ELEMENTARY DIFFERENTIAL GEOMETRY

YONG-GEUN OH

– Based on the lecture note of Math 621-2020 in POSTECH –

Contents

Part 1. Riemannian Geometry

1. Parallelism and Ehresman connection

For a given vector bundle $\pi : E \to M$, we have natural exact sequence

$$0 \to \ker d_e\pi \to T_eE \stackrel{d_e\pi}{\to} T_{\pi(e)}M \to 0.$$

The kernel $\ker d\pi \subset TE$ forms a subbunle of rank $k = \operatorname{rank} E$ which defines an integrable distribution on the manifold E. Its integral submanifolds are nothing but the fibers $\pi^{-1}(x)$, $x \in M$. We denote $VT_eE = \ker d_e\pi$ and call it as the *vertical tangent space* of E at e, and the union

$$VTE = \bigcup_{e\in E} VT_eE =: V_eE$$

the *vertical subbundle.* There is no canonical notion of *horizontal subspaces.*

Definition 1.1. An *Ehresmann connection* is an assignment of subspaces $H_e \subset T_eE$ complementary to V_eE in T_eE at each $e \in E$ such that

(1) $HE := \cup_{e\in E} H_e$ is a subbundle of TE such that $T_eE = H_eE \oplus V_eE$,
(2) For any $e \in E$ and $\lambda \in \mathbb{R}$, $H_{\lambda e} = dm_\lambda(H_e)$ where $m_\lambda : E \to E$ is the scalar multiplication by λ.

Statement (2) in particular implies $H_{0_x} = T_{0_x}0_E$ for the zero section $0_E \cong M$ and $0_x \in 0_E$ is the point corresponding to $x \in M$.

Postponing the discussion on the existence of Ehresmann connection till later, we proceed. Denote an Ehresmann connection of E as a splitting

$$\Gamma : TE = HE \oplus VE$$

for $HE \subset TE$ the aforementioned subbundle.

We now recall the definition of the pull-back bundle f^*E for a smooth map $f : N \to M$:

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

Denote by $[n, e]$ the element of f^*E represented by $(n, e) \in N \times E$. Then we can express

$$T_{[n,e]}(f^*E) = \{(v, \xi) \in T_nN \times T_eE \mid df(v) = d\pi(\xi)\}.$$

Definition 1.2 (Pull-back connection)**.** Let $\pi : E \to M$ be a vector bundle and $f : N \to M$ be a smooth map. The pull-back connection $f^*\Gamma$ is given by the choice of a horizontal subbunle of $T(f^*E)$ given by

$$H_{[n,e]}(f^*E) := \{(v, \xi) \in T_nN \times T_eE \mid df(v) = d\pi(\xi),\, \xi \in H_{f(n)}E\} \cong H_{f(n)}E.$$

Let $\pi : E \to M$ be a the vector bundle of rank k. Consider the pull-back bundle $\gamma^*E \to I = [0,1]$ for the given path $\gamma : I \to M$.

Let me start with more abstract terms. The parallel transport has the following interpretation. Let $F = \gamma^*E \to [0,1] = B$ be the above pull-back bundle and $TF = H \oplus V$ its splitting induced by the pull-back connection. Note that H is a one-dimensional distribution on F and that the projection $d\pi : H \to TB$ induces an isomorphism that $TB = T[0,1]$ carries a canonical global frame $\{\frac{\partial}{\partial t}\}$. Therefore this lifts to a smooth (autonomous) vector field X on $F = \gamma^*E$ (regarded as a manifold) given by

$$X(e) := (d_e\pi|_H)^{-1}\left(\frac{\partial}{\partial t}\right)$$

on F. By definition

$$d_e\pi(X(e)) = \frac{\partial}{\partial t}\Big|_t.$$

Therefore we can use t itself as the *time* for the ODE $\dot{e} = X(e)$ on F.

We consider the integral curve $\widetilde{\gamma}$ of X issued at $e \in \gamma^*E$, i.e., satisfying the ODE

$$\frac{d\widetilde{\gamma}}{dt} = X(\widetilde{\gamma}(t)),\ \widetilde{\gamma}(0) = e.$$

This exists by the existence theorem of the first-order ODE on whole $[0,1]$ and is smooth by the smooth dependence of the solution of ODE on its initial data.

By denoting the solution by $\widetilde{\gamma}_e$, we have $\widetilde{\gamma}_e(t) \in \gamma^*E|_t$ and interpret is as a curve in E noting that

$$\gamma^*E|_t = E_{\gamma(t)}$$

Definition 1.3. The parallel transport $\Pi_\gamma : E_{\gamma(0)} \to E_{\gamma(1)}$ is defined by

$$\Pi_\gamma(e) := \widetilde{\gamma}_e(1).$$

Proposition 1.4. *The map $\Pi_\gamma : E_{\gamma(0)} \to E_{\gamma(1))}$ is a linear map.*

Proof. We start with proving $\Pi_\gamma(\lambda e) = \lambda\Pi_\gamma(e)$. Let $\widetilde{\gamma}_e$ is the horizontal lift of γ with $\widetilde{\gamma}_e(0) = e$. We consider the curve $\ell : [0,1] \to E$ defined by $\ell(t) = \lambda\widetilde{\gamma}_e(t)$. Clearly $\ell(t) = \lambda e$ and $\frac{d\widetilde{\gamma}_e}{dt}|_t \in H_{\widetilde{\gamma}_e(t)}$ since $\widetilde{\gamma}_e$ is the horizontal lift of γ. We compute the derivative

$$\dot{\ell}(t) = \frac{d}{dt}(\lambda\widetilde{\gamma}_e)\Big|_t = dm_\lambda\left(\frac{d\widetilde{\gamma}_e}{dt}\Big|_t\right) \in H_{\lambda\widetilde{\gamma}_e(t)}$$

by the defining property of the horizontal subspace in the definition of Ehresmann connection.

It is also a lift of γ. Therefore by the uniqueness of the lift, we must have

$$\lambda\widetilde{\gamma}_e(t) = \widetilde{\gamma}_{\lambda e}(t)$$

for all $t \in [0,1]$. In particular, we have $\lambda\widetilde{\gamma}_e(1) = \widetilde{\gamma}_{\lambda e}(1)$. By definition, the last is equivalent to $\Pi_\gamma(\lambda e) = \lambda\Pi_\gamma(e)$.

Next we quote the following lemma

Lemma 1.5. *Let V, W be vector spaces and $f : V \to W$ be a map differentiable at $0 \in V$. If f satisfies $f(\lambda v) = \lambda f(v)$ for all $\lambda \in \mathbb{R}$ and $v \in V$. Then $f = df(0)$. In particular f is a linear map.*

Proof. First by setting $\lambda = 0$, we obtain $f(0) = 0$. Therefore for $\lambda \neq 0$, we have

$$f(v) = \frac{1}{c} f(cv) = \frac{1}{c}(f(cv) - f(0)).$$

By letting $c \to 0$, we obtain

$$\lim_{c \to 0} \frac{1}{c}(f(cv) - f(0)) = df(0)(v)$$

since f is assumed to be differentiable at 0. This finishes the proof. □

Combining the above, we have proved the proposition. □

Now we introduce the notion of covariant derivative.

Definition 1.6. Let $\Gamma : TE = HE \oplus VE$ be an Ehresmann connection. Let $s \in \Gamma(E)$ be a section of E. Then we define the *covariant derivative* $\nabla_v s$ of s along $v \in T_x M$ to be

$$\nabla_v s := \frac{d}{dt}\Big|_{t=0} (\Pi_\gamma|_0^t)^{-1}(s(\gamma(t)))$$

for a (and so any) germ of curves $\gamma : (-\epsilon, \epsilon) \to M$, where $\Pi_\gamma|_0^t$ is the parallel transport along γ from 0 to t.

It can be checked that this definition does not depend on the choice of γ satisfying $\gamma'(0) = v$.

Remark 1.7. We can also write

$$\nabla_v s = \Pi_\Gamma^v(ds(v)) \in VT_{s(x)} \cong E_x$$

i.e., 'the covariant derivative is the vertical projection of $ds(v) \in T_{s(x)}E$ with respect to the Ehresmann connection $\Gamma : TE \oplus HE \oplus VE$ after identification of $VT_{s(x)} \cong E_x$."

The following properties can be also checked easily.

Lemma 1.8. *Let $x \in M$. The assignment*

$$s \in \Gamma(E) \to E_x$$

satisfies

(1) $\nabla_{v_1+v_2} s = \nabla_{v_1} s + \nabla_{v_2} s$ *for* v_1, $v_2 \in T_x M$,
(2) $\nabla_{cv} s = c\nabla_c s$ *for all* $c \in \mathbb{R}$ *and* $v \in T_x M$,
(3) $\nabla_v(fs) = f(x)\nabla_v s + v[f]\nabla_v s$ *for any* $f \in C^\infty(M)$, *where* $v[f]$ *is the directional derivative of* f *along* v *at* x,
(4) $\nabla_v(s_1 + s_2) = \nabla_v s_1 + \nabla_v s_2$ *for* $s_i \in \Gamma(E)$ *and* $v \in T_x M$.

We would like to mention that when E is the trivial line bundle $M \times \mathbb{R}$, ∇_v is the same as the tangent vector v as a derivation at x with values in $\mathbb{R}$.

2. Affine connections on vector bundles

Definition 2.1. Let $x \in M$. A linear map $D : \Gamma(E) \to E_x$ is called defines a *derivation at* x *with values in* E if D satisfies the properties (1) - (4) above. A smooth family of such D over M is called a *differential operator of order 1* on $\Gamma(E)$.

An example of a differential operator of order 1 on $\Gamma(E)$ is ∇_X for any vector field X on M.

Definition 2.2 (Affine connection)**.** An *affine connection* on $E \to M$ is the assignment of $X \mapsto \nabla_X$ that satisfies

(1) $\nabla_{X_1+X_2} s = \nabla_{X_1} s + \nabla_{X_2} s$,
(2) $\nabla_{cX} s = c\nabla_X s$,
(3) $\nabla_X(fs) = f\nabla_X s + X[f]\nabla_X s$ for any $f \in C^\infty(M)$,
(4) $\nabla_X(s_1 + s_2) = \nabla_X s_1 + \nabla_X s_2$ for $s_i \in \Gamma(E)$.

In particular, ∇ defines a differential operator of order 1

$$\nabla : \Omega^0(E) = \Gamma(E) \to \Omega^1(E) = \Gamma(T^*M \otimes E).$$

If we denote by $\Omega^k(E)$ the E-valued differential k-form on M $\Gamma(\Lambda^k(T * M) \otimes E)$ for $k = 0, 1, \ldots,$, and extend ∇ to

$$d^\nabla : \Omega^k(E) \to \Omega^{k+1}(E)$$

as a derivation under the wedge product of T^*M by anti-symmetrization as for the exterior derivative $d : \Omega^k(M) \to \Omega^{k+1}(M)$ with $d^\nabla = \nabla$ on $\Omega^0(E)$.

Lemma 2.3. *Denote the space of connections on $E \to M$ by $\mathcal{A}(E)$. Then $\mathcal{A}(E)$ is an (infinite dimensional) affine space modelled by $\Omega^1(End(E))$.*

Proof. Let ∇, ∇' be two affine connections of E. We will show that $X \to \nabla_X - \nabla'_X$ is a tensor, i.e.,

$$(\nabla_X - \nabla'_X)(fs) = f(\nabla_X - \nabla'_X)(s)$$

for all $s \in \Gamma(E)$ and $f \in C^\infty(M)$. But

$$\begin{aligned}(\nabla_X - \nabla'_X)(fs) &= \nabla_X(fs) - \nabla'_X(fs) = (f\nabla_X s + X[f]s) - (f\nabla'_X s + X[f]s) \\ &= f(\nabla_X s - \nabla'_X s) = f(\nabla_X - \nabla'_X)(s).\end{aligned}$$

This proves the assignment $s \mapsto (\nabla - \nabla')s$ defines a one-form valued at $\mathrm{End}(E)$. □

In other words, for a given affine connection ∇, any other affine connection ∇' can be written as

$$\nabla' = \nabla + \omega$$

for an element $\omega \in \Omega^1(End(E))$.

Now we prove existence of an affine connection.

Theorem 2.4. *For any vector bundle $E \to M$, there exists an affine connection and so $\mathcal{A}(E)$ is a nonempty affine space modeled by $\Omega^1(\mathrm{End}(E))$.*

Proof. Let $\{U_\alpha\}$ be an atlas of M and $\{(\Phi_\alpha, U_\alpha)\}$ a compatible system of trivialization of E. We denote

$$\Phi(e_x) = (x, h_\alpha(x)), \quad h_\alpha : U_\alpha \to \mathbb{R}^k.$$

For given s, locally we have $\Phi_\alpha \circ s(x) = (x, s_\alpha(x))$ for $s_\alpha : U_\alpha \to \mathbb{R}^k$. Then we have the compatibility relation

$$(x, s_\beta(x)) = (x, g_{\alpha\beta} s_\alpha(x))$$

where $g_\alpha\beta$ is the transition map defined by

$$\Phi_\beta \circ \Phi_\alpha(x, e_\alpha) = (x, g_{\alpha\beta}(x)e_\beta)$$

where $(x, e_\alpha) = \Phi(x, e) = (x, e_\beta)$. By definition, we have local representatives of s which satisfy

$$s_\beta(y) = g_{\alpha\beta}(y)s_\alpha(y)$$

for any $x \in U_\alpha \cap U_\beta$.

Define
$$\nabla_v(s)(x) := \Phi_\alpha^{-1}(x, v[s_\alpha])$$
for $x \in U_\alpha$. By taking the directional derivative of this in v at x,
$$\nabla_v^\alpha(s)(x) := \Phi_\alpha^{-1}(x, v[s_\alpha])$$
and the sum
$$\nabla_v(s)(x) := \sum_\alpha \chi_\alpha(x)\Phi_\alpha^{-1}(x, v[s_\alpha]).$$
for a partitions of unity $\{\chi_\alpha\}$ adapted to $\{U_\alpha\}$. We check all the defining properties of a connection. □

Definition 2.5. Let ∇ be an affine connection on $E \to M$.

(1) The *dual connection* on E^* is define by the derivation relation
$$X[\alpha(s)] = \nabla_X \alpha(x) + \alpha(\nabla_X s)$$
i.e.,
$$\nabla_X \alpha(x) := X[\alpha(s)] - \alpha(\nabla_X s)$$

(2) The induced connection on $\mathcal{T}_k^\ell(E) = \Gamma(E^{\otimes k} \otimes (E^*)^{\otimes \ell})$ is given by the Leibnitz rule starting from action on $C^\infty(M)$, $\Gamma(E)$, $\Gamma(E^*)$. We just denote this extended connection by ∇ by an abuse of notation.

2.1. **Local expression of covariant derivatives.** Let s be any given local section on U and represent
$$\Phi \circ s(p) = (p, s^\Phi(p))$$
where $s^\Phi : U \to \mathbb{R}^k$ be the local representative map of s. Then we can express
$$\Phi(\nabla s|_p) = (p, d_p s^\Phi + \omega_p^\Phi s^\phi(p)).$$
We set $(\nabla s)^\Phi = ds^\Phi + \omega_p^\Phi s^\phi$ which is the local representative of ∇s and which is a $\mathbb{R}^k$-valued one-form. We usually write this as
$$\nabla s^\Phi = ds^\Phi + \omega^\Phi s^\Phi$$
where $\omega^\Phi = (\omega_\alpha^\beta)_{1 \leq \alpha,\beta \leq k}$ is a matrix valued one-form.

In terms of the local frame $\{E_\alpha\}$, we can determine $\omega^\Phi = (\omega_\alpha^\beta)$ by the formula
$$\nabla E_\alpha = \sum_\beta \omega_\alpha^\beta E_\beta. \tag{2.1}$$

We call $\omega^\Phi = (\omega_\alpha^\beta)$ the *connection matrix* and ω_α^β are connection forms of the frame $\{E_\alpha\}$ (associated to Φ.)

In particular, the (local) section
$$\nabla_{\frac{\partial}{\partial x^j}} s$$
has the local representative given by $\frac{\partial s^\Phi}{\partial x^j}(p) + \omega_p^\Phi(\frac{\partial}{\partial x^j})s^\Phi(p)$. In other words,
$$(\nabla_{\frac{\partial}{\partial x^j}}(s))^\Phi = \frac{\partial s^\Phi}{\partial x^j} + \omega^\Phi\left(\frac{\partial}{\partial x^j}\right)s^\Phi$$
with $\Gamma_{j\alpha}^\beta = \omega_\alpha^\beta\left(\frac{\partial}{\partial x^j}\right)$.

For general section $s = \sum_{\alpha=1}^{k} s_\alpha E_\alpha$, we denote

$$\nabla_j s = \sum_{\alpha=1}^{k} (\nabla_j s_\alpha) E_\alpha$$

where we write the α's coefficient of $\nabla_j s$ denoted by $\nabla_j s_\alpha$ are given by

$$\nabla_j s_\alpha := \frac{\partial s_\alpha}{\partial x_j} + \sum_{\beta=1}^{k} \Gamma^\beta_{j\alpha} s_\beta \tag{2.2}$$

for $k = 1, \dots, n$ and $\alpha = 1, \dots, k$. The $\{\Gamma^\beta_{j\alpha}\}$ are called the *Christoeffel's symbols.*

2.2. **Affine connection recovers Ehresmann connection.** We can recover the Ehresmann connection Γ out of the affine connection ∇ as follows.

Definition 2.6. We say a germ of section $s \in \Gamma(E)$ is parallel in direction $v \in T_x M$, if $\nabla_v s = 0$, and just parallel if it is parallel in all direction of $T_x M$.

