-1}(V) \cap U$ is a flow-invariant neighborhood of $\bar{m}$ as required. (Recall that $\bar{\psi}_t$ denotes the flow on M/G induced by the flow of the Hamiltonian vector field of h on M. □

We now apply Theorem 149 to the Lagrange top. Recall from Lecture 22 that after one reduction we can think of Lagrange top as a family of Hamiltonian systems $(T^*S^2, \omega_\lambda, h_\lambda)$ where $h_\lambda(q,p) = \frac{1}{2}g^*(q)(p,p) + q_3$ is the spherical pendulum Hamiltonian (g^* is the dual of the round metric on S^2) and ω_λ is the sum the standard symplectic form and of the pull-back of a multiple of an area form on S^2: $\omega_\lambda = \omega_{T^*S^2} + \lambda\nu$, where $\nu = (q_1 dq_2 \wedge dq_3 + q_2 dq_3 \wedge dq_1 + q_3 dq_1 \wedge dq_2)|_{S^2}$. We would like to understand stability of the North pole $(0,0,1,0,0,0)$ for various values of λ. We do it by computing in coordinates.

Consider coordinates on the upper hemisphere of S^2 induced by the projection $(q_1, q_2, q_3) \mapsto (q_1, q_2)$. The inverse map ϕ is given by $\phi(x,y) = (x, y, \sqrt{1 - x^2 - y^2})$. Since

$$\phi^*(q_1 dq_2 \wedge dq_3 + q_2 dq_3 \wedge dq_1 + q_3 dq_1 \wedge dq_2) = \frac{1}{z} dx \wedge dy,$$

where $z = \sqrt{1 - x^2 - y^2}$, the symplectic form ω_λ in the canonical coordinates corresponding to (x,y) is

$$\omega_\lambda = dx \wedge dp_x + dy \wedge dp_y + \frac{\lambda}{z} dx \wedge dy.$$

The systems $(T^*S^2, \omega_\lambda, h_\lambda)$ have SO(2) as a symmetry group: SO(2) acts by the lift of the rotation of S^2 about the vertical axis. The action is Hamiltonian; we denote a corresponding moment map by Φ_λ. In coordinates the action of $\mathrm{SO}(2) = S^1$ is given by

$$e^{i\theta} \cdot (x, y, p_x, p_y) = \left(\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} p_x \\ p_y \end{pmatrix} \right).$$

The moment map Φ_λ (which is defined up to an additive constnat) is then

$$\Phi_\lambda(x, y, p_x, p_y) = y p_x - x p_y - \lambda z + \lambda.$$

Note that the North pole in these coordinates is the origin and that the moment map is normalized so that $\Phi_\lambda(0) = 0$.

The metric g by assumption is the one that comes from the embedding of S^2 as a round sphere in $\mathbb{R}^3$. We now normalize the metric by setting the radius of the sphere to 1. Then the sphere is cut out by the equation

$$\sum q_i^2 = 1, \tag{28}$$

In our coordinates the metric g is given by

$$(g_{ij}) = \begin{pmatrix} 1 + \frac{x^2}{z^2} & \frac{xy}{z^2} \\ \frac{xy}{z^2} & 1 + \frac{y^2}{z^2} \end{pmatrix}.$$

Consequently the dual metric $g*$ is given by

$$(g^{ij}) = z^2 \begin{pmatrix} 1+\frac{y^2}{z^2} & -\frac{xy}{z^2} \\ -\frac{xy}{z^2} & 1+\frac{x^2}{z^2} \end{pmatrix}.$$

Hence the Hamiltonian h_λ is

$$h_\lambda(x,y,p_x,p_y) = \frac{1}{2}\left(\left(1+\frac{y^2}{z^2}\right)p_x^2 - 2\frac{xy}{z^2}p_xp_y + \left(1+\frac{x^2}{z^2}\right)p_y^2\right) + z.$$

We are now in position to compute the reduced space (which is not going to be smooth!) and the reduced Hamiltonian.