Theorem 2.7. *Let ∇ be an affine connection on the vector bundle $\pi : E \to M$. At each point $e \in E$, define the subset*

$$H_e E = \{\xi \in T_e E \mid \xi = ds(\pi(e)),\, \nabla_v s = 0,\, \forall v \in T_{\pi(e)} M\}.$$

Prove that $T_e E = H_e TE \oplus V_e TE$ where $V_e TE = \ker d_e \pi$ for the tangent map $d\pi : TE \to TM$ and the bundle $HE \to E$ is $\mathbb{R}$-equivariant, i.e., $dR_c(H_e TE) = H_{ce} TE$.

Proof. It is easy to check that indeed $H_e TE$ forms a subspace of $T_e E$. Next we show that $T_e E = H_e TE \oplus V_e TE$, i.e., $H_e TE + V_e TE = T_e E$ and $H_e TE \cap V_e TE = \{0\}$. Denote $p = \pi(e) \in M$.

Step 1: We prove $H_e TE \cap V_e TE = \{0\}$. Suppose $\xi = d_x s(v)$ for some local section s satisfying $\nabla_u s = 0$ for all $u \in T_p M$ and $d_e \pi(\xi) = 0$. But we have $0 = d_e \pi(\xi) d_p s(v) = d_p(\pi \circ s)(v) = d_p(id)(v) = v$. Hence finishes the proof.

Step 2: We next prove that for any given $\xi \in T_e E$, there exists some local section s at p such that

$$\nabla s|_p = 0,\, s(p) = e,\, \xi - d_x s(d_e \pi(\xi)) \in V_e TE = \ker d_e \pi. \tag{2.3}$$

In other words, once we find such s, ξ can be decomposed into

$$\xi = d_x s(d_e \pi(\xi)) + (\xi - d_x s(d_e \pi(\xi))) \in H_e TE + V_e TE.$$

This combined with Step 1 then proves

$$T_e E = H_e TE \oplus V_e TE.$$

To solve (2.3), we take a trivialization $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$ where we may assume U is a coordinate neighborhood at p in M. Let $\varphi = (x^1, \dots, x^n)$ be the associated coordinate chart on U centered at p. It will be enough to find a set of local sections $\{s_1, \cdots, s_n\}$ satisfying

$$\nabla s_j|_p = 0,\, s_j(p) = e,\, d\pi(ds_j(p)) = \frac{\partial}{\partial x^j},\, j = 1, \cdots, n. \tag{2.4}$$

Let s be any given local section on U and represent

$$\Phi \circ s(p) = (p, s^\Phi(p))$$

where $s^\Phi : U \to \mathbb{R}^k$ be the local representative map of s. Then

$$(\nabla_{\frac{\partial}{\partial x^k}}(s))^\Phi = \frac{\partial s^\Phi}{\partial x^k} + \omega^\Phi\left(\frac{\partial}{\partial x^k}\right) s^\Phi. \tag{2.5}$$

Therefore $\nabla s|_p = 0$ which is equivalent to

$$(\nabla_{\frac{\partial}{\partial x^k}} s)^\Phi|_p = 0$$

for for all $k = 1, \cdots, n$. This is in turn equivalent to

$$\frac{\partial s^\Phi}{\partial x^k}(p) + \omega_p^\Phi\left(\frac{\partial}{\partial x^k}\right) s^\Phi(p) = 0 \tag{2.6}$$

for each given $k = 1, \cdots, n$, which we want to solve at the given point p.

Denote $\omega^\Phi(\frac{\partial}{\partial x^j}) =: \omega_j^\Phi$ which is a linear map $E_p \to E_p$.

Recalling s^Φ is a vector valued (smooth) function $s^\Phi : U \to \mathbb{R}^k$, we can take the Tayler expansion

$$s^\Phi(y) = s^\Phi(p) + \sum_{j=1}^n x^j \frac{\partial s^\Phi}{\partial x^j} + o(|x|)$$

where $x = (x^1, \cdots, x^n)$ at p in terms of the coordinates $(x^1, \cdots, x^n)$. Therefore (2.6) can be written as

$$\frac{\partial s^\Phi}{\partial x^j}(p) + \omega_j^\Phi(p) s^\Phi(p) = 0.$$

Therefore if we set $s(p) = e$, this equation determines the first derivative of s^Φ at p by

$$\frac{\partial s^\Phi}{\partial x^j}(p) = -\omega_j^\Phi(p) e^\Phi$$

in terms of the given e at p for any section s satisfying (2.6) and $s(p) = e$. (*Note here that the value $s^\Phi(p)$ can be arbitrarily prescribed, which was set to be e as one of the standing conditions* (2.3).)

In conclusion, we take s_j by its local expression

$$s_j^\Phi(x) := e^\Phi - x^j \omega_j^\Phi(p) e^\Phi. \text{ (“ **No summation over } j \text{ involved!**”)}$$

Note that this is a function of $x = (x^1, \cdots, x^n)$ depending only on x^j. Now we need to show that s_j indeed satisfies

$$\nabla_{\frac{\partial}{\partial x^k}} s^j|_p = 0$$

for all $k = 1, \cdots, n$. But this can be easily checked by construction using the general local formula (2.5) applied to $s = s_j$, which is omitted. (**But you should check it!**)

Step 3: Prove the $\mathbb{R}$-equivariance. Let $\xi \in H_e TE$. We will show $dR_c(\xi) \in T_{ce}E$ is contained in $H_{ce}TE$. By definition, there exists a section s defined near p such that

$$s(p) = e, \, \xi = ds(p)(v)$$

for some s and $v \in T_p M$. Now consider the section $R_c(s) = cs$. Obviously we have

$$R_c(s)|_p = cs(p) = ce.$$

By the linearity of the affine connection over $\mathbb{R}$, we have

$$\nabla(cs) = c\nabla s$$

In particular if $\nabla s|_p = 0$, then $\nabla(cs)|_p = 0$. Furthermore

$$d_p(cs)(v) = d_p(R_c \circ s)(v) = d_e R_c d_p s(v) = d_e R_c(\xi)$$

which shows that $d_e R_c(\xi) \in H_{ce}TE$, and hence $dR_c(H_e TE) \subset H_{ce}TE$.

For $c \neq 0$, R_c has its inverse $(R_c)^{-1} = R_{1/c} : H_{ce}TE \to H_eTE$. By the same argument applied to $1/c$, we prove

$$dR_{1/c}(H_{ce}TE) \subset H_eTE$$

which is equivalent to $H_{ce}TE \subset dR_c(H_eTE)$. This proves $H_{ce}TE = dR_c(H_eTE)$ also when $c \neq 0$.

On the other hand, when $c = 0$, R_0 is the zero map and $ce = 0$ for any e. Considering $e = o_p$, we have $o_p = R_c(o_p)$. Denote by $\underline{0}$ the zero section of $E \to M$. Obviously $R_0 \circ \underline{0} = \underline{0}$ which again defines a section. Therefore we have $d\pi(d_{o_p}R_0 d_p\underline{0}(v)) = d\pi d(R_0 \circ \underline{0})(v) = v$ since $R_0 \circ \underline{0}$ is again a (zero) section. This equation then implies $\dim dR_0(H_{o_p}TE) \geq n = \dim T_pM$. By dimension counting, we must have $dR_0(H_{o_p}TE) = H_{o_p}TE$.

This finishes the proof. □

2.3. **Curvature.** Recall that an Ehresmann connection associates the horizontal distribution $H \subset TE$ on E. In general this distribution may not be integrable. How much non-integrable this distribution is can be quantified by the notion of curvature.

Let (E, ∇) be a vector bundle equipped with an affine connection.

Definition 2.8. The curvature R of ∇ is defined by the formula

$$R(X,Y)s = \nabla_X\nabla_Y s - \nabla_Y\nabla_X s - \nabla_{[X,Y]}s.$$

We call ∇ a *flat* connection if $R = 0$.

Proposition 2.9. *R is a tensor field in that the assignment $(X, Y, s) \mapsto R(X, Y)s$ are linear over $C^\infty(M)$ for all 3 arguments which is skew-symmetric over X, Y.*

Remark 2.10.

(1) In terms of the operation $d^\nabla : \Omega^k(E) \to \Omega^{k+1}(E)$, we have

$$d^\nabla d^\nabla = \Omega \in \Omega^2(\mathrm{End}(E))$$

where Ω is the E-valued two-form defined by its values

$$\Omega(X,Y) \cdot s := R(X,Y)s.$$

(2) In fact H is integrable if and only if the curvature of its associated affine connection is flat.

When we are given a frame $\{E_\alpha\}$ and its dual frame $\{\theta^\alpha\}$, the local representative of Ω is a matrix valued two form (Ω_α^β).

Proposition 2.11. *The curvature two-form is given by*

$$\Omega_\alpha^\beta = d\omega_\alpha^\beta + \sum_\gamma \omega_\gamma^\beta \wedge \omega_\alpha^\gamma.$$

2.4. **Metrics and Euclidean connections.**

Definition 2.12. A metric on E, denoted by $g = \langle\cdot,\cdot\rangle$, is a smooth assignment $x \mapsto g(x) = \langle\cdot,\cdot\rangle_x$ of a positive definite bilinear form on E_x, i.e., a linear map $E_x \otimes E_x \to \mathbb{R}$.

Proposition 2.13. *Any vector bundle $E \to M$ over a (paracompact) manifold carries a metric.*

Definition 2.14. A connection ∇ on (E, g) is called *Euclidean* if it preserves the metric g in that:

$$X\langle s_1, s_2\rangle = \langle \nabla_X s_1, s_2\rangle + \langle s_1, \nabla_X s_2\rangle,$$

which is equivalent $\nabla_X g = 0$.

Let $\{E_\alpha\}$ be an orthonormal frame of (E, g). Express

$$\nabla E_\alpha = \sum_\beta \omega_\alpha^\beta E_\beta$$

where ω_α^β are the associated connection one-forms.

Proposition 2.15. *A connection is Euclidean if and only if the matrix $(\omega_\alpha^\beta(X))$ is skew-symmetric for any vector field X.*

Proof. Since E_α is orthonormal, we have $\langle E_\alpha, E_\beta\rangle = \delta_{\alpha\beta}$. Therefore we have

$$\begin{aligned} 0 &= X\langle E_\alpha, E_\beta\rangle = \langle \nabla_X E_\alpha, E_\beta\rangle + \langle E_\alpha, \nabla_X E_\beta\rangle \\ &= \sum_\gamma \omega_\alpha^\gamma(X)\delta_{\beta\gamma} + \sum_\gamma \omega_\beta^\gamma(X)\delta_{\alpha\gamma} = \omega_\alpha^\beta(X) + \omega_\beta^\alpha(X). \end{aligned}$$

This finishes the proof. □

3. Riemannian metrics and Levi-Civita connection

Now let us consider the tangent bundle $TM \to M$. Let ∇ be an affine connection on TM.

Definition 3.1 (Torsion tensor)**.** The torsion of ∇ is a $(2,1)$-tensor defined by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]$$

for any vector fields X, Y on M.

In terms of the coordinate frames

$$E_j = \frac{\partial}{\partial x^j}, \quad j = 1, \ldots, n,$$

we have

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x_k} = \sum_i \Gamma_{jk}^i \frac{\partial}{\partial x^i}$$

for the associated Christoefl symbols Γ_{jk}^i.

Proposition 3.2. *An affine connection ∇ on M is torsion-free if and only if Γ_{jk}^i is symmetric for j, k.*

Because of this, a torsion-free connection on M is also called a symmetric connection.

Definition 3.3. An Euclidean connection on (M, g) is called a Riemannian connection.

Recall that for an Euclidean connection, Γ_{jk}^i is skew-symmetric for i and k.

Definition 3.4. We call a torsion-free Riemannian connection on (M, g) a Levi-Civita connection of the metric g.

Theorem 3.5 (Levi-Civita connection)**.** *Let (M, g) be a Riemnnaian manifold. Then there exists a unique torsion-free Riemannian connection.*

Proof. We have only to determine $\nabla_X Y$ for all vector fields X, Y or its pairing $\langle \nabla_X Y, Z\rangle$ for any Z. Assume that there is such a connection ∇. We will express $\langle \nabla_X Y, Z\rangle$ in terms of the known quantities and then check its torsion freeness and Rimannian property afterwards.

Start from the Riemannian property

$$\begin{aligned} X\langle Y, Z\rangle &= \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle \\ Y\langle Z, X\rangle &= \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle \\ Z\langle X, Y\rangle &= \langle \nabla_Z Y, X\rangle + \langle X, \nabla_Z Y\rangle. \end{aligned}$$

Add the first two and subtract the third therefrom, apply the torsion freeness and then get

$$\begin{aligned} & X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle \\ = & \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle + \langle \nabla_Y Z, X\rangle + \langle Z, \nabla_Y X\rangle \\ & -\langle \nabla_Z X, Y\rangle - \langle X, \nabla_Z Y\rangle \\ = & 2\langle \nabla_X Y, Z\rangle + (\langle Z, \nabla_Y X\rangle - \langle \nabla_X Y, Z\rangle) \\ & +(\langle Y, \nabla_X Z\rangle - \langle Y, \nabla_Z X\rangle) + (\langle \nabla_Y Z, X\rangle - \langle \nabla_Z Y, X\rangle) \\ = & 2\langle \nabla_X Y, Z\rangle + \langle Z, [Y, X]\rangle + \langle Y, [X, Z]\rangle + \langle [Y, Z], X\rangle. \end{aligned}$$

This suggests the definition of ∇ which must be given by

$$\begin{aligned} \frac{1}{2}\langle \nabla_X Y, Z\rangle &= X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle \\ & \quad -(\langle Z, [Y, X]\rangle + \langle Y, [X, Z]\rangle + \langle [Y, Z], X\rangle). \end{aligned} \tag{3.1}$$

We note that the right hand side is already determined when X, Y, Z and g are given and hence it uniquely determines $\nabla_X Y$.

It remains to check that the ∇ defined by this last identity satisfies the defining property of the affine connection and the resulting connection are both torsion free and Riemannian. But this immediately follows from its definition.

The connection property can be directly checked from this defining identity. Once this is done both Riemannian and torsion-freeness follow from the definition. □

We write

$$g = \sum_{i,j} g_{ij} dx^i \otimes dx^j$$

in terms of the local coordinates $(x_1, \ldots, x_n)$. Then the formula (3.1) also provides the coordinate expression of the Levi-Civita connection as follows: Set

$$Y = \frac{\partial}{\partial x^i},\ X = \frac{\partial}{\partial x^j},\ Z = \frac{\partial}{\partial x^\ell}.$$

Then it becomes

$$\langle \partial_\ell, \sum_k \Gamma_{ij}^k \partial_k\rangle = \frac{1}{2}\{\frac{\partial g_{i\ell}}{\partial x^j} + \frac{\partial g_{j\ell}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^\ell}$$

and hence

$$\sum_k \Gamma_{ij}^k g_{k\ell} = \frac{1}{2}\{\frac{\partial g_{i\ell}}{\partial x^j} + \frac{\partial g_{j\ell}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^\ell}\}.$$

If we denote by $(g^{k\ell})$ the inverse matrix of $)(g_{\ell k}$, then we obtain

Proposition 3.6. *The Christoefl symbols are determined by the metric by the formula*

$$\Gamma_{ij}^k = \frac{1}{2}\sum_\ell \{\frac{\partial g_{i\ell}}{\partial x^j} + \frac{\partial g_{j\ell}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^\ell}\} g^{\ell k}.$$

This is the classicical expression of the Christoffel symbols. In terms of the notation we introduced before, the covariant derivative $\nabla_i X$ for $X = \sum_j v^j \frac{\partial}{\partial x^j}$, is expressed as $\nabla_j X = \sum_k (\nabla_k v^j)\frac{\partial}{\partial x^k}$ with

$$\nabla_i v^j = \frac{\partial v^j}{\partial x^i} + \Gamma_{ik}^j v^k.$$

3.1. **Examples of Riemannian manifolds.**

Definition 3.7. Let (M, g) and (N, h) be two Riemannian manifolds. An *isometry* from $(M, g) \to (N, h)$ is a map $f : M \to N$ is a diffeomorphism that preserves the metric, i.e., if $g = f^*h$. More explicitly, it is an isometry if

$$h(df(v), df(w)) = g(v, w)$$

for all $x \in M$. When f is not a diffemorphism, it is called an *isometric immersion.* If f is an embedding, it is called an *isometric embedding.*

Definition 3.8 (Induced metric)**.** Let (N, h) be a Riemannian manifold. For a given immersion, $f : M \to N$, we define the *induced metric* $g = f^*h$ by

$$g(v, w) := h(df(v), df(w)). \tag{3.2}$$

The injectivity $d_x f : T_x M \to T_{f(x)} N$ makes this bilinear form is again positive definite. Obviously it is symmetric and so defines a metric.

Exercise 3.9. Consider the unit sphere $S^n \subset \mathbb{R}^{2n+1}$ and equip $\mathbb{R}^{2n+1}$ with the standard metric given by the standard constant inner product $\langle \cdot, \cdot \rangle$

$$h_x(v, w) = \langle v, w \rangle$$

at each $x \in \mathbb{R}$. Let $(y^1, \ldots, y^n)$ the coordinate functions on $U = S^n \setminus \{southpole\}$ of the stereographic projection $\varphi : U \to \mathbb{R}^n$. Express the induced metric φ^*h in this coordinates

$$g = \sum_{i,j} g_{ij} dy^i dy^j$$

i.e., express the function $g_{ij} = g_{ij}(y^1, \cdots, y^n)$ in terms of the coordinates $(y^1, \ldots, y^n)$.

Example 3.10. Consider the torus $T \subset \mathbb{R}^3$ a surface of revolution obtained by revolving the circle

$$\{(0, y, z) \mid (y-1)^2 + z^2 = \frac{1}{4}\}.$$

It carries the induced metric g. This metric is not flat: Its Gaussian curvature is not zero.

Definition 3.11 (Product metric)**.** Consider two Riemannian manifolds (M_1, g_1) and (M_2, g_2). The product manifold $M_1 \times M_2$ has a natural metric $g = g_1 \times g_2$ defined by

$$g((v_1, u_1), (v_2, u_2)) := g(v_1, v_2) + g(u_1, u_2)$$

for each pair $(v_i, u_i) \in T_{(x,y)}(M_1 \times M_2) = T_x M_1 \oplus T_y M_2$.

Example 3.12. Consider the unit circle $S^1 \subset \mathbb{R}^2$ equipped with the induced metric h. The product metric $h \times h$ on $S^1 \times S^1$ is isometric to the induced metric on $S^1 \times S^1 \subset \mathbb{R}^2 \times \mathbb{R}^2$, but not isometric to the $T \subset \mathbb{R}^3$ because this product metric is flat.

3.2. **Covariant derivative along the curve.**

Definition 3.13 (Vector field over a map $f : M \to N$)**.** We call a section $\xi \in \Gamma(f^*TN)$ a *vector field over the map* f.

Recall the definition of pull-back

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

The defining condition is equivalent to saying $e \in E_{f(x)}$. Then the tangent space $T_{(x,e)}(f^*E)$ is represented by the pair $(v, \xi) \in T_xM \times T_e(f^*E)$ satisfying

$$df(v) = d\pi(\xi), \quad v \in T_xM,\, \xi \in T_e(f^*E)$$

at (x, e). For given Ehresmann connection $\Gamma : TE = HTE \oplus VTE$, we denote by Π^v_Γ (resp. Π^h_Γ) the projections to the vertical tangent space (resp. to the horizontal space). The pull-back connection $f^*\Gamma$ is provided by the splitting

$$f^*\Gamma : T(f^*E) = HT(f^*E) \oplus VT(f^*E)$$

where we have

$$\begin{aligned} HT_{(x,e)}(f^*E) &= \{(v, \xi) \mid df(v) = d\pi(\xi),\, \xi \in HT_eE\} \\ VT_{(x,e)}(f^*E) &= \{(v, \xi) \mid df(v) = d\pi(\xi),\, \xi \in VT_eE\} \end{aligned}$$

at $(x, e) \in f^*E$. We denote by $f^*\nabla$ the associated affine connection.

Proposition 3.14. *Let $(E \to N, \nabla)$ be a vector bundle with affine connection, and let $f : M \to N$. Consider the pull-back connection $f^*\nabla =: \nabla^f$ on $f^*E =: F$. Let s be a section of f^*E and let $v \in T_xM$. Then for given section $s : M \to F$, we have*

$$\nabla^f_v s = \nabla_{df(v)}\widetilde{s}$$

for any local section $\widetilde{s}$ around $f(x)$.

Proof. Now by definition if $(x, e) = (x, \widetilde{s}(f(x))$ for some local section $\widetilde{s}$ around $f(x)$, we have

$$\begin{aligned} \nabla^f_v s &= \Pi^v_{f^*\Gamma}(d(x, e)(v)) = \Pi^v_{f^*\Gamma}(v, d(\widetilde{s} \circ f)(v)) = \Pi^v_\Gamma(d(\widetilde{s} \circ f)(v)) \\ &= \Pi^v_\Gamma(d(\widetilde{s}(df)(v))) = \Pi^v_\Gamma(d(\widetilde{s})(df(v)) = \nabla_{df(v)}\widetilde{s} \end{aligned}$$

This finishes the proof. □

We apply the above to curves $\gamma : I \to M$ be a curve. A *vector field along* γ is a section of the pull-back bundle $\gamma^T M$. In other words, it is an assignment

$$t \mapsto V(t) \in T_{\gamma(t)}M$$

A good example is the tangent vector $\dot\gamma$ which is given by

$$\dot\gamma(t) = \frac{d\gamma}{dt}(t) \in T_{\gamma(t)}M.$$

Proposition 3.15 (Covariant derivative of V along γ)**.** *Let M be equipped with an affine connection ∇. For any curve $\gamma : I \to M$, the pull-back affine connection on γ^*TM is equivalent to the following assignment*

$$V \mapsto \frac{DV}{dt}; \quad \Gamma(\gamma^*TM) \to \Gamma(\gamma^*TM)$$

called the covariant derivative along γ *which satisfies*

(1) $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$ *for any* $V, W \in \Gamma(\gamma^*TM)$,
(2) $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$ *for all function* $f : I \to \mathbb{R}$,
(3) *If V is induced by a vector field X on M locally near $\gamma(t)$, i.e., if $V(t) = X(\gamma(t))$ on $(t_0-\epsilon, t_0+\epsilon)$, then*
$$\frac{DV}{dt}(t) = \nabla_{\dot\gamma(t)}X.$$

Recall we have

$$\nabla_w s = \Pi^v_\Gamma(ds(w))$$

in general where Π^v_Γ is the vertical projection with respect to the splitting $\Gamma : TE = HTE \oplus VTE$.

Let $V(t) = X(\gamma(t))$ on $(t_0-\epsilon, t_0+\epsilon)$. We express $X = \sum_j X^j \frac{\partial}{\partial x^j}$ for a coordinate system $(x^1, \ldots, x^n)$ on U containing $\gamma(t_0)$, and assume that $\gamma(t_0-\epsilon, t_0+\epsilon) \subset U$. We can also express

$$V(t) = \sum_j v^j(t)\frac{\partial}{\partial x^j}\Big|_{\gamma(t)}$$

with $v^j(t) = X^j(\gamma(t))$. By applying the defining property of the pull-back connection $\frac{D}{dt}$ along γ, we compute

$$\begin{aligned}
\frac{D}{dt}V &= \frac{D}{dt}\left(\sum_j X^j(\gamma(t))\frac{\partial}{\partial x^j}\Big|_{\gamma(t)}\right)\\
&= \sum_j \frac{D}{dt}\left(X^j(\gamma(t))\frac{\partial}{\partial x^j}\Big|_{\gamma(t)}\right)\\
&= \sum_j \left(\frac{d(X^j\circ\gamma))}{dt}(t)\right)\frac{\partial}{\partial x^j}\Big|_{\gamma(t)} + \sum_j X^j(\gamma(t))\frac{D}{dt}\left(\frac{\partial}{\partial x^j}\Big|_{\gamma}\right).
\end{aligned}$$

However we have

$$\frac{D}{dt}\left(\frac{\partial}{\partial x^j}\Big|_{\gamma}\right)(t) = \nabla_{\dot\gamma(t)}\left(\frac{\partial}{\partial x^j}\Big|_{\gamma}\right) = \sum_i \dot\gamma^i \nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} = \gamma^i\Gamma^k_{ij}\frac{\partial}{\partial x^k}.$$

Therefore we have derived the coordinate formula

$$\frac{DV}{dt} = \dot v^j\frac{\partial}{\partial x^j} + \dot\gamma^i v^j\Gamma^k_{ij}\frac{\partial}{\partial x^k} \tag{3.3}$$

when $\dot\gamma = \dot\gamma^i\frac{\partial}{\partial x^i}$ and $V = v^j\frac{\partial}{\partial x^j}$.

Definition 3.16. A vector field V along γ is called *parallel* if $\frac{DV}{dt} = 0$ for all $t \in I$.

In coordinates, $V(t) = v^j(t)\frac{\partial}{\partial x^j}|_{\gamma(t)}$, the equation is equivalent to

$$\dot v^j + \dot\gamma^i v^k\Gamma^j_{ik}(\gamma(t)) = 0, \quad j = 1, \cdots, n. \tag{3.4}$$

This is a system of linear first order ODE which defines the *parallel transport along* γ.

4. Riemann curvature tensor

Recall that a curvature for an affine connection of a vector bundle $E \to M$ is a $C^\infty(M)$-bilinear operator $R : \Gamma(TM) \times \Gamma(TM) \to \Gamma(\mathrm{End}(E))$ defined by

$$R(X, Y)s = \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X,Y]} s$$

for X, $Y \in \Gamma(TX)$, $s \in \Gamma(E)$.

For given coordinates $\varphi = (x^1, \ldots, x^n)$ of M and a local frame $\mathcal{F} = \{e_1, \ldots, e_k\}$ on $U = \mathrm{dom}(\varphi) \subset M$, we write

$$R(\partial_i, \partial_j)e_\alpha = R_{ij\alpha}^\beta e_\beta$$

and call $R_{ij\alpha}^\beta$ *the components of* R *with respect to the coordinates* φ *and the frame* $\mathcal{F}$.

A straightforward computation shows the formula

$$R_{ij\alpha}^\beta = \frac{\partial}{\partial x^i}\Gamma_{j\alpha}^\beta - \frac{\partial}{\partial x^j}\Gamma_{i\alpha}^\beta + \Gamma_{j\alpha}^\gamma \Gamma_{i\gamma}^\beta - \Gamma_{i\alpha}^\gamma \Gamma_{j\gamma}^\beta.$$

Now we specialize to the case of $E = TM$ and the Levi-Civita connection ∇ and the coordinate frame $\mathcal{F} = \{\frac{\partial}{\partial x^i}, \ldots, \frac{\partial}{\partial x^n}\}$. Then all the Greek letters become the same roman letters and do not make difference between them.

Proposition 4.1. *Let* R *be the curvature of the Levi-Civita connection of* (M, g). *Then*

(1) *(Trilinearity) The assignment* $(X, Y, Z) \mapsto R(X, Y)Z$ *is trilinear over* $C^\infty(M)$,
(2) *(Bianchi identity) For all triples* (X, Y, Z),

$$R(X, Y)Z + R(Y, Z)X + R(Z, X)Y = 0.$$

Proof. By definition,

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

we derive

$$\begin{aligned}
& R(X, Y)Z + R(Y, Z)X + R(Z, X)Y \\
= \; & \nabla_X(\nabla_Y Z - \nabla_Z Y) + \nabla_Y(\nabla_Z X - \nabla_X Z) + \nabla_Z(\nabla_X Y - \nabla_Y X) \\
& -\nabla_{[X,Y]} Z - \nabla_{[Y,Z]} X - \nabla_{[Z,X]} Y \\
= \; & \nabla_X([Y, Z]) + \nabla_Y([Z, X]) + \nabla_Z([X, Y]) \\
& -\nabla_{[X,Y]} Z - \nabla_{[Y,Z]} X - \nabla_{[Z,X]} Y \\
= \; & [Z, [X, Y] + [X, [Y, Z]] + [Y, [Z, X]] = 0.
\end{aligned}$$

□

Using the metric, we pair $R(X, Y)Z$ with W to get a function

$$\langle R(X, Y)Z, W\rangle.$$

Proposition 4.2. *We have*

$$\begin{aligned}
\langle R(X, Y)Z, W\rangle &= -\langle R(Y, X)Z, W\rangle \\
\langle R(X, Y)W, Z\rangle &= -\langle R(X, Y)Z, W\rangle \\
\langle R(W, X)Y, Z\rangle &= \langle R(Y, Z)W, X\rangle.
\end{aligned}$$

Proof. We will just prove the third equality. We write the Bianchi identities for X, Y, Z, W by permuting them and add them up to get

$$2\langle R(W,X)Y,Z\rangle = \langle R(Y,Z)W,X\rangle$$

which finishes the proof. □

Alternatively, we can look at the quad-linear function

$$(X,Y,Z,W) = \langle R(X,Y)Z,W\rangle =: R(X,Y,Z,W)$$

In coordinates, we write

$$R_{ijkl} = \langle R(\partial_i,\partial_j)\partial_k,\partial_\ell\rangle = g_{m\ell}R^m_{ijk} :$$

Then the Bianchi identity corresponds to

$$R_{ijk\ell} + R_{jki\ell} + R_{kij\ell} = 0$$

and similarly we can write other symmetries of $R_{ijk\ell}$.

Definition 4.3 (Sectional curvatures)**.** The *section curvature function* at p is a quantity $K(\sigma)$ associated to each 2-dimensional subspace σ of T_pM defined by

$$K(\sigma) = \frac{\langle R(X,Y)Y,X\rangle}{|X\wedge Y|^2}$$

where $\sigma = \text{span}\{X,Y\}$.

By the calculation similar to the proof of the polarization identity, we have derived that the sectional curvature functions uniquely determine the Riemannian curvature.

Proposition 4.4. *The sectional curvature function determine the curvature R in the following sense: Let V be an inner vector space of dimension ≥ 2 with inner product $\langle\cdot,\cdot\rangle$. Suppose R, $R' : V\times V\times V \to V$ be two tri-linear mappings such that both R and R' satisfy the identities*

$$\langle R(X,Y)Z,W\rangle + \langle R(Y,Z)X,W\rangle + \langle R(Z,X)Y,W\rangle = 0$$

and

$$\begin{aligned}\langle R(X,Y)Z,W\rangle &= -\langle R(Y,X)Z,W\rangle\\ \langle R(X,Y)W,Z\rangle &= -\langle R(X,Y)Z,W\rangle\\ \langle R(W,X)Y,Z\rangle &= = \langle R(Y,Z)W,X\rangle.\end{aligned}$$

If $K(\sigma) = K'(\sigma)$ for all σ, then $R = R'$.

This provides the following characterization of Riemann curvature for the constant sectional curvatures.

Corollary 4.5. *Let (M,g) be a Riemannian manifold and consider a trilinear mapping $R_0' : TM\times TM\times TM \to TM$ by*

$$\langle R_0'(X,Y)Z,W\rangle = \langle X,W\rangle\langle Y,Z\rangle - \langle X,Z\rangle\langle Y,W\rangle.$$

Then M has constant sectional curvatures c if and only if $R = cR_0'$.

Proof. One easily checks that R_0' satisfies all the aforementioned symmetries. On the other hand, if R has constant sectional curvature, then we have

$$\frac{\langle R(X,Y)Y, X\rangle}{|X\wedge Y|^2} = c$$

i.e., we have

$$\langle R(X,Y)Y, X\rangle = c|X\wedge Y|^2 = c(\langle X, X\rangle\langle Y, Y\rangle - \langle X, Y\rangle^2)$$

for all X, Y. This is equivalent to saying $K(\sigma) = cK_0'(\sigma)$. Then by the proposition, we obtain $R = cR_0'$. □

5. Raising and lowering indices and contractions

If $g = \langle\cdot,\cdot\rangle$ is an inner product on a vector space V, its dual vector space V^* carries the dual inner product. It is described as follows: Consider the isomorphism

$$\widetilde{g} : V \to V^*; \quad v \mapsto \langle v, \dot{\rangle}$$

Then the dual metric is nothing but the pushforward

$$\langle\cdot,\cdot\rangle^* = \widetilde{g}_*\langle\cdot,\cdot\rangle.$$

In other words, we define it by choosing an orthonormal basis $\{E_1, \cdots, E_n\}$ of V, considering its dual basis $\{f^1, \cdots, f^n\}$ and then declaring it to be orthonormal for the inner product.

For a Riemannian metric g on TM, let $g = g_{ij}dx^i \otimes dx^j$, then its dual metric $h = (\widetilde{g})_*g$ on T^*M can be expressed as

$$h = g^{ij}\frac{\partial}{\partial x^i}\otimes\frac{\partial}{\partial x^j}$$

with (g^{ij}) is the inverse matrix of (g_{ij}). The bundle map $\widetilde{g} : TM \to T^*M$ has its inverse $T^*M \to TM$. Combining the two, we have the induced isomorphisms

$$\widetilde{g} : \mathcal{T}^r_s(M) \to \mathcal{T}^s_r(M).$$

In coordinates, the map corresponds to

$$t^{i_1\cdots i_r}_{j_1\cdots j_s} \mapsto g_{j_1k_1}\cdots g_{j_sk_s}t^{k_1\cdots k_r}_{\ell_1\cdots\ell_r}g^{i_1\ell_1}\cdots g^{i_r\ell_r}$$

when applied to the tensor field

$$T = t^{i_1\cdots i_r}_{j_1\cdots j_s}dx^{j_1}\cdots dx^{j_s}\otimes\frac{\partial}{\partial x^{i_1}}\cdots\frac{\partial}{\partial x^{i_r}}.$$

Using the multi-indices, $I = (i_1, \cdots, i_r)$ and etc, we can simplify the writing as

$$t^I_J \mapsto g_{JK}t^K_L g^{LI}.$$

Example 5.1. (1) For a vector field $X = X^i\frac{\partial}{\partial x^i}$, we have its dual field $X^\flat = \langle X, \cdot\rangle$ has the expression

$$X^\flat = X^\flat_i dx^i, \quad \text{for } X^\flat_i = g_{ij}X^j.$$

(2) For the curvature tensor $R = R^\ell_{ijk}$, the lowering operation corresponds to the tensor

$$(X, Y, Z, W) \mapsto \langle R(X,Y)Z, W\rangle \Longleftrightarrow R^\ell_{ijk} \mapsto R_{ijkl} := g_{i\ell}R^\ell_{ijk}.$$

Next, we have a canonical isomorphism

$$\mathrm{Hom}(V,V) \cong V^* \otimes V$$

and an element $A \in \mathrm{Hom}(V,V)$ is written as $A = A^i_j dx^j \otimes \frac{\partial}{\partial x^i}$. Then we have

$$\mathrm{tr}(A) = A^i_i \left(= \sum_{i=1}^n A^i_i \right).$$

For a general tensor field of type (r,s), this *contraction* operation gives rise to a tensor of type $(r-1, s-1)$

$$t^{i_1 \cdots i_r}{j_1 \cdots j_s} \mapsto t^{i_1 \cdots k \cdots j_r}_{j_1 \cdots k \cdots j_s} \left(= \sum_{k+1}^n t^{i_1 \cdots k \cdots j_r}_{j_1 \cdots k \cdots j_s} \right).$$

Definition 5.2 (Ricci curvature)**.** For each given vector fields (X,Y), we have the linear map $R_{X,Y} : W \mapsto R(W,X)Y$ which we call the *Ricci operator*. The *Ricci curvature tensor* Ric_g is defined to be

$$\mathrm{Ric}_g(X,Y) = \mathrm{tr}(R_{X,Y}) = \sum_{i=1}^n \langle R(E_i, X)Y, E_i \rangle$$

where $\{E_1, \cdot, E_n\}$ is an (and so any) orthonormal frame fields.