Observe that $\Phi_\lambda^{-1}(0)\cap\{y=0\}$ is the set

$$\{(x,0,p_x,p_y) \mid p_y = \lambda\frac{(1-\sqrt{1-x^2})}{x}\}.$$

Note that

$$1-\sqrt{1-x^2} = 1-(1-\frac{1}{2}x^2 - \frac{1}{8}x^4 + \mathcal{O}(x^6)) = \frac{1}{2}x^2 + \frac{1}{4}x^4 + \mathcal{O}(x^6),$$

so

$$\lambda\frac{(1-\sqrt{1-x^2})}{x} = \frac{\lambda}{2}(x+\frac{1}{8}x^3 + \mathcal{O}(x^5))$$

is smooth at $x=0$. Consider now a map $\psi : (-1,1)\times\mathbb{R}\to\Phi_\lambda^{-1}(0)$, $(u,p_u)\mapsto(u,0,p_u,\lambda(1-\sqrt{1-u^2})/u)$. Given any point point (x,y,p_x,p_y) there is an angle θ so that $e^{i\theta}\cdot(x,y,p_x,p_y)$ is of the form $(*,0,*,*)$. It follow that the image of ψ in $\Phi_\lambda^{-1}(0)$ intersects every orbit of SO(2). In fact the image intersects every orbit except $\{(0,0,0,0)\}$ in exactly two points: $(u,0,p_u,\lambda(1-\sqrt{1-u^2})/u)$ and $(-u,0,-p_u,-\lambda(1-\sqrt{1-u^2})/u)$. Hence the composition

$$(-1,1)\times\mathbb{R}\xrightarrow{\psi}\Phi_\lambda^{-1}(0)\to\Phi_\lambda^{-1}(0)/SO(2)$$

is a branched double cover. Therefore, to understand the extremal points of the reduced Hamiltonian, it is enough to understand the extremal points of ψ^*h_λ. It is easy to see that

$$\begin{aligned}\psi^*h_\lambda(u,p_u) &= \frac{1-u^2}{2}\left[p_u^2 + \left(1+\frac{u^2}{1-u^2}\right)\left(\lambda\frac{1-\sqrt{1-u^2}}{u}\right)^2\right] + \sqrt{1-u^2} \\ &= \frac{1}{2}(1-u^2)p_u^2 + \frac{1}{2}\lambda^2\left(\frac{1-\sqrt{1-u^2}}{u}\right)^2 + \sqrt{1-u^2}.\end{aligned}$$

Hence the critical points of the function $\psi^*h_\lambda(u,p_u)$ are of the form $(u^*,0)$ where u^* is a critical point of the effective potential

$$U_\lambda(u) = \frac{1}{2}\lambda^2\left(\frac{1-\sqrt{1-u^2}}{u}\right)^2 + \sqrt{1-u^2}.$$

Moreover, relative maxima of U_λ correspond to unstable equilibria of H_λ and relative minima to stable equilibria.

Remark 151. Since we are working on the branched double cover of the reduced system, any pair of critical points of U_λ of the form $\pm u$, $u \neq 0$ correspond to the same critical point of the reduced system. Therefore only the non-negative critical points u^* of the potential need be considered.

We now concentrate on $u^* = 0$. Since

$$\sqrt{1-u^2} = 1 - \frac{1}{2}u^2 + \mathcal{O}(u^4)$$

$$\left(\frac{1-\sqrt{1-u^2}}{u}\right)^2 = \frac{1}{2}u + \mathcal{O}(u^3))^2 = \frac{1}{4}u^2 + \mathcal{O}(u^3).$$

Therefore

$$U_\lambda(u) = \frac{1}{2}\lambda^2\frac{1}{4}u^2 + 1 - \frac{1}{2}u^2 + \mathcal{O}(u^3) = 1 + \frac{1}{2}\left(\frac{\lambda^2}{4} - 1\right)u^2 + \mathcal{O}(u^3).$$

We conclude that if $\frac{\lambda^2}{4} > 1$ then the North pole is a stable equilibrium of the Hamiltonian systems $(T^*S^2, \omega_\lambda, h_\lambda)$. Hence the Lagrange top in the upright position is SO(2)-stable if it is spinning sufficiently fast.

Homework Problem 26. if $1 < \lambda^2 < 4$ then the potential U_λ has another positive critical point u^*. Find u^* and deterine the stability of the corresponding relative equilibrium.