So $\mathrm{Ric} = \mathrm{Ric}_g$ is a covariant symmetric 2-tensor which defines a symmetric bilinear form on the tangent space T_pM at each $p \in M$.

Definition 5.3. (1) (Ricci curvature) We call

$$\frac{1}{n-1} \mathrm{Ric}(X,X)|_p$$

the *Ricci curvature* of (M,g) along X at $p \in M$.

(2) (Scala curvature) We call

$$K(p) = \frac{1}{n(n-1)} \sum_{j=1}^n \mathrm{Ric}(E_i, E_j) = \frac{1}{n(n-1)} \sum_{j=1}^n \sum_{i=1}^n \langle R(E_i, E_j)E_j, E_i \rangle$$

for a (and so any) orthonormal basis $\{E_i\}$.

The components of Ric_g are given by the contraction

$$R_{ij} := R^k_{kij} (= R_{kij\ell} g^{k\ell})$$

when we write the Ricci tensor as $\mathrm{Ric}_g = R_{ij} dx^i \otimes dx^j$. The Ricci curvature along $(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$ is given by

$$\frac{1}{n-1} \mathrm{Ric}(\partial_i, \partial_j)|_p = \frac{1}{n-1} R_{ij}.$$

Proposition 5.4. *In coordinates, we have*

$$K(p) = \frac{1}{n-1} R_{ij} g^{ij}.$$

Proof. Let $\{E_1, \cdots, E_n\}$ be an orthonormal frame and express

$$E_i = a^k_i \frac{\partial}{\partial x^k}.$$

Then we evaluate

$$\sum_{j=1}^{n} \mathrm{Ric}(E_j, E_j) = \sum_{j=1}^{n} \mathrm{Ric}(a_j^k \partial_k, a_j^\ell \partial_\ell) = \sum_j R_{k\ell} a_j^k a_j^\ell.$$

By the orthonormality $\langle E_i, E_j \rangle = \delta_{ij}$ we obtain

$$g_{k\ell} a_i^k a_j^\ell = \delta_{ij}$$

and hence

$$a_i^k a_j^\ell = g^{k\ell} \delta_{ij}.$$

Substituting this into above, we have derived

$$\sum_{j=1}^{n} R_{k\ell} a_j^k a_j^\ell = \sum_{j=1}^{n} R_{k\ell} g^{k\ell} \delta_{jj} = n R_{k\ell} g^{k\ell}$$

Therefore we have

$$K(p) = \frac{1}{n(n-1)} \sum_{j=1}^{n} \mathrm{Ric}(E_i, E_j) = \frac{1}{n-1} R_{k\ell} g^{k\ell}.$$

□

Exercise 5.5. Prove that for 3-dimensional Riemannian manifolds (M, g), the Ricci curvature determines the curvature tensor of (M, g). In coordinates, express the curvature tensor R_{ijkl} in terms of the Ricci curvatures $\{R_{ij}\}$.

6. Geodesics and exponential maps

By applying $V = \dot\gamma$, we obtain the notion of a geodesic.

Definition 6.1 (Geodesic)**.** A curve $\gamma : I \to M$ is called a *gedesic* if $\dot\gamma$ is parallel, i.e., if $\frac{D\dot\gamma}{\partial t} = 0$ on I.

In coordinates, we substitute $v^j = \dot\gamma^j$ into (3.4) and obtain

$$\ddot\gamma^k + \Gamma_{ij}^k \dot\gamma^i \dot\gamma^j = 0. \tag{6.1}$$

This is a system of nonlinear second-order ODE. To uniquely solve the equation, we need to put the initial conditions $\gamma(0) = x_0$, $\dot\gamma(0) = v$. (In general, the equation may not be able to solve for all time.)

To solve this 2nd-order ODE, we transform it to a first-order ODE by lifting the equation to an equation on TM. For this purpose, we introduce the *canonical coordinates* of TM.

Definition 6.2 (Canonical coordinates of TM)**.** The *canonical coordinates* of TM (or T^*M) associated to the coordinates $(x^1, \ldots, x^n)$ on U is the coordinates of $TU = \pi^{-1}(U)$ associated to

$$d\varphi = (x^1, \ldots, x^n, v^1, \ldots, v^n) : TU \to \varphi(U) \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n$$

where

$$x^i = s^i \circ (\pi_1 \circ d\varphi), \quad v^j = s^j \circ \pi_2, \circ d\varphi.$$

For any $v_x \in T_xM$ for $x \in U$, we can *uniquely* express

$$v_x = \sum v^j \frac{\partial}{\partial x^j}|_x.$$

Remark 6.3. (1) Similarly, we define the canonical coordinates denoted by

$$(q^1, \dots, q^n, p_1, \dots, p_n)$$

by

$$q^i = s^i \circ (\pi_{*,1} \circ d(\varphi^{-1})^*, \quad p_j = s^j \circ \pi_{*,2} \circ (d\varphi^{-1})^* :$$

For any element $\alpha \in T^*M|_U$, we can uniquely express

$$\alpha = \sum_i p_i(\alpha) dq^i|_x.$$

(2) Note that both v^j and p_j are fiberwise linear functions on $TM|_U$ and $T^*M|_U$ respectively.

In the canonical coordinates $(x^1, \dots, x^n, v^1, \dots, v^n)$ of TM, the geodesic equation can be reduced to the first-order ODE

$$\dot x^k = v^k, \quad \dot v^k = -\Gamma^k_{ij} v^i v^j : \tag{6.2}$$

For any geodesic γ, if we set $v^k(t) = \dot\gamma^k(t)$, then (γ^k, v^k), satisfy (6.2), and vice versa.

Exercise 6.4. Prove that RHS of the equation is indeed the coordinate expression of a vector field on TM, i.e., a section of $T(TM) \to TM$.(**Hint:** Prove that it satisfies the transformation rule of vector fields on $T(TM)$.)

We call the flow of this vector field on TM the *geodesic flow* of (M, g).

By the general existence, uniqueness and continuous dependence theorems on the first-order ODE, we have proven.

From now on, we will restrict ourselves to the Levi-Civita connection of the Riemannian manifold (M, g).

Lemma 6.5 (from ODE theory)**.** *For each $p \in M$, there exists an open subset $\mathcal{U} \subset TM$ with $(p, 0) \in \mathcal{U}$, $\delta > 0$ and a smooth map*

$$\varphi : (-\delta, \delta) \times \mathcal{U} \to TM$$

such that $t \mapsto \varphi(t, q, v)$ is the unique trajectory of the vector field G with $\varphi(0, q, v) = (q, v)$. Choosing V and $\epsilon_1 > 0$ small enough, we may assume

$$\{(q, v) \mid q \in V, v \in T_qM, |v| < \epsilon_1\}$$

where V is a neighborhood of $p \in M$.

Translating this into the original geodesic equation, we obtain

Proposition 6.6. *Given $p \in M$, there exist an open subset $V \subset M$ with $p \in V$, $\delta > 0$, $\epsilon_1 > 0$ and a smooth map*

$$\gamma : (-\delta, \delta) \times \mathcal{U} \to M, \quad \mathcal{U} \text{ as above}$$

such that the curve $t \mapsto \gamma(t, q, v)$, $t \in (-\delta, \delta)$ is the unique geodesic of M which at $t = 0$ passes through q with velocity v for each given $(q, v) \in TM|_V$ with $|v| < \epsilon_1$.

By the uniqueness, we have

$$\gamma(at, q, v) = \gamma(t, q, av)$$

whenever both sides are defined. Therefore by letting ϵ_1 small enough, we may assume $\delta = 2$ in the above proposition.

By choosing $\epsilon > 0$ sufficiently small so that $\gamma : (-2, 2) \times \mathcal{U} \to M$ is defined for all $q \in V$ and v with $|v| < \epsilon$.

Definition 6.7 (Exponential map)**.** Let $\mathcal{U}$ be as above. We define a map, called the *exponential map at* q, $\exp_q : T_qM \to M$ by

$$\exp_q(v) := \gamma(1, q, v)$$

when it is defined, i.e, if $|v| < \epsilon_q$.

Proposition 6.8. *Given* $q \in M$*, there exists some* $\epsilon = \epsilon_q > 0$ *such that* $\exp_q$: $B_\epsilon(0) \subset T_qM \to M$ *is a diffeomorphism of* $B_\epsilon(0)$ *onto an open subset of* M*.*

Proof. By the implicit function theorem, it is enough to prove that the derivative $d_0 \exp_q : T_0(T_qM) \to T_qM$ is an isomorphism.

By the canonical identification of $T_0(T_qM) \cong T_qM$, we have

$$d_0 \exp_q(v) = \frac{d}{dt}\Big|_{t=0} \exp_q(tv) = \frac{d}{dt}\Big|_{t=0} \gamma(1, q, tv) = \frac{d}{dt}\Big|_{t=0} \gamma(t, q, v) = v.$$

Therefore we have $d_0 \exp_q = id|_{T_qM}$ under the identification of $T_0(T_qM) \cong T_qM$. □

Example 6.9.

(1) $M = \mathbb{R}^n$: Since the covariant derivative coincides with the usual derivative, $\widetilde{\nabla_X Y} = \widetilde{X}[\widetilde{Y}]$, the geodesics are the straight lines with constant speed parameterization. Therefore $\exp_q(v) = q + v$.

(2) $M = S^n \subset \mathbb{R}^{n+1}$ equipped with the induced metric: In this case, the geodesics are the great circles parameterized by arc length.

Definition 6.10 (Arc length)**.** Let $\gamma : I \to M$ be a differentiable curve.

(1) The arc length of the curve $\gamma : I \to M$ is defined by

$$L(\gamma) = \int_{t_0}^{t_1} \left| \frac{d\gamma}{dt}(u) \right| du.$$

(2) By varying t_1, we define the arc-length function $s : I \to \mathbb{R}_+$ by

$$s(t) = \int_{t_0}^{t} \left| \frac{d\gamma}{dt}(u) \right| du =: L(\gamma|_{[t_0,t]}).$$

We can also define the arc-length of a piecewise differentiable function.

If $\dot\gamma(t) \neq 0$, then it is a local diffeomorphism at t and so the function s has a local inverse $t = t(s)$. The new parameterization $\widetilde{\gamma}(s) := \gamma(t(s))$ is called the *arc-length parameterization* of γ. With respect to this parameterization, we have

$$|\dot{\widetilde{\gamma}}(s)| = \left| \frac{d\gamma}{dt} \frac{dt}{ds} \right| \equiv 1.$$

We state a basic property of geodesics.

Proposition 6.11. *If* γ *is a geodesic, then the parameter* t *is the same with the arc-length up to a multiplication by a nonzero constant* c *and a translation.*

Proof. Enough to show that any geodesic has constant speed. □

We say a geodesic is *normalized* if it is parameterized by the arc-length.

Definition 6.12 (Distance)**.** For given points x, $y \in M$, we define the function $d : M \times M \to \mathbb{R}_+$ by

$$d(p, q) = \inf_{\gamma}\{L(\gamma) \mid \gamma(0) = p, \gamma(1) = q, \gamma \text{ piecewise differentiable}\}.$$

Theorem 6.13. *Then d defines a metric on M, i.e., it satisfies*

1. $d(p,q) = d(q,p)$,
2. $d(p,q) + d(q,r) \geq d(p,r)$,
3. $d(p,q) \geq 0$ *and equality holds only when* $p = q$.

The points (1), (2) are obvious. But we will need preliminary works to prove (3).

7. First variation of arc-length

Let $\gamma : [a,b] \to M$ be a smooth curve and recall the definition of the arc length of the curve $\gamma : [a,b] \to M$ is defined by

$$L(\gamma) = \int_a^b \left| \frac{d\gamma}{dt}(t) \right| dt.$$

This function is defined for a piecewise smooth curve and does not depend on the parametrization of the curve.

A more useful quantity than the length function for the variational study of geodesics is the (kinetic) *energy*

Definition 7.1 (Energy functional)**.** Let $\gamma : [a,b] \to M$ be a (piecewise) differentiable curve

$$E(\gamma) = \frac{1}{2} \int_a^b |\dot{\gamma}|^2 \, dt.$$

Lemma 7.2. *For any differentiable curve $\gamma : [a,b] \to M$, we have*

$$L(\gamma)^2 \leq 2|b-a| E(\gamma)$$

and the equality holds when $|\dot{\gamma}|$ is constant.

Proof. A direct calculation shows

$$L(\gamma)^2 = \left(\int_a^b \left| \frac{d\gamma}{dt}(t) \right| dt \right)^2 \leq \left| \int_a^b 1^2 \, dt \right| \cdot \left| \int_a^b |\dot{\gamma}(t)|^2 \, dt \right| = 2|b-a| E(\gamma)$$

and equality holds when $|\dot{\gamma}(t)|$ is constant. □

A *variation* of a curve γ is a a smooth map $C : (-\epsilon, \epsilon) \times [a,b] \to M$ satisfying $C(0,t) = \gamma(t)$. Denote

$$\frac{\partial C}{\partial s}(0,t) =: V(t)$$

which we call an *infinitesimal variation* along γ, which is nothing but a vector field along γ.

Writing $\gamma_s := C(s, \cdot)$, we will compute

$$\frac{d}{ds}\Big|_{s=0} E(\gamma_s) = \frac{1}{2} \int_a^b \left| \frac{\partial C}{\partial t} \right|^2 dt.$$

We compute

$$\begin{aligned} \frac{1}{2} \frac{\partial}{\partial s} \left| \frac{\partial C}{\partial t} \right|^2 &= \frac{\partial}{\partial s} \left\langle \frac{\partial C}{\partial t}, \frac{\partial C}{\partial t} \right\rangle = \left\langle \frac{D}{\partial s} \frac{\partial C}{\partial t}, \frac{\partial C}{\partial t} \right\rangle \\ &= \left\langle \frac{D}{\partial t} \frac{\partial C}{\partial s}, \frac{\partial C}{\partial t} \right\rangle = \frac{\partial}{\partial t} \left\langle \frac{\partial C}{\partial s}, \frac{\partial C}{\partial t} \right\rangle - \left\langle \frac{\partial C}{\partial s}, \frac{D}{\partial t} \frac{\partial C}{\partial t} \right\rangle. \end{aligned}$$

Here we use the following for the penultimate equality.

Exercise 7.3. Let ∇ be any torsion-free affine connection on M. Consider a smooth map $C : (-\epsilon, \epsilon) \times [a, b] \to M$ and denote by $\frac{D}{\partial t} = \nabla_{\frac{\partial C}{\partial t}}$, $\frac{D}{\partial s} = \nabla_{\frac{\partial C}{\partial s}}$ the covariant derivatives along C in the direction of $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial s}$ respectively. Prove

$$\frac{D}{\partial t}\frac{\partial C}{\partial s} = \frac{D}{\partial s}\frac{\partial C}{\partial t}.$$

Therefore we have derived

Proposition 7.4 (The first variation of the energy)**.** *Let γ be a differentiable curve and $C : (-\epsilon, \epsilon) \times [a, b] \to M$ be a variation of γ. Denote $V(t) := \frac{\partial C}{\partial t}(0, t)$. Then*

$$\frac{d}{ds}\Big|_{s=0} E(\gamma_s) = -\int_a^b \left\langle V(0,t), \frac{D\dot\gamma}{\partial t}(0,t)\right\rangle dt + \langle V(0,b), \dot\gamma(0,b)\rangle \Big|_a^b. \qquad (7.1)$$

Corollary 7.5. *Consider the variation $C = \{\gamma\}_{s\in(-\epsilon,\epsilon)}$* with the same end points. *Then we have*

$$\frac{d}{ds}\Big|_{s=0} E(\gamma_s) = -\int_a^b \left\langle V(0,t), \frac{D\dot\gamma}{\partial t}(0,t)\right\rangle dt.$$

Define the subset of maps

$$\mathcal{P} = \mathcal{P}_{[a,b]}(p, q) = \{\gamma : [a, b] \to M \mid \gamma(a) = p,\ \gamma(b) = q,\ \gamma \, smooth\}.$$

Proposition 7.6. *Let γ be a curve that minimizes the energy among curves in $\mathcal{P}$. Then γ satisfies $\frac{D}{dt}\dot\gamma = \nabla_{\dot\gamma}\dot\gamma = 0$, i.e., γ is a geodesic.*

Since the geodesic has constant speed, any energy minimizing curve is also length minimizing by the lemma. Indeed, it also uniquely determines a parametrization of constant speed.

Proof of Proposition 7.6. We start with the following lemma

Lemma 7.7. *For any given variation V along γ with $V(a) = V(b) = 0$, there exists a variation $\{\gamma_s\} \subset \mathcal{P}_{[a,b]}(p, q)$ with $\frac{\partial\gamma_s}{\partial s}\Big|_{s=0} = V(t)$.*

Proof. Just take $\gamma_s(t) = \exp_{\gamma(t)} sV(t)$. □

Since γ minimizes E on $\mathcal{P}$ the function $s \mapsto E(\gamma_s)$ takes its minimu at $s = 0$. Therefore we have

$$\frac{d}{ds}E(\gamma_s)\Big|_{s=0} = 0$$

for any variation and in particular for $\gamma_s = \exp_\gamma sV$. By the first variation formula, we derive

$$\int_a^b \langle V(t), \frac{D\dot\gamma}{dt}\rangle \, dt = 0$$

for all infinitesimal variation V with $V(a) = V(b) = 0$.

The proposition will follow from

Exercise 7.8. Let W_0 be a smooth (and so continuous) vector field along γ such that

$$\int_a^b \langle V, W_0\rangle \, dt = 0$$

for all variations V along γ with $V(a) = V(b) = 0$. Prove $W_0 = 0$.

□

Now we are ready to study the length minimizing property of geodesics. Here is a key lemma whose proof is an application of the first variation formula.

Theorem 7.9 (Gauss lemma)**.** *If $\rho(t) = tv$ is a ray through the origin in T_pM in the direction of v and if $w \in T_v(T_pM)$ is perpendicular to $\dot\rho(1) \in T_v(T_pM)$, then*

$$d_v \exp_p(\dot\rho(1)) \perp d_v \exp_p(w).$$

Proof. Let $v(s)$ be a curve in T_pM such that $v(0) = v$, $v'(0) = w$ and $|v(s)| = const.$. Define

$$\alpha(s,t) = \exp_p(\rho_s(t)), \quad \rho_s(t) := tv(s)$$

where $\rho_s : [0,1] \to T_pM$ is the ray segment from $\vec{0}$ to $v(s)$ in T_pM. We put

$$\gamma_v(t) := \exp_p(tv)$$

which is the geodesic with $\dot\gamma_v(0) = v$.

Then $\gamma_{v(s)}$ satisfy

(1) $\gamma_{v(s)}(0) = p$ and so $\frac{\partial\alpha}{\partial s}(s,0) = 0$ for all s,
(2) $\dot\gamma_{v(s)}(t) = d_{\rho_s(t)} \exp_p(\dot\rho_s(t))$,
(3) $|\dot\gamma_{v(s)}(t)| = |v(s)| = |v|$ for all s, t since $\gamma_{v(s)}$ is a geodesic with initial velocity $v(s)$.

Therefore the energy function

$$s \mapsto \frac{1}{2}\int_0^1 |\dot\gamma_{v(s)}|^2\, dt = \frac{1}{2}\int_0^1 \left|\frac{\partial\alpha}{\partial s}\right|^2 dt$$

is a constant function. By the first variation formula, we derive

$$\begin{aligned} 0 &= \frac{d}{dt}\Big|_{s=0} \frac{1}{2}\int_0^1 \left|\frac{\partial\alpha}{\partial s}\right|^2 dt \\ &= \left\langle \frac{\partial\alpha}{\partial s}(0,t), \dot\gamma_{v(0)}(t)\right\rangle\Big|_0^1 = \left\langle \frac{\partial\alpha}{\partial s}(0,1), \dot\gamma_v(t)\right\rangle \\ &= \langle d_v \exp_p(w), d_v \exp_p(v)\rangle. \end{aligned}$$

Here the term $\dot\rho_v(1) = v$ through the canonical identification $T_v(T_pM) \cong T_pM$. □

Exercise 7.10. Let N, $\overline{N}$ be submanifolds of a Riemannian manifold (M,g) and let $\gamma : [0,1] \to M$ be a geodesic from N to $\overline{N}$. Prove that γ satisfies

$$\dot\gamma(0) \perp T_{\gamma(0)}N, \quad \dot\gamma(1) \perp T_{\gamma(1)}\overline{N}.$$

For a piecewise smooth curve γ with possible kinks at

$$a = t_0 < t_1 < \ldots < t_{N-1} < t_N = b$$

we consider the piecewise smooth variation V along γ which is continuous especially at t_i's. Then

Proposition 7.11. *Let γ and V as above. Then*

$$\frac{d}{ds}\Big|_{s=0} E(\gamma_s) = -\int_a^b \langle V, \frac{D\dot\gamma}{dt}\rangle\, dt - \sum_{i=0}^{N-1} \langle V, \dot\gamma(t_i + 0) - \dot\gamma(t_i - 0)\rangle.$$

We denote

$$\frac{d}{ds}\Big|_{s=0} E(\gamma_s) =: \delta E(\gamma)(V).$$

Corollary 7.12. *If γ is any piecewise smooth energy minimizing curve with fixed end points, it is a smooth geodesic.*

Proof. Since γ is energy minimizing, we have

$$\delta E(\gamma)(V) = 0$$

for all such variation.

By considering the infinitesimal variation V such that

$$V(t_i) = 0$$

for all i, we show $-\int_a^b \langle V, \frac{D\dot\gamma}{dt}\rangle\, dt = 0$. This proves that

$$\frac{D\dot\gamma}{dt} = 0$$

and so a geodesic on (t_i, t_{i+1}) for all i. By the hypothesis, γ is continuous.

By substituting this into the variation formula, we have derived

$$0 = \delta E(\gamma)(V) = -\sum_{i=0}^{N-1} \langle V, \dot\gamma(t_i + 0) - \dot\gamma(t_i - 0)\rangle.$$

for all infinitesimal variation of γ. In particular by considering a variation with

$$V(t_i) = \gamma(t_i + 0) - \dot\gamma(t_i - 0),$$

we obtain

$$0 = -\sum_{i+1}^{N-1} |\gamma(t_i + 0) - \dot\gamma(t_i - 0)|^2$$

and hence $\gamma(t_i + 0) = \dot\gamma(t_i - 0)$ must hold.

Once this is proved, we derive that its second derivative continuously extends to the points t_i by the geodesic equation

$$\ddot\gamma^i + \Gamma^i_{jk}(\gamma^1, \ldots, \gamma^n)\dot\gamma^i\dot\gamma^j = 0.$$

Similarly by differentiating the equation, we derive that all of its higher derivatives continuously extends to whole interval $[a, b]$. □

Remark 7.13. Such an argument is called the bootstrap argument to increase the regularity of a solution of a differential equation.

8. Geodesic normal coordinates and geodesic balls

Definition 8.1 (Normal coordinates)**.** A local coordinate system $(x^1, \ldots, x^n)$ is called *normal coordinates* at p, if it satisfies

(1) $x^i(p) = 0$,
(2) $g_{ij}(p) = \delta_{ij}$,
(3) $\frac{\partial g_{ij}}{\partial x^k}(p) = 0$ for all i, j, k.

In this coordinates, we can Taylor-expand g_{ij} as

$$g_{ij}(x) = \delta_{ij} + \text{“2nd or higher order terms of } x^i\text{”}.$$

Definition 8.2 (Geodesic normal coordinates)**.** Let $p \in M$ and choose an orthonormal coordinates $(\overline{x}^1, \ldots, \overline{x}^n)$ at 0 of $T_pM \cong \mathbb{R}^n$. Define a coordinate system by

$$x^i = \overline{x}^i \circ \exp_p^{-1}.$$

We call this the *geodesic normal coordinates* at p.

Proposition 8.3. *Geodesic normal coordinates are indeed normal.*

Proof. (1) is obvious. For (2), we compute

$$\begin{aligned} g_{ij}(p) &= \langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\rangle_p = \left\langle d_0 \exp_p(\frac{\partial}{\partial \overline{x}^i}\Big|_0), d_0 \exp_p(\frac{\partial}{\partial \overline{x}^j}\Big|_0)\right\rangle \\ &= \left\langle \frac{\partial}{\partial \overline{x}^i}\Big|_0, \frac{\partial}{\partial \overline{x}^j}\Big|_0 \right\rangle = \delta_{ij} \end{aligned}$$

For the proof of (3), we compute

$$\frac{\partial g_{ij}}{\partial x^k}(p) = \frac{\partial}{\partial x^k}\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\rangle|_p = \langle \nabla_k \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\rangle|_p + \langle \frac{\partial}{\partial x^i}, \nabla_k \frac{\partial}{\partial x^j}\rangle|_p.$$

Now we claim

$$\nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^i}\Big|_p = 0 \tag{8.1}$$

for all i, k.

We first recall

$$\frac{\partial}{\partial x^k}\Big|_p = d_0 \exp(\frac{\partial}{\partial \overline{x}^k}).$$

Since $t \mapsto \exp_p(t\frac{\partial}{\partial \overline{x}^i}) =: \gamma_i$ is a geodesic and $d_0 \exp(\frac{\partial}{\partial \overline{x}^k}) = \frac{d\gamma}{dt}(0)$, $\frac{D\dot\gamma_i}{dt} = 0$. On the other hand, by definition we have

$$\dot\gamma_i = \frac{\partial}{\partial x^i}\Big|_{\gamma_i(t)}.$$

Therefore

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^i}\Big|_p = \nabla_{\dot\gamma_i(0)} \frac{\partial}{\partial x^i} = \frac{D\dot\gamma_i}{dt}(0) = 0$$

for all i. Since $\frac{\partial}{\partial x^k}|_p + \frac{\partial}{\partial x^j}|_p$ is the tangent vector at 0 for the geodesic

$$t \mapsto \exp_p t\left(\frac{\partial}{\partial \overline{x}^k} + \frac{\partial}{\partial \overline{x}^j}\right)$$

we also obtain

$$\nabla_{\frac{\partial}{\partial x^k} + \frac{\partial}{\partial x^j}}(\frac{\partial}{\partial x^k} + \frac{\partial}{\partial x^j})\Big|_p = 0$$

for all k, j. This proves $\nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j}\Big|_p + \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k}\Big|_p = 0$. By the torsion freeness of the Levi-Civita connection and $[\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^j}] = 0$, we obtain

$$\nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} + \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^k} = 2\nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j}.$$

Combining the above discussions, we have proved the claim. □

Exercise 8.4. Prove that Condition (3) for the normal coordinate is equivalent to the vanishing of the Christoeffl symbols:

$$\Gamma_{ij}^k(p) = 0$$

for all i, j, k.

Corollary 8.5 (Geodesic polar coordinates). *Let $\overline{r}, \overline{\theta}_1, \dots, \overline{\theta}_{n-1}$ be a spherical coordinates of $T_pM \cong \mathbb{R}^n$ and define*

$$r = \overline{r} \circ \exp_p^{-1}, \quad \theta_j = \overline{\theta}_j \circ \exp_p^{-1}, \quad j = 1, \dots, n-1.$$

Then we can express the metric

$$ds^2 = dr^2 + r^2 \sum_{j=1}^{n-1} h_{ij}(r,\theta) d\theta^i d\theta^j$$

for a positive definite matrix $(h_{ij}(r,\theta))$ such that

$$h_{ij}(r,\theta) = h_{ij}^0 + O(r) \tag{8.2}$$

with $h_{ij}^0 \equiv h_{ij}(0,\cdot)$.

Proof. By the Gauss lemma, we have $\frac{\partial}{\partial r} \perp \frac{\partial}{\partial \theta^i}$ for all i. i.e., $\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^i} \rangle = 0$ for all i.

Next, we claim

$$\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial r} \rangle = \langle \frac{\partial}{\partial \overline{r}}, \frac{\partial}{\partial \overline{r}} \rangle = 1.$$

We first prove $\frac{\partial}{\partial r} = d\exp_p(\frac{\partial}{\partial \overline{r}})$. By definition, we have $dr = d\overline{r} \circ d\exp_p^{-1}$ and hence

$$1 = dr\left(\frac{\partial}{\partial r}\right) = d\overline{r}\left(d\exp_p^{-1} \frac{\partial}{\partial r}\right).$$

Since $\left(d\exp_p^{-1} \frac{\partial}{\partial r}\right) \perp \frac{\partial}{\partial \overline{\theta}^j}$ again by the Gauss lemma, we prove $\frac{\partial}{\partial r} = d\exp_p(\frac{\partial}{\partial \overline{r}})$.

We also have

$$\frac{1}{r^2}\langle \frac{\partial}{\partial \theta^i}, \frac{\partial}{\partial \theta^j} \rangle =: h_{ij}(r,\theta)$$

for some positive definite symmetric matrix $(h_{ij}(r,\theta))$ for $r > 0$. Since

$$\langle \frac{1}{r}\frac{\partial}{\partial \theta^i}, \frac{1}{r}\frac{\partial}{\partial \theta^j} \rangle\Big|_{p=0} = \langle \frac{1}{\overline{r}}\frac{\partial}{\partial \overline{\theta}^i}, \frac{1}{\overline{r}}\frac{\partial}{\partial \overline{\theta}^j} \rangle\Big|_0 = h_{ij}^0$$

where (h_{ij}^0) is a positive definite $(n-1) \times (n-1)$ matrix. This proves

$$\frac{1}{r^2} h_{ij}(r,\theta) = h_{ij}^0 + O(r).$$

□

Suppose that

$$\exp_p : B_p(\epsilon) \to M$$

is a diffeomorphism onto its image. We denote $B_p(\epsilon) = \exp_p(B_0(\epsilon))$ and call it a *normal ϵ-ball* at p. Any open neighborhood $W \subset B_p(\epsilon)$ of p a *normal neighborhood* of p.

Definition 8.6 (Injectivity radius). Define

$$\mathrm{inj}(p) = \sup\{\epsilon \mid \exp_p : B_0(\epsilon) \to M \text{ is injective}\}$$

is injective. We call $\mathrm{inj}(p)$ the *injectivity radius of p*. We then call

$$\mathrm{inj}(M) := \inf_{p \in M} \mathrm{inj}(p)$$

the *injectivity radius* of M.

We mention that $\mathrm{inj}(M)$ could be 0.

Example 8.7. The injectivity radius of $\mathbb{R}^n$ is ∞. The injectivity radius of $S^n(1)$ is 2π.

Proposition 8.8. *For each point $p \in M$, there exists $\epsilon_p > 0$ such that any geodesic issued at p ending at $q \in B_p(\epsilon_p)$ is the shortest curve joining p and q in $B_p(\epsilon_p)$. In particular, geodesics are locally distance minimizing.*

Proof. Put $\epsilon_p = \mathrm{inj}(p)$. Let $q \in B_p(\epsilon_p)$ and $v = \exp_p^{-1}(q) \in T_pM$. Consider the geodesic

$$\gamma_v : [0,1] \to M, \quad \gamma_v(t) = \exp_p(tv).$$

Then we have $\gamma_v(0) = p$, $\gamma_v(1) = q$ and $L(\gamma_v) = |v|$.

Let $c : [0,1] \to U$ be any curve $c(0) = p$, $c(1) = q$. In geodesic normal coordinates at p, we express

$$\dot c = \frac{d(r\circ c)}{dt}\frac{\partial}{\partial r} + \sum_{j=1}^{n-1} \frac{d(\theta^j \circ c)}{dt}\frac{\partial}{\partial \theta^j}.$$

Write $c^r = r \circ c$ and $c^j = \theta^j \circ c$ for $j = 1, \ldots, n-1$. Therefore we obtain

$$\begin{aligned} L(c) &= \int_0^1 |\dot c(t)|\, dt = \int_0^1 \sqrt{(\dot c^r)^2 + \sum_{j=1}^{n-1} h_{ij}\dot c^i \dot c^j}\, dt \\ &\geq \int_0^1 \sqrt{(\dot c^r)^2}\, dt \geq \int_0^1 \dot c^r\, dt = c^r(1) - c^r(0) = |v| \end{aligned}$$

for $v = \exp_p^{-1} q$.

The first equality holds when $c^j = \theta^j \circ c$ are constants and the second equality holds when $|\dot c^r| = \dot c^r$ i.e., $c^r = r \circ c$ is monotonically increasing. Therefore we have

$$L(c) \geq L(\gamma_v)$$

for any smooth curve joining p and q *inside* $B_p(\epsilon_p)$ and equality holds only when $c = \gamma_v$. □

Corollary 8.9. *The distance function d is nondegenerate, i.e., $d(p,q) = 0$ iff $p = q$.*

Proof. Let p be given. Suppose $q \neq p$. We will show that $d(p,q) > 0$. We consider two cases, $q \in B_p(\epsilon)$ and $q \notin B_p(\epsilon_p)$ separately.

First consider the case $q \in B_p(\epsilon)$. Then denote $v = \exp_p^{-1}(q) \in T_pM$ which is well-defined. Suppose that $c : [0,1] \to M$ is any piecewise smooth curve with $c(0) = p$ and $c(1) = q$. If $\operatorname{Image} c \subset B_p(\epsilon_p)$, then we have

$$L(c) \geq L(\gamma_v) = |v| > 0. \tag{8.3}$$

If $\operatorname{Image} c \not\subset B_p(\epsilon_p)$, , let t_0 be the first exit time from $B_p(\epsilon_p)$ such that

$$c([0,t_0]) \subset B_p(\epsilon_p).$$

Then the above proof shows that

$$L(c) \geq L(c|_{[0,t_0)}) \geq L(\gamma_{v'}) = \epsilon_p \tag{8.4}$$

where $v' = \exp_p^{-1}(c(t_0))$.

If $q \in B_p(\epsilon_p)$, then any curve $c : [0,1] \to M$ connecting p and q must not be contained in $B_p(\epsilon_p)$. By the same argument as the second half of the case $q \in B_p(\epsilon_p)$, we conclude $L(c) \geq \epsilon_p$ again.

Combining the above three, we have proved that either $L(c) \geq \epsilon_p$ or $L(c) \geq |\exp_p^{-1}(q)|$. This proves that if $d(p,q) = 0$, then $p = q$. □

Next we ask the following two questions: Let $p \in M$ be given and let $\epsilon_p = \mathrm{inj}(p)$.

(1) Is there a geodesic between any two given points $q,\ q' \in B_p(\epsilon_p)$?
(2) If so, is the geodesic fully contained in $B_p(\epsilon)$?

The answer is no in general, unless we make the radius of the normal ball smaller.

Proposition 8.10. *For any $p \in M$, there is a open neighborhood W of p and $\delta = \delta_p > 0$ such that for each $q \in W$, $\exp_q : B_0^q(\delta) \to T_qM \to M$ is a diffeomorphism onto its image, and $W \subset \exp_q(B_0^q(\delta)) = B_q(\delta)$.*

Proof. Let $\epsilon > 0$ and $V \subset M$ be an open neighborhood p. Consider $\mathcal{U} \subset TM$ given by

$$\mathcal{U} = \{(q, v) \mid q \in V,\ v \in T_qM, |v| < \epsilon\} = \bigcup_{q \in V} \{q\} \times B_0^q(\epsilon) \subset TM.$$

Define a map $\mathrm{Exp} : \mathcal{U} \to M \times M$ by

$$\mathrm{Exp}(q, v) = (q, \exp_q v).$$

We compute

$$d_{(p,0)} \mathrm{Exp} : T_{(p,0)}(TM) \cong T_pM \oplus T_pM \to T_pM \oplus T_pM$$

by computing its partial derivatives in the horizontal and vertical directions:

$$d_1 \mathrm{Exp}_{(p,0)} = id|_{T_pM} \oplus id|_{T_pM}, \quad d_2 \mathrm{Exp}_{(p,0)} = 0 \oplus id|_{T_pM},$$

i.e., we have

$$d\,\mathrm{Exp}_{(p,0)} = \begin{pmatrix} id & 0 \\ id & id \end{pmatrix}.$$

Therefore $d_{(p,0)} \mathrm{Exp}$ is an isomorphism. By the inverse function theorem, there is a neighborhood $\mathcal{U}' \subset \mathcal{U}$ of $(p, 0)$ such that $\mathrm{Exp}|_{\mathcal{U}'}$ is a diffeomorphism onto its image. Since M is locally compact, we can choose $\mathcal{U}'$ as a box neighborhood of the type

$$\mathcal{U}' = \{(q, v) \mid q \in V', \quad |v| < \delta\}$$

for some smaller neighborhood $V' \subset V$ of p. Now we choose a sufficiently small $W \subset V'$ so that $W \times W \subset \mathrm{Exp}(\mathcal{U}')$. Then the choice we made for W and $\delta > 0$ will do our purpose. □

The proposition tells us in particular that there is a unique geodesic between $q_1,\ q_2 \in W$ contained in $B_{q_1}(\delta) \cap B_{q_2}(\delta)$. In general, the whole geodesic may not be completely contained in W, i.e., W may not be *convex*.