24. Lecture 24. Completely integrable systems

Definition 152. Let (M, ω) be a symplectic manifold and let h be a smooth function on M. The Hamiltonian system (M, ω, h) is **completely integrable** if there exists $n = \frac{1}{2}\dim M$ smooth functions $f_1 = h, f_2, \dots, f_n$ on M such that

1. the functions Poisson commute: $\{f_i, f_j\} = 0$ for all i and j;
2. the differentials df_i are linearly independent on a dense open subset M_{reg} of M.

By abuse of language we will refer to the collection $(M, \omega, f_1, \dots f_n)$ as a completely integrable system.

Proposition 153. *Let (M, ω) be a symplectic manifold. Suppose that there exist k smooth functions $f_1, \dots, f_k$ on M which Poisson commute (i.e.,$\{f_i, f_j\} = 0$ for all i and j). If $c \in \mathbb{R}^k$ is a regular value of the map $f : M \to \mathbb{R}^k$ defined by $f(m) = (f_1(m), \dots, f_k(m))$, then the submanifold $Z := f^{-1}(c)$ is coisotropic. In particular, if $k = \frac{1}{2}\dim M$, then Z is Lagrangian.*

Proof. Fix $z \in Z$. The tangent space T_zZ is $\ker df(z) = \cap \ker df_i(z)$. Let V be the subspace of T_zM spanned by the Hamiltonian vector fields $X_{f_i}(z)$. We claim that T_zZ is the symplectic perpendicular to V. Indeed

$$\begin{aligned} v \in T_z Z \iff & & v \in \ker df_i(z) \quad & \text{for all } i \\ \iff & & 0 = \langle df_i(z), v\rangle = \omega(X_{f_i}, v) \quad & \text{for all } i \\ \iff & v \in V^\omega & & \end{aligned}$$

On the other hand, since

$$0 = \{f_i, f_j\} = -\omega(X_{f_i}, X_{f_j}),$$

the subspace V is isotropic: $V \subset V^\omega$. Hence $T_z Z^\omega = (V^\omega)^\omega = V \subset V^\omega = T_z Z$, i.e., $T_z Z$ is coisotropic. Finally note that the coisotropic submanifolds of half the dimension of the manifold are Lagrangian. ☐

Consider again a completely integrable system $(M, \omega, f_1, \dots, f_n)$. We now make two blanket assumptions:

1. The map $f = (f_1, \dots, f_n) : M \to \mathbb{R}^n$ is proper.
2. The level sets of the map $f : M_{reg} \to \mathbb{R}^n$ are connected.

Note the properness of f implies that the level sets $f^{-1}(r)$ of f are all compact.

Theorem 154 (Eheresman fibration theorem)**.** *Let X and Y be manifolds. Assume that Y is connected. Suppose $f : X \to Y$ be a proper submersion. Then $f : X \to Y$ is a fiber bundle with a typical fiber $f^{-1}(y)$, for some $y \in Y$.*

Proof. See, for example, [Bröker and Jänich]. ☐

We conclude that under our assumptions the study of completely integrable systems leads to a study of fiber bundles with the following properties: the total space is symplectic and the fibers are compact connected Lagrangian submanifolds.

Theorem 155. *Let (Q, ω) be a symplectic manifold. Suppose $F \to Q \xrightarrow{\rho} B$ is a fiber bundle with compact connected Lagrangian fibers. Then the fibers are tori. In fact, for every $b \in B$, the abelian group $T_b^* B$ acts transitively on the fiber $F_b = \rho^{-1}(b)$ with a zero dimensional isotropy group.*

Remark 156. Recall that if a group G acts on a space X, then then the isotropy groups of two points on the same orbit are conjugate: if $y = g \cdot x$ then $G_y = g G_x g^{-1}$. Hence if G is abelian, the isotropy group of a point is the same for all points in the orbit. It follows that the isotropy group $L_b \subset T_b^* B$ for the action of $T_b^* B$ on the fiber F_b depends only on $b \in B$. The set $L = \cup_{b \in B} L_b$ is called the **period lattice**.

Proof. 1. We start by defining a linear map v from the vector space $T_b^* B$ to the space of vector fields $\chi(F_b)$ on the fiber $F_b = \rho^{-1}(b)$.