Definition 8.11. An open set $U \subset M$ is called *(geodesically) convex* if for any $p,\ q \in U$, there exists a unique minimal unit speed geodesic joining p and q completely contained in U. It is called *strongly convex* if the same holds for any $q_1,\ q_2 \in \overline{U}$ except possibly the end points of the geodesic.

Theorem 8.12. *For each $p \in M$, there exists $\bar\epsilon_p > 0$ such that the geodesic $\bar\epsilon_p$-ball $B_p(\bar\epsilon_p)$ is strongly convex. Moreover $\bar\epsilon_p$ can be chosen uniformly over on a compact neighborhood of p.*

Corollary 8.13. *Any manifold has good covering $\{U_\alpha\}$: Any finite intersection of U_α's are contractible.*

Proof of Theorem 8.12. Fix a geodesic normal coordinates $\varphi = (x^1, \ldots, x^n)$ centered at p. We define a function $\| \cdot \|_\varphi : TM|_{B_p(\epsilon_0)} \to \mathbb{R}$ by

$$\|v\|_\varphi := \sum_{i=1} (\dot{x}^i(v))^2$$

for $v = \sum_{i=1}^n \dot{x}^i(v) \frac{\partial}{\partial x^i}$. On the other hand, we have

$$\|v\| = \sqrt{g_{ij} \dot{x}^i(t) \dot{x}^j(t)}$$

for $g_{ij} = \delta_{ij} + O(r^2)$. Therefore we can choose $\epsilon_0 > 0$ such that

$$\frac{\epsilon_0}{5} < 1$$

and there is some $K \geq 1$ depending only on ϵ_0 such that

$$\frac{1}{K} \|v\|_\varphi \leq \|v\| \leq K \|v\|_\varphi$$

for all $v \in TM|_{B_p(\epsilon_0)}$. (Here we use the fact that any two norms are equivalent on a given vector space.)

Then we choose $0 < \epsilon_p < \frac{\epsilon_0}{5} < 1$ so small that

(1) For any $q \in B_p(\epsilon_p)$, geodesics issued at q are distance minimizing in $B_q(3\epsilon_p)$. Choose $3\epsilon_p < \delta$ in the above proposition.
(2) $\|g_{ij} - \delta_{ij}\|_{C^0; B_p(3\epsilon_p)} < \frac{1}{10K^2}$,
(3) $\|\Gamma_{ij}^k\|_{C^0; B_p(3\epsilon_p)} < \frac{1}{10K^2 n^3}$.

We claim that with these choices $B_p(\epsilon_p)$ is indeed strongly convex.

Suppose q, $q' \in B_p(\epsilon_p)$ and let $\gamma : [0, a] \to M$ be a unit speed distance minimizing geodesic with $\gamma(0) = q$, $\gamma(a) = q' \in B_p(\epsilon_p)$. By (1), Image $\gamma \subset B_q(3\epsilon_p)$. Then we have

$$L(\gamma) \leq d(p, q) + d(p, q') \leq 2\epsilon_p$$

and hence $a \leq 2\epsilon_p$.

By the choices we made above, $B_q(3\epsilon_p) \subset B_p(5\epsilon_p)$ and

$$\gamma(t) \in B_q(3\epsilon_p)$$

for all $t \in [0, a]$, and $d(p, \gamma(0)$, $d(p, \gamma(1)) \leq \epsilon_p$.

Now we put

$$f(t) = (d(p, \gamma(t))^2 = \gamma^1(t)^2 + \ldots + c^n(t)^2, \; \gamma^i = x^i \circ \gamma.$$

It suffices to show that

$$f''(t) > 0 \quad \text{for all } t \in (0, a).$$

We compute

$$\frac{1}{2} f''(t) = \sum_{i=1} \dot{\gamma}^i(t)^2 + \sum_{j=1} \gamma^i(t) \ddot{\gamma}^i(t)$$

and

$$\ddot{\gamma}^j(t) + \Gamma_{km}^j(\gamma(t)) \dot{\gamma}^k(t) \dot{\gamma}^m(t) = 0.$$

Therefore

$$\frac{1}{2} f''(t) = \|\dot{\gamma}\|_\varphi^2 - \sum_j \gamma^j \left(\Gamma_{km}^j(\gamma(t)) \dot{\gamma}^k(t) \dot{\gamma}^m(t) \right).$$

Then

$$\begin{aligned}\frac{1}{2}f''(t) &\geq \left(\|\dot\gamma\| - \frac{1}{10K^2}K^2\|\dot\gamma(t)\|^2\right)^2 - \left|\sum_j \gamma^j(t)\left(\Gamma^j_{km}(\gamma(t))\dot\gamma^k(t)\dot\gamma^m(t)\right)\right| \\ &\geq \left(\frac{81}{100} - -3\epsilon_p\frac{1}{10K^2n^3}(K^2n^3)\right)\|\dot\gamma(t)\|^2 \\ &= \frac{81}{100} - \frac{3\epsilon_p}{10} > \frac{81}{100} - \frac{3}{10} > 0.\end{aligned}$$

This finishes the proof. □

9. Hopf-Rinow Theorem

Question 9.1. (1) When is the exponential map $\exp_p$ defined on whole T_pM?
(2) Let p, $q \in M$. Is it possible to join p and q by a length minimizing geodesic?

Theorem 9.2 (Hopf-Rinow)**.** *Let (M, g) be a connected Riemannian manifold. Let $p \in M$ be given. TFAE:*

(1) $\exp_p$ *is defined on all of T_pM.*
(2) *The closed and bounded sets of M are compact.*
(3) *M is Cauchy complete with respect to the Riemannian distance.*
(4) *M is geodesically complete: any geodesic is defined for all time.*

In addition, any of the above implies

(5) *For any $q \in M$, there exists a gedesic γ joining p to q with $L(\gamma) = d(p, q)$.*

Proof. We will prove the theorem in the following chain of logical sequence:

$$1) \Longrightarrow 5), \quad 1)\,\&\,5) \Longrightarrow 2) \Longrightarrow 3) \Longrightarrow 4) \Longrightarrow 1).$$

1) $\Longrightarrow$ 5): Let $d(p, q) = r$ and let $B_p(\delta)$ be a geodesic normal ball at p with $S = S_p(\delta) = \partial B_p(\delta)$. Let $x_0 \in S$ be a point where the continuous function $x \mapsto d(q, x)$, $x \in S$ attains a minimum. Then $x_0 = \exp_p(\delta v)$ for some $v \in T_pM$ with $|v| = 1$. Then we consider the geodesic $\gamma = \gamma_v$ defined by $\gamma_v(s) = \exp_p sv$ which is defined for all s by the Hypothesis 1). We will show $\gamma_v(r) = q$. To prove this, we consider the equation

$$d(\gamma_v(s), q) = r - s \tag{9.1}$$

for $s \in [\delta, r]$. Define the subset

$$A = \{s \in [\delta, r] \mid (9.1) \text{ holds}\}.$$

We will show $A = [\delta, r]$ by proving that A is nonempty, open and closed in $[0, r]$.

Since curve connecting p and q intersect $\partial B_p(\delta)$, we have

$$d(p, q) = \min_{p' \in \partial B_p(\delta)} (d(p, p') + d(p', q)) = \delta + d(p_0, q)$$

and hence $d(p_0, q) = d(p, q) - \delta = r - \delta$. This implies $\delta \in A$ and so $A \neq \emptyset$. Since d, γ are continuous, A is also closed.

To prove A is open in $[\delta, r]$, let $s_0 \in A$. If $s_0 = r$, we will be done. Suppose $\delta \leq s_0 < r$. Consider geodesic normal balls $B_{\gamma(s_0)}(\delta')$ for all sufficiently small $\delta' > 0$. Similarly as before we put $S' = \partial B_{\gamma(s_0)}(\delta')$ and let $x_0' \in S'$ be a point at which $d(x, q)$ attains a minimum in S'.

Again we gave

$$d(\gamma(s_0), q) = \min_{x \in S'}(d(\gamma(s_0), x) + \min_{x \in S'} d(x, q) = \delta' + d(x_0', q),$$

from which we derive

$$d(x_0', q) = r - s_0 - \delta' = r - (s_0 + \delta').$$

Therefore

$$d(p, x_0') \geq d(p, q) - d(x_0', q) = r - (r - (s_0 + \delta')) = s_0 + \delta'.$$

On the other hand, the path from p to $x_0 = \gamma(s_0)$ followed by the minimal geodesic from x_0 to x_0' has length $s_0 + \delta'$ and hence we have proved

$$d(p, x_0') = s_0 + \delta'.$$

This implies that the concatenated path has length same as the distance between p and x_0', it must be indeed smooth by Corollary 7.12, and in particular $x_0' = \gamma(s_0 + \delta')$ for all sufficiently small $\delta' > 0$ and hence A is open. This finishes the proof of 1) $\Longrightarrow$ 5).

1) & 5) $\Longrightarrow$ 2): Let $A \subset M$ be closed and bounded. Since A is bounded, we may assume $A \subset B_p^d(R)$ for some metric ball of radius R at p. Then we have

$$B_p^d(R) \subset \exp_p(\overline{B_0(R)})$$

for the Euclidean ball $B_0(R) \subset T_pM$. But $\overline{B_0(R)}$ is compact by Heine-Borel. Then since $\exp_p$ is continuous, $\exp_p(\overline{B_0(R)})$ is compact and hence the closed set $A \subset \exp_p(\overline{B_0(R)})$ is compact.

2) $\Longrightarrow$ 3): is easy since any Cauchy sequence is bounded.

3) $\Longrightarrow$ 4): Let γ be a geodesic with $\gamma(0) = p$ and $\dot\gamma(0) = v$ and define

$$s_0 = \sup_s \{s \in \mathbb{R}_+ \mid \gamma(t) \text{ is defined for all } 0 \leq t \leq s\}.$$

If $s_0 = \infty$, we are done. So suppose $s_0 < \infty$.

Take an increasing sequence $s_j \nearrow s_0$. Then $\gamma(s_j)$ is a Cauchy sequence and so converges to $p_0 \in M$. Let W, $\delta > 0$ be the pair of a neighborhood of p and a positive constant fo p_0 given as in Proposition 8.10.

Take sufficiently large $N > 0$ such that if n, $m > N$ with $n > m$, then

$$\gamma(s_n), \gamma(s_m) \in W, \quad s_n \geq s_m$$

and $|s_n - s_m| < \frac{\delta}{4}$. Then there exists a unique geodesic $\overline{\gamma}$ between $\gamma(s_n)$ and $\gamma(s_m)$ whose length is $\leq \delta$.

Since $\exp_{\gamma(s_m)}$ is a diffeomorphism on $B_0(\delta) \subset T_{\gamma(s_m)}M$ and $\exp_{\gamma(s_m)}(B_0(\delta)) \supset W$, $\overline{\gamma}$ indeed extends γ to the interval $[0, s_m + \delta)$. Since $s_m \to s_0$ as $m \to \infty$, we have $s_m + \frac{\delta}{2} > s_0$ eventually, a contradiction to the definition of s_0.

4) $\Longrightarrow$ 1): Obvious. □

Definition 9.3. We call (M, g) a *complete* Riemannian manifold if it satisfies any (and so all) of 1) - 4).

10. Classification of constant curvature surfaces

Example 10.1. (1): $S^2(1) \subset \mathbb{R}^3$ equipped with the induced metric. In geodesic polar coordinates around the south pole S with $r(S) = 0$, the metric is given by

$$ds^2 = dr^2 + \sin^2 r d\theta^2, \quad 0 < r \le \pi, 0 \le \theta < 2\pi,$$

and in the stereographic coordinates on $S^2(1) \setminus \{N\}$, we have $ds^2 = \frac{dx^2+dy^2}{1+\frac{1}{4}r^2}$. This metric has its sectional curvature $K = 1$.
(2): $\mathbb{R}^2$ with $ds^2 = dx^2 + dy^2$. In geodesic polar coordinates,

$$ds^2 = dr^2 + r^2 d\theta^2, \quad K = 0.$$

(3): $\mathbb{H}^2 = \{(x, y) \mid r > 0\}$ equipped with

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

In the disc model, it is given by

$$ds^2 = \frac{dx^2 + dy^2}{1 - \frac{1}{4}r^2}.$$

In geodesic polar coordinates

$$ds^2 = dr^2 + \sinh^2 r d\theta^2, \quad K = -1.$$

Theorem 10.2. *Any constant curvature surfaces with $K = 1, 0 - 1$ are locally isometric to one of the aboves.*

Recall that in geodesic polar coordinates (r, θ), $\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} = 0$. We will also need to compute $\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta}$.

Lemma 10.3. *Write $J = \frac{\frac{\partial}{\partial \theta}}{|\frac{\partial}{\partial \theta}|}$. Then J is parallel along the radial curve, i.e.,*

$$\nabla_{\frac{\partial}{\partial r}} J = 0.$$

Proof. It is enough to prove

$$\langle \nabla_{\frac{\partial}{\partial r}} J, J \rangle = 0 = \langle \nabla_{\frac{\partial}{\partial r}} J, \frac{\partial}{\partial r} \rangle.$$

The first follows since $\langle J, J \rangle = 1$. For the second, we compute

$$\langle \nabla_{\frac{\partial}{\partial r}} J, \frac{\partial}{\partial r} \rangle = \frac{\partial}{\partial r} \langle J, \frac{\partial}{\partial r} \rangle - \langle J, \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} \rangle = 0.$$

□

Corollary 10.4. *For any function $g \in C^\infty(M)$, we have*

$$\nabla_{\frac{\partial}{\partial r}} (g \cdot J) = \frac{\partial g}{\partial r} J.$$

Proof of 10.2. Consider a geodesic polar coordinate (r, θ) on M so that we can write

$$ds^2 = dr^2 + f^2(r, \theta) d\theta^2.$$

Then we have

$$\langle \partial_r \partial_r \rangle = 1, \quad \langle \partial_\theta, \partial_\theta \rangle = f^2(r, \theta), \quad \langle \partial_r \partial_\theta \rangle = 0.$$

Therefore we have the (unique) sectional curvature

$$\langle R(\partial_r,\partial_\theta)\partial_\theta,\partial_r\rangle = c\left(\langle\partial_R,\partial_r\rangle\langle\partial_\theta,\partial_\theta\rangle - \langle\partial_r,\partial_\theta\rangle^2\right) = cf^2(r,\theta). \tag{10.1}$$

On the other hand, using the definition of R, we compute

$$\begin{aligned}\langle R(\partial_r,\partial_\theta)\partial_\theta,\partial_r\rangle &= -\langle R(\partial_r,\partial_\theta)\partial_r,\partial_\theta\rangle\\ R(\partial_r,\partial_\theta)\partial_r &= \nabla_r\nabla_\theta\partial_r - \nabla_\theta\nabla_r\partial_r = \nabla_r\nabla_r\partial_\theta\\ &= \nabla_r\nabla_r(fJ) = \frac{\partial^2 f}{\partial r^2}J.\end{aligned}$$

Therefore we obtain

$$\langle R(\partial_r,\partial_\theta)\partial_\theta,\partial_r\rangle = \left\langle\frac{\partial^2 f}{\partial r^2}\frac{1}{f}\partial_\theta,\partial_\theta\right\rangle = \frac{\partial^2 f}{\partial r^2}\frac{1}{f}\langle\partial_\theta,\partial_\theta\rangle = f\frac{\partial^2 f}{\partial r^2}. \tag{10.2}$$

Combining (10.1) and (10.2), we have obtained

$$f\frac{\partial^2 f}{\partial r^2} + cf^2 = 0 \Longrightarrow \frac{\partial^2 f}{\partial r^2} + cf = 0.$$

Combining the fact that $f^2(r,\theta) = r^2 + O(r)$, we obtain the boundary condition

$$\lim_{r\to 0} f(r,\theta) = 0, \quad \lim_{r\to} \frac{\partial f}{\partial r} = 1$$

for all $\theta = 0$. Applying the uniqueness of the solution to the initial value problem

$$\begin{cases}\frac{\partial^2 f}{\partial r^2} + cf = 0\\ f(0,\theta) = 0,\ \frac{\partial f}{\partial r}(0,\theta) = 1\end{cases}$$

we derive that f is independent of θ. Therefore we conclude

(1) (Case $c = 0$) $f(r,\theta) = r \Longrightarrow ds^2 = dr^2 + r^2 d\theta^2$,
(2) (Case $c = 1$) $f(r,\theta) = \sin r \Longrightarrow ds^2 = dr^2 + \sin^2 r d\theta^2$,
(3) ICase $c = -1$) $f(r,\theta) = \sinh r \Longrightarrow ds^2 = dr^2 + \sinh^2 r d\theta^2$.

□

11. Second variation of energy

Next we recall the general second variation formula.

Theorem 11.1 (Second variation formula)**.** *Let $\gamma : [0,a] \to (N,g)$ be a geodesic on a Riemannian manifold. Let $C : (-\epsilon,\epsilon)\times[0,a] \to N$ be a variation of γ i.e, a map satisfying $C(0,t) = \gamma(t)$ Denote $V(s,t) = \frac{\partial C}{\partial s}(s,t)$ and $c_s := C(s,\cdot)$. The second variational formula at a geodesic γ is given by*

$$\begin{aligned}\frac{d^2E(c_s)}{ds^2}\Big|_{s=0} &= \int_0^1\left\langle\frac{DV}{\partial t},\frac{DV}{\partial t}\right\rangle - \int_0^1\langle R(V,\dot\gamma)\dot\gamma,V\rangle\,dt\\ &\quad -\left\langle\frac{DV}{\partial s}(0),\dot\gamma(0)\right\rangle + \left\langle\frac{DV}{\partial s}(a),\dot\gamma(a)\right\rangle. \qquad (11.1)\\ &= -\int_0^1\left\langle\frac{D^2V}{\partial t^2} + R(V,\dot\gamma)\dot\gamma,V\right\rangle dt\\ &\quad -\left\langle\frac{DV}{\partial s}(0),\dot\gamma(0)\right\rangle + \left\langle\frac{DV}{\partial s}(a),\dot\gamma(a)\right\rangle\\ &\quad -\left\langle\frac{DV}{\partial t}(0),V(0)\right\rangle + \left\langle\frac{DV}{\partial t}(a),V(a)\right\rangle \qquad (11.2)\end{aligned}$$

Proof. We start with the following lemma

Lemma 11.2. *Let ξ be a vector field over a map $f:(-\epsilon,\epsilon)\times[0,a]\to M$. Then*

$$\frac{D}{ds}\frac{D\xi}{dt}-\frac{D}{\partial t}\frac{D\xi}{\partial s}=R(\frac{\partial f}{\partial s},\frac{\partial f}{\partial t})\xi.$$

Applying this to C and $\xi=\frac{\partial C}{\partial s}$, we compute the derivatives of $E(c_s)=\frac{1}{2}\int_0^a|\dot{c}_s|^2$,

$$\frac{dE(c_s)}{ds}=\int_0^a\left\langle\frac{D}{\partial s}\frac{\partial C}{\partial t},\frac{\partial C}{\partial t}\right\rangle dt$$

and

$$\begin{aligned}\frac{d^2E(c_s)}{ds^2} &= \int_0^a\left\langle\frac{D}{\partial s}\frac{\partial C}{\partial t},\frac{D}{\partial s}\frac{\partial C}{\partial t}\right\rangle dt+\int_0^a\left\langle\frac{D}{\partial s}\frac{D}{\partial s}\frac{\partial C}{\partial t},\partial C\partial t\right\rangle dt\\ &= \int_0^a\left\langle\frac{D}{\partial t}\frac{\partial C}{\partial s},\frac{D}{\partial t}\frac{\partial C}{\partial s}\right\rangle dt+\int_0^a\left\langle\frac{D}{\partial s}\frac{D}{\partial t}\frac{\partial C}{\partial s},\frac{\partial C}{\partial t}\right\rangle dt.\end{aligned}$$

The first term becomes

$$\begin{aligned}\int_0^a\left\langle\frac{D}{\partial s}\frac{\partial C}{\partial t},\frac{D}{\partial s}\frac{\partial C}{\partial t}\right\rangle dt &= \int_0^a\left\langle\frac{D}{\partial t}\frac{\partial C}{\partial s},\frac{D}{\partial t}\frac{\partial C}{\partial s}\right\rangle dt\\ &= \int_0^1\left\langle\frac{DV}{\partial t},\frac{DV}{\partial t}\right\rangle dt.\end{aligned}$$

For the second term, using the lemma, we rewrite

$$\int_0^a\left\langle\frac{D}{\partial s}\frac{D}{\partial t}\frac{\partial C}{\partial s},\frac{\partial C}{\partial t}\right\rangle dt=\int_0^a\left\langle\frac{D}{\partial t}\frac{D}{\partial s}\frac{\partial C}{\partial s},\frac{\partial C}{\partial t}\right\rangle dt+\int_0^a\left\langle R\left(\frac{\partial C}{\partial s},\frac{\partial C}{\partial t}\right)\frac{\partial C}{\partial s},\frac{\partial C}{\partial t}\right\rangle dt.$$

We do integration by pars,

$$\begin{aligned}\int_0^a\left\langle\frac{D}{\partial t}\frac{D}{\partial s}\frac{\partial C}{\partial s},\frac{\partial C}{\partial t}\right\rangle dt &= \int_0^a\frac{\partial}{\partial t}\langle\frac{D}{\partial s}\frac{\partial C}{\partial s},\frac{\partial C}{\partial t}\rangle-\langle\frac{\partial C}{\partial s},\frac{D}{\partial t}\frac{\partial C}{\partial t}\rangle\,dt\\ &= \langle\frac{\partial C}{\partial t},\frac{D}{\partial s}\frac{\partial C}{\partial s}\rangle\Big|_0^a.\end{aligned}$$

By evaluating $s=0$, we derive (11.1).

For (11.2), we further apply the integration by parts to

$$\begin{aligned}\int_0^1\left\langle\frac{DV}{\partial t},\frac{DV}{\partial t}\right\rangle dt &= \int_0^1\frac{d}{dt}\left\langle\frac{DV}{\partial t},V\right\rangle dt-\int_0^1\left\langle\frac{D^2V}{\partial t^2},V\right\rangle dt\\ &= -\int_0^1\left\langle\frac{D^2V}{\partial t^2},V\right\rangle dt+\left\langle\frac{DV}{\partial t},V\right\rangle\Big|_0^a.\end{aligned}$$

□

Remark 11.3. We would like to emphasize that the general second variation formula *allowing the end points to move* contains the boundary terms appearing in the last line of (11.1) which is the same as

$$-\left\langle\frac{D}{\partial s}\frac{\partial C}{\partial s},\dot{\gamma}\right\rangle(0,0)+\left\langle\frac{D}{\partial s}\frac{\partial C}{\partial s},\dot{\gamma}\right\rangle(0,1).$$

(These terms will not appear in (11.1) *when the variation is fixed at the end* $t=0$, 1. See e.g., [Sp2, p.303] or [dC, Remark 2,10] for such a discussion.)

Under the fixed boundary condition i.e., with $c_s(0) = p$, $c_s(a) = q$, or the periodic boundary condition, i.e., with $c_s(0) = c_s(a)$ and $\dot c_s(0) = \dot c_s(a)$, the boundary terms drop out and the formula (11.1) and (11.2) are reduced to

$$\int_0^a \left\langle \frac{DV}{\partial t}, \frac{DV}{\partial t} \right\rangle - \langle R(V,\dot\gamma)\dot\gamma, V\rangle \, dt$$

and

$$-\int_0^a \left\langle \frac{D^2V}{\partial t^2} + R(V,\dot\gamma)\dot\gamma, V \right\rangle dt$$

respectively.

Definition 11.4 (Jacobi field)**.** A variation V over a geodesic γ is called a *Jacobi field* if V satisfies the equation

$$\frac{D^2V}{dt^2} + R(V,\dot\gamma)\dot\gamma = 0.$$

Definition 11.5 (Index form)**.** For two infinitesimal variations of a geodesic γ, we define the *index form* I_a by

$$I_a(V,W) = \int_0^a \langle \frac{DV}{\partial t}, \frac{DV}{\partial t}\rangle - \langle R(V,\dot\gamma)\dot\gamma, W\rangle \, dt.$$

This form defines a symmetric quadratic form on the tangent space of the set of paths with the following boundary conditions:

- (Two-point boundary condition)

$$\mathcal{P}_{q,p}(a) = \{c : [0,a] \to M \mid c(0) = p,\, c(a) = q\}$$

- (Periodic boundary condition)

$$\Omega(a) = \{c : \mathbb{R}\mathbb{Z} \to M \mid c \text{ is } a\text{-periodic } \}$$

Equip $\mathcal{V}$ with the L^2-inner product

$$\langle\langle V, W\rangle\rangle = \int_0^1 \langle V(t), \langle W(t)\rangle \, dt$$

We decompose $\mathcal{V} = \mathcal{V}^+ \oplus \mathcal{V}^0 \oplus \mathcal{V}^-$ into the positive, zero and negative eigenspaces of the symmetric (**Why is it symmetric?**) quadratic form I_a on $\mathcal{V}$.

Proposition 11.6. *Consider the index form*

$$I_a(V,W) = \int_0^1 \left\langle \frac{DV}{\partial t}(t), \frac{DW}{\partial t}(t) \right\rangle - \langle R(V(t),\dot\gamma(t))\dot\gamma(t), W(0,t)\rangle \, dt$$

on the set $\mathcal{V}$ of vector fields along γ with either of the following boundary conditions:

$$V(0) = V(a) = 0, \quad V(0) = V(a)$$

Then the index and the nullity of I_a are finite.

Proof. We will use the following well-known fact in the functional analysis

Lemma 11.7. *A Hilbert space $\mathbb{H}$ (i.e., a complete inner product space) whose closed unit ball is compact if and only if* $\dim \mathbb{H}$ *is finite.*

To apply this lemma, we will take the L^2-completion of $\mathcal{V}^0 \oplus \mathcal{V}^-$ and consider the closed L^2 unit-ball thereof. Denote the closed L^2 unit-ball by

$$B^1_{L^2}(\mathcal{V}) = \{V \in \mathcal{V} \mid \|V\|_{L^2} \leq 1\}.$$

We will show that

$$B := \mathcal{V}^0 \oplus \mathcal{V}^- \cap B^1_{L^2}(\mathcal{V})$$

is pre-compact in L^2. This in particular implies that the L^2-closure

$$\overline{B} = \overline{\mathcal{V}^0 \oplus \mathcal{V}^-} \cap B^1_{L^2}(\mathcal{V})$$

is compact which implies $\overline{\mathcal{V}^0 \oplus \mathcal{V}^-}$ is finite-dimensional by the above lemma. This in turn implies that $\mathcal{V}^0 \oplus \mathcal{V}^-$ itself is finite dimensional which we wanted to prove.

It remains to prove that B is pre-compact in L^2 on $[0,a]$. Actually we will prove that B is pre-compact in C^0-topology on $[0,a]$. We first observe that if $V \in \mathcal{V}^0 \oplus \mathcal{V}^-$ with $\|V\|_{L^2} \leq 1$,

$$0 \geq I_a(V,V) = \int_0^1 \left\langle \frac{DV}{dt}(t), \frac{DV}{dt}(t) \right\rangle - \langle R(V(t),\dot\gamma(t))\dot\gamma(t), V(0,t)\rangle \, dt$$

and hence we have

$$\int_0^1 \left\langle \frac{DV}{dt}(t), \frac{DV}{dt}(t) \right\rangle dt \leq \langle R(V(t),\dot\gamma(t))\dot\gamma(t), V(0,t)\rangle \, dt$$

and so

$$\left\| \frac{DV}{dt} \right\|^2_{L^2} \leq \|R(\gamma)\|_{C^0} \|\dot\gamma\|^2_{C^0} \|V\|^2_{L^2}$$

where $\|R(\gamma)\|_{C^0}$ is the C^0-norm of the function $t \mapsto R(\gamma(t))$. In particular, we have

$$\left\| \frac{DV}{dt} \right\|_{L^2} < C < \infty$$

for all $\|V\|_{L^2} \leq 1$ with

$$C = \sqrt{\|R(\gamma)\|_{C^0}} \|\dot\gamma\|_{C^0} < \infty$$

which is independent of V but depends only on g and γ. Let $t_0 \in [0,a]$ be any given point. Then

$$(\Pi^t_{t_0})^{-1}(V(t)) = V(t_0) + \int_{t_0}^t \left((\Pi^s_{t_0})^{-1}(\frac{DV}{dt}(s) \right) ds$$

Lemma 11.8. *The section $V \in \Gamma(\gamma^*TM)$ is equi-continuous in the sense that the function $t \mapsto (\Pi^t_{t_0})^{-1}(V(t)) \in T_{\gamma(t_0)}M$ is equicontinuous on $[0,a]$ uniformly over B.*

Proof. We compute

$$\begin{aligned} |(\Pi^t_{t_0})^{-1}(V(t)) - V(t_0)| &\leq \int_{t_0}^t \left| (\Pi^s_{t_0})^{-1}(\frac{DV}{dt}(s)) \right| ds \\ &\leq \sqrt{|t-t_0|} \sqrt{\int_{t_0}^t \left| \frac{DV}{dt}(t) \right|^2 dt} \\ &\leq \sqrt{|t-t_0|} \left\| \frac{DV}{dt} \right\|_{L^2} < C\sqrt{|t-t_0|}. \end{aligned}$$

Here we use Hölder's inequality for the second inequality and the property that the parallel transport map $\Pi_{t_0}^t$ preserves the norm. Noting that C does not depend on V, t_0 and t, we have finished the proof. □

Since $[0, a]$ is compact, Ascoli-Arzela theorem implies that B is pre-compact in C^0-topology. This finishes the proof. □

Remark 11.9. (1) In the above proof, we in fact have

$$\overline{\mathcal{V}^0 \oplus \mathcal{V}^-} \cap B^1_{L^2}(\mathcal{V}) = (\mathcal{V}^0 \oplus \mathcal{V}^-) \cap B^1_{L^2}(\mathcal{V})$$

since the latter is a dense subspace of the former in L^2 which is finite dimensional.

(2) The argument used in the above proof is the essence of the proofs of the Sobolev embedding theorem and the Reillich compactness theorem

$$C^0 \hookrightarrow W^{1,p}, \quad 1 - \frac{n}{p} > 0$$

in functional analysis. The current case corresponds to $p = 2$ and $n = 1$.

We have the following comparison theorem which is the beginning of global Riemannian geometry.

Theorem 11.10 (Bonnet-Myers)**.** *Let M be a complete Riemannian manifold. Suppose that the Ricci-curvature of M positively-pinched, i.e., satisfies*

$$\mathrm{Ric}_p(v, v) \geq \frac{1}{r^2} > 0$$

for all $p \in M$ and for all $v \in T_pM$. Then M is compact and has the diameter $\leq \pi r$.

Proof. Let p and q be any pair of points in M. Since M is complete, there exists a length-minimizing geodesic γ between them. It is enough to prove that $L(\gamma) =: \ell \leq \pi r$.

Suppose to the contrary that $L(\gamma) > \pi r$. Choose parallel fields $E_1(t), \ldots, E_{n-1}(t)$ which are orthonormal and perpendicular to $\gamma'(t)$. Write $E_n(t) = \frac{\gamma'(t)}{\ell}$. Then $\{E_1(t), \ldots, E_n(t)\}$ form an orthonormal parallel frame of TM along γ. Consider the vector fields along γ defined by

$$V_j(t) = (\sin \pi t) E_j(t), \quad j = 1, \ldots, n-1.$$

Surely we have $V_j(0) = V_j(1) = 0$. Therefore the second variation formula gives rise to

$$E_j''(0) = -\int_0^1 \langle \frac{D^2 V_i}{dt^2} + R(V_j, \dot\gamma), V_j \rangle \, dt.$$

But we evaluate

$$\langle \frac{D^2 V_i}{dt^2} + R(V_j, \dot\gamma), V_j \rangle = -\pi^2 \sin^2 \pi t + \ell^2 K_j(t) \sin^2 \pi t$$

where $K_j(t) := \langle R(E_j(t), E_n(T))E_n(t), E_j(t) \rangle$, and hence

$$E_j''(0) = \int_0^1 \ell^2 \sin^2 \pi t(-\pi^2 + K_j(t)) \, dt.$$

By summing up over $j = 1, \ldots, n-1$, we have derived

$$\begin{aligned}\sum_{j=1}^{n-1} E_j''(0) &= \int_0^1 (n-1)\ell^2\pi^2 \sin^2 \pi t - \ell^2 \sin^2 \pi t \sum_{j=1}^{n-1} K_j(t)\, dt \\ &= \int_0^1 \left(\ell^2(n-1)\pi^2 \sin^2 \pi t - (n-1)\ell^2 \operatorname{Ric}_{\gamma(t)}(E_n(t), E_n(t))\right)\, dt.\end{aligned}$$

Since $\operatorname{Ric}_{\gamma(t)}(E_n(t)) \geq \frac{1}{r^2}$ and $\ell > \pi r$, we have

$$(n-1)\ell^2 \operatorname{Ric}_{\gamma(t)}(E_n(t), E_n(t)) > (n-1)\pi^2.$$

This implies $E''(0) < 0$ and hence there is some $j = 1$ such that $E_j''(0) < 0$. This contradicts that γ is a minimizing geodesic between p and q. □

Since the curvature assumption is local, the same applies to the universal covering space of M.

Corollary 11.11. *Under the same hypothesis, the universal covering space M is compact. In particular M has finite fundamental group.*

Example 11.12.
(1) T^n cannot be given a metric of positive Ricci-curvature.
(2) If M has sectional curvatures $\geq \frac{1}{r^2} > 0$, then M is compact and $\operatorname{diam}(M) \leq \pi r$.
(3) We cannot weaken the hypothesis $K \geq \frac{1}{r^2} > 0$ to just $K > 0$, since the paraboloid
$$\{(x, y, z) \in \mathbb{R}^3 \mid z = x^2 + y^2\}$$
has positive curvature $K > 0$ which is complete and non-compact.

Part 2. Symplectic Geometry

12. Geometry of cotangent bundles

Let N be any smooth manifold and T^*N be its cotangent bundle. When $N = \mathbb{R}^n$,

$$T^*N \cong \mathbb{R}^n \times (\mathbb{R}^n)^* \cong \mathbb{R}^{2n} \cong \mathbb{C}^n.$$

The first isomorphism arises by the canonical coordinates of an element $\alpha \in T^*N$

$$\alpha \mapsto (q(\alpha), p(\alpha)), \quad q = (q_1, \ldots, q_n),\ p = (p_1, \ldots, p_n),$$

and the second follows by

$$(q_i, p_i) \mapsto q_i + \sqrt{-1}p_i.$$

In classical mechanics, $\mathbb{R}^{2n} \cong T^*\mathbb{R}^n$ equipped with *standard canonical coordinates* $(q_1, q_2, \ldots, q_n, p_1, \ldots, p_n)$.