Let $p \in T_b^* B$ be a covector. Then there exists a smooth function f on B such that $df(b) = p$ (For example if $p = \sum a_i dx_i$ in coordinates, then $\tilde{f} = \sum a_i x_i$ has the desired property. Extend $\tilde{f}$ to a smooth function on all of B by multiplying it by an appropriate compactly supported function). Consider the Hamiltonian vector field $X_{\rho^* f}$ of the function $\rho^* f$ on (Q, ω). Then for $q \in F_b$

$$(\omega_q)^{\#}(X_{\rho^* f}(q)) = d(\rho^* f)(q) = d\rho(q) \circ df(b) = d\rho(q) \circ p.$$

Since F_b is Lagrangian in (Q, ω), for any $q \in F_b$ the map $(\omega_q)^{\#}$ sends the tangent space $T_q F_b$ bijectively onto the annihilator $(T_q F_b)^\circ$ of $T_q F_b$ in $T_q^* Q$. For any $p \in T_b^* B$, the covector $d\rho(q) \circ p$ lies in $(T_q F_b)^\circ$. Consequently the vectors $X_{\rho^* f}(q) = ((\omega_q)^{\#})^{-1}(d\rho(q) \circ p)$ form a vector field on the fiber F_b. Note that this vector field depend only on the covector p and not on the function f. We therefore get a map $v : T_b^* B \to \chi(F_b)$ defined by

$$p \mapsto (q \mapsto ((\omega_q)^{\#})^{-1}(d\rho(q) \circ p).$$

Note well that v is **linear** and a **bijection**.

2. Next we argue that for any covectors $p, p' \in T_b^* B$ the vector fields $v(p)$ and $v(p')$ commute. Choose $f' \in C^\infty(B)$ such that $df'(b) = p'$. Then since the vector field $X_{\rho^* f}$ is tangent to the fibers of ρ and the function $\rho^* f'$ is constant on the fibers of ρ,

$$0 = X_{\rho^* f}(\rho^* f') = \{\rho^* f, \rho^* f'\}.$$

Since the map $h \mapsto X_h$, $C^\infty(Q) \to \chi(Q)$, is a Lie algebra map, it follows that $[v(p), v(p')] = [X_{\rho^* f}, X_{\rho^* f'}] = X_{\{\rho^* f, \rho^* f'\}} = 0$.

3. Let φ_t^p denote the time t flow of the vector field $v(p)$ on the fiber F_b. Since the fiber is compact, the flow exists for all time. Since $[v(p), v(p')] = 0$ the flows φ_t^p and $\varphi_s^{p'}$ commute:

$$\varphi_t^p \circ \varphi_s^{p'} = \varphi_s^{p'} \circ \varphi_t^p. \tag{29}$$

Also, since v is linear,

$$\varphi_t^{ap} = \varphi_{at}^p.$$

We now define a map $T_b^* B \times F_b \to F_b$ by

$$(p, q) \mapsto p \cdot q = \varphi_1^p(q).$$

It follows from (29) that the map is an **action** of the abelian Lie group $G = T_b^* B$ on the manifold F_b.

4. It remains to argue that the action is transitive on F_b and has zero dimensional isotropy group. These are consequences of the fact that the map v is bijective.

Since the fiber F_b is connected and has the same dimension as the group $T_b^* B$, it is enough to prove that the orbits are open. For the latter it is enough to show that the induced map from the Lie algebra $\mathfrak{g}$ of $G = T_b^* B$ to a tangent space $T_q(F_b)$ is a bijection. Now for any $p \in T_b^* B$ the curve

$$\gamma_p(s) = ps$$

is a one-parameter subgroup of the Lie group $T_b^* B$. For any $q \in F_b$

$$\frac{d}{dt}\Big|_{t=0} \gamma_p(t) \cdot q = \frac{d}{dt}\Big|_{t=0} \varphi_1^{tp}(q) = \frac{d}{dt}\Big|_{t=0} \varphi_{t \cdot 1}^p(q) = v(p)(q).$$

Since the map v is bijective, the induced map $\mathfrak{g} \to T_q(F_b)$ is an isomorphism and we are done. $\square$