Definition 12.1 (Lagrange-Poisson bracket)**.** For given pair F, G of functions on $\mathbb{R}^{2n}$, the Lagrange-Poisson bracket is defined to be

$$\{F, G\}_{LP} = \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial q_i}\frac{\partial F}{\partial p_i} \right) \tag{12.1}$$

Regard the bracket operation $(f, g) \mapsto \{f, g\}_{LP}$ as a bilinear map

$$C^\infty(\mathbb{R}^{2n}) \times C^\infty(\mathbb{R}^{2n}) \to C^\infty(\mathbb{R}^{2n}).$$

Proposition 12.2. *The bracket $\{\cdot, \cdot\} = \{\cdot, \cdot\}_{LP}$ satisfies the following properties:*

(1) *(Skew symmetry)* $\{F,G\} = -\{G,F\}$,
(2) *(Bilinearity)* $\{F+G,H\} = \{F,H\} + \{G,H\}$,
(3) *(Leibniz rule)* $\{FG,H\} = F\{G,H\} + \{F,H\}G$,
(4) *(Jacobi identity)* $\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\} = 0$,
(5) *(Nondegeneracy)* $\{F,G\} = 0$ *for all* G *if and only if* F *is a constant function.*

Definition 12.3. We call a (local) coordinate $(y^1,\cdots,y^{2n})$ is called a *Hamiltonian canonical coordinate* if it can be permuted to $(Q_1,P_1,Q_2,P_2,\cdots,Q_n,P_n)$ such that

$$\{Q_i,Q_j\}_{LP} = 0 = \{P_i,P_j\}_{LP}, \quad \{Q_i,P_j\}_{LP} = \delta_{ij}.$$

Exercise 12.4. Let $(Q_1,P_1,Q_2,P_2,\cdots,Q_n,P_n)$ be any coordinate system on $\mathbb{R}^{2n}$. Prove that the coordinate system is Hamiltonian-canonical if and only if for any function F, $G:\mathbb{R}^{2n}\to\mathbb{R}$

$$\sum_{i=1}^n\left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial q_i}\frac{\partial F}{\partial p_i}\right) = \sum_{i=1}^n\left(\frac{\partial F}{\partial Q_i}\frac{\partial G}{\partial P_i} - \frac{\partial G}{\partial Q_i}\frac{\partial F}{\partial P_i}\right).$$

In particular, under the identification of $\mathbb{R}^{2n}\cong T^*\mathbb{R}^n$, we have

Proposition 12.5. *Under the identification of* $\mathbb{R}^{2n}\cong\mathbb{R}^n\times(\mathbb{R}^n)^* = T^*\mathbb{R}^n$, *any canonical coordinates associated to a coordinate* $(x^1,\cdots,x^n)$ *on* $\mathbb{R}^n$ *is Hamiltonian canonical.*

Proof. We denote by $(q_1,\ldots,q_n,p_1,\cdots,p_n)$ the canonical coordinates associated to the standard coordinates of $\mathbb{R}^n$.

Recall a general canonical coordinates is associated to a coordinates $(x^1,\cdots,x^n)$ of the base $\mathbb{R}^n$ is defined by $(Q_1,\ldots,Q_n,P_1,\ldots,P_n)$ is defined to be

$$Q_i(\alpha) = x_i\circ\pi(\alpha),\ P_j = \alpha\left(\frac{\partial}{\partial x^j}\right).$$

We compute $\{Q_i,Q_j\}$, $\{Q_i,P_j\}$ and $\{P_i,P_j\}$ separately. By definition, Q_i does not depend on p_i's but depends only on q_i's. Therefore we have $\{Q_i,Q_j\}=0$.

By definition of canonical coordinates, we have

$$\sum_i P_i dQ_i = \sum_j p_j dq_j. \tag{12.2}$$

(**Why does this follow from the definition?**) The RHS can be written as

$$\sum_i\sum_j p_j\frac{\partial q_j}{\partial Q_i}dQ_i = \sum_i\left(\sum_j p_j\frac{\partial q_j}{\partial Q_i}\right)dQ_i.$$

This proves

$$P_i = \sum_j p_j\frac{\partial q_j}{\partial Q_i}. \tag{12.3}$$

We compute

$$\frac{\partial P_i}{\partial p_j} = \frac{\partial q_j}{\partial Q_i}. \tag{12.4}$$

Therefore

$$\begin{aligned}\{P_k, P_\ell\}_{LP} &= \sum_{i=1}^n \left(\frac{\partial P_k}{\partial q_i}\frac{\partial P_\ell}{\partial p_i} - \frac{\partial P_\ell}{\partial q_i}\frac{\partial P_k}{\partial p_i} \right) \\ &= \sum_{i=1}^n \left(\frac{\partial P_k}{\partial q_i}\frac{\partial q_i}{\partial Q_\ell} - \frac{\partial P_\ell}{\partial q_i}\frac{\partial q_i}{\partial Q_k} \right) \\ &= \frac{\partial P_k}{\partial Q_\ell} - \frac{\partial P_\ell}{\partial Q_k} = 0 - 0 = 0 \qquad (12.5)\end{aligned}$$

where the penultimate equality follows since $\frac{\partial p_k}{\partial q_j} = 0$ for all j and so $\frac{\partial p_k}{\partial Q_j} = 0$ as $\frac{\partial}{\partial Q_j}$ is a linear combination of $\frac{\partial}{\partial q_i}$'s not involving $\frac{\partial}{\partial p_\ell}$'. The last equality from the definition of canonical coordinates $(Q_1, \dots, Q_n, P_1, \dots, P_n)$.

Finally we compute

$$\{Q_k, P_\ell\}_{LP} = \sum_{i=1}^n \left(\frac{\partial Q_k}{\partial q_i}\frac{\partial P_\ell}{\partial p_i} - \frac{\partial P_\ell}{\partial q_i}\frac{\partial Q_k}{\partial p_i} \right). \qquad (12.6)$$

Substituting (12.4) into (12.6), we obtain

$$\{Q_k, P_\ell\}_{LP} = \sum_{i=1}^n \left(\frac{\partial Q_k}{\partial q_i}\frac{\partial q_i}{\partial Q_\ell} - \frac{\partial q_i}{\partial Q_\ell}\frac{\partial Q_k}{\partial p_i} \right) = \sum_{i=1}^n \frac{\partial Q_k}{\partial q_i}\frac{\partial q_i}{\partial Q_\ell} = \frac{\partial Q_k}{\partial Q_\ell} = \delta_{k\ell}.$$

Here again the first equality follows by the same reason and the second equality from the vanishing $\frac{\partial Q_k}{\partial p_i} = 0$ since $Q_k = Q_k(q_1, \cdots, q_n)$. This finishes the proof. □

This indicates existence of some geometric structure associated to the bracket operation on the cotangent bundle T^*N any smooth manifold N, in particular on the classical phase space $\mathbb{R}^{2n}$.

Indeed (12.2) shows that there is a globally defined one-form on T^*N whose coordinate expression given by $\sum_{i=1} p_i dq_i$ and a two-form given by $\sum_i dp_i \wedge dq_i$.

Definition 12.6 (Canonical symplectic form)**.** The *Liouville one-form* on T^*N is defined by the formula

$$\theta_\alpha(\xi) = \alpha(d\pi(\xi))$$

for $\xi \in T_\alpha(T^*N)$. The canonical symplectic form on T^*N is defined by

$$\omega_0 := -d\theta.$$

One can check that in coordinates we have $\theta = \sum_i p_i dq_i$ and $\omega_0 = \sum_i dq_i \wedge dp_i$. By definition, ω_0 is a closed (in fact exact) two-form and *nondegenerate*.

Definition 12.7. A two form ω on a manifold M is called nondegenerate if the bundle map

$$\widetilde{\omega} : v \mapsto v\rfloor\omega; \quad T_xM \to T_x^*M$$

is an isomorphism.

Since ω is a skew-symmetric bilinear form on each tangent space, M must be even dimensional.

13. Poisson manifolds and Schouten-Nijenhuis bracket

We start with the notion of Lie algebra.

Definition 13.1 (Lie algebra)**.** A Lie algebra is a vector space $\mathfrak{g}$ equipped with a bilinear map

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

that satisfies the following properties:

(1) (skew-symmetry) $[a, b] = -[b, a]$,
(2) (Jacobi identity)

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0.$$

When $\dim \mathfrak{g} < \infty$, we can provide the coordinate description as follows: Let $\{e_1, \cdots, e_n\}$ be a basis. Then we have

$$[e_i, e_j](\mu) = C_{ijk} e_k$$

for some constants $\{C_{ijk}\}$ called the *structure constants*. Then they satisfy the structure equations

$$C_{ijk} = -C_{jik},$$
$$\sum_{h=1}^{n} (C_{ijh} C_{hk\ell} + C_{jkh} C_{hi\ell} + C_{kih} C_{hj\ell}) = 0.$$

When the elements a, b carries grading and the above two properties hold in the graded sense, we call such a Lie algebra a *graded Lie algebra*.

Definition 13.2 (Poisson algebra)**.** A *Poisson algebra* is an algebra A that also carries a Lie bracket, denoted by $\{\cdot,\cdot\} : A \times A \to A$, that is compatible with the algebra structure, i.e., that satisfies the properties (1) - (4) stated in Proposition 12.2. Such a bracket is called a *Poisson bracket*.

When the associated Lie algebra and algebra are graded, the resulting graded Poisson algebra is also called a *super Poisson algebra*.

Definition 13.3. A smooth manifold P for which $A := C^\infty(M)$ carries a Poisson bracket $\{\cdot,\cdot\}$ compatible with the usual product operation of $C^\infty(P)$ is called a *Poisson manifold*. When the Poisson bracket is nondegenerate, we call it a *symplectic manifold*.

Definition 13.4 (Poisson maps)**.** A Poisson map is a C^∞ map $\varphi : (P_1, \{\cdot,\cdot\}_1) \to (P_2, \{\cdot,\cdot\}_2)$ such that the map $\varphi^* : C^\infty(P_2) \to C^\infty(P_1)$ satisfies

$$\{\varphi^* f, \varphi^* g\}_1 = \varphi^*\{f, g\}_2$$

for all f, $g \in C^\infty(P_2)$.

By definition, the map $f \mapsto \{\cdot, f\}$ is a derivation on $C^\infty(P)$ and so defines a vector field on P.

Definition 13.5. We call the vector field associated to the derivation $\{\cdot, f\}$ the *Hamiltonian vector field* of f which is denoted by X_f.

13.1. **Poisson tensor and Jacobi identity.** A bi-vector field is a section of $TP \otimes TP$. Given a Poisson bracket $\{\cdot, \cdot\}$, we can define a bi-vector Π so that the equality

$$\{f, g\} = \Pi(df, dg) \quad f,\, g \in C^\infty(P)$$

holds. More precisely, we have

Definition 13.6. Define a bilinear map $\Pi_x : T_x^*P \times T_x^*P \to \mathbb{R}$ as follows: For each given $\alpha_1,\, \alpha_2 \in T_x^*P$, we define

$$\Pi_x(\alpha_1, \alpha_2) := \{f, g\}(x)$$

for $f,\, g \in C^\infty(P)$ such that $df(x) = \alpha_1,\, dg(x) = \alpha_2$.

Because $\{C, \cdot\} = 0$ for any constant function, this definition indeed defines a skew-symmetric bi-vector field Π. What would correspond to the Jacobi identity for the bi-vector field Π, called the Poisson tensor associated to the Poisson bracket?

To understand this we need to consider the whole of (dual) exterior algebra, $\Gamma(\wedge^*(TP))$, the set of multi-vector fields, and an extension of the Lie algebra of vector fields to the *graded Lie algebra* of multi-vector fields.

Definition 13.7 (Schouten-Nijenhuis bracket)**.** is a (graded) bilinear map

$$[\cdot, \cdot] : \Gamma(\wedge^k(TP)) \otimes \Gamma(\wedge^\ell(TP)) \to \Gamma(\wedge^{k+\ell-1}(TP))$$

for $k,\, \ell \geq 0$ that satisfies the following properties:

(1) If $X \in \Gamma(TP) = \Gamma(TP)$,[X,Y]= $\mathcal{L}_X Y$,
(2) (Graded skew-symmetric) $[X^{(k)}, Y^{(\ell)}] = -(-1)^{(k-1)(\ell-1)}[Y^{(\ell)}, X^{(k)}]$,
(3) (Graded derivation)

$$[X^{(k)}, Y^{(\ell)} \wedge Z^{(m)}] = [X^{(k)}, Y^{(\ell)}] \wedge Z^{(m)} + (-1)^{(k-1)\ell} Y^{(\ell)} \wedge [X^{(k)}, Z^{(m)}].$$

(4) (Graded Jacobi identity) For $A = X^{(k)}$, $B = Y^{(\ell)}$, $C = Z^{(m)}$,

$$(-1)^{(k-1)(m-1)}[A, [B, C]] + (-1)^{(\ell-1)(k-1)}[B, [C, A]] + (-1)^{(m-1)(\ell-1)}[C, [A, B]] = 0$$

Hence the multi-bracket puts a super-Poisson algebra structure on $\Gamma(\wedge^*(TP))$ with respect to the wedge product and grading $|X^{(k)}|' = k - 1$.

Exercise 13.8. Define $\mathrm{ad}_A = [A, \cdot]$ and $[\mathrm{ad}_A, \mathrm{ad}_B] = \mathrm{ad}_A \mathrm{ad}_B - (-1)^{(a-1)(b-1)} \mathrm{ad}_B \mathrm{ad}_A$.

(1) Check that the graded Jacobi identity is equivalent to $[\mathrm{ad}_A, \mathrm{ad}_B] = \mathrm{ad}_{[A,B]}$.
(2) We know $\mathrm{ad}_A(B \wedge C) = \mathrm{ad}_A B \wedge C + (-1)^{(a-1)b} B \wedge \mathrm{ad}_A C$. What is the formula for $\mathrm{ad}_A[B, C]$?

Proposition 13.9. *Let Π be the Poisson tensor associated to the Poisson structure on P.*

(1) *For a function f on P, let $X_f = \{\cdot, f\}$ be the Hamiltonian vector field associated to f. Then $[\Pi, f] = X_f$.*
(2) *The Jacobi identity on $C^\infty(P)$ is equivalent to $[\Pi, \Pi] = 0$.*

Proof. Let $(x_1, \cdots, x_n)$ be a coordinates and express

$$\Pi = \frac{1}{2} \pi_{ij} \frac{\partial}{\partial x_i} \wedge \frac{\partial}{\partial x_j}.$$

We first express the Hamiltonian vector field $X_f = \{\cdot, f\}$ in coordinates. For this we compute $\{x_i, f\}$ for the function f. By definition of Π, we compute

$$\begin{aligned}
\{x_i, f\} &= \Pi(dx_i, df) = \Pi(dx_i, \frac{\partial f}{\partial x_j} dx_j) \\
&= \frac{\partial f}{\partial x_j} \frac{1}{2} \pi_{k\ell} \left(\frac{\partial}{\partial x_k} \wedge \frac{\partial}{\partial x_\ell} \right) (dx_i, dx_j) \\
&= \frac{\partial f}{\partial x_j} \frac{1}{2} \pi_{k\ell} \begin{vmatrix} dx_i(\frac{\partial}{\partial x_k}) & dx_j(\frac{\partial}{\partial x_k}) \\ dx_i(\frac{\partial}{\partial x_\ell}) & dx_j(\frac{\partial}{\partial x_\ell}) \end{vmatrix} \\
&= \frac{\partial f}{\partial x_j} \frac{1}{2} \pi_{k\ell} (\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}) = \frac{\partial f}{\partial x_j} \pi_{ij}
\end{aligned}$$

and hence we have $X_f = \pi_{ij} \frac{\partial f}{\partial x_j} \frac{\partial}{\partial x_i}$.

We then compute

$$\begin{aligned}
2[\Pi, f] &= [\pi_{ij} \frac{\partial}{\partial x_i} \wedge \frac{\partial}{\partial x_j}, f] = \pi_{ij} \frac{\partial}{\partial x_j} \wedge [\frac{\partial}{\partial x_i}, f] - \pi_{ij} \frac{\partial}{\partial x_i} \wedge [\frac{\partial}{\partial x_j}, f] \\
&= \pi_{ij} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} - \pi_{ji} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} = 2\pi_{ij} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} = 2X_f.
\end{aligned}$$

This proves (1).

Then a direct calculation shows

$$[\Pi, \Pi] = -\frac{1}{2} \left(\frac{\partial \pi_{ik}}{\partial x_j} \pi_{j\ell} + \text{`cyclic sum over } ik\ell\text{'} \right) \frac{\partial}{\partial x^i} \wedge \frac{\partial}{\partial x^k} \wedge \frac{\partial}{\partial x_\ell}.$$

On the other hand, we have

$$\pi_{ij} = \{x_i, x_j\}$$

and so

$$\{x_i, \{x_j, x_k\}\} = \{x_i, \pi_{jk}\} = -X_{x_i}(\pi_{jk}) = -[\Pi, x_i](\pi_{jk}) = -\pi_{\ell i} \frac{\partial \pi_{jk}}{\partial x_\ell}.$$

By comparing the two, we have proved $[\Pi, \Pi] = 0$ if and only if

$$\{x_i, \{x_j, x_k\}\} + \text{`cyclic sum over } ijk\text{'} = 0$$

for all i, j, k. This finishes the proof. □

13.2. **Lie-Poisson space.** Let $\mathfrak{g}$ be a (finite dimensional) Lie algebra and $\mathfrak{g}^*$ its dual space. Regarding $P = \mathfrak{g}^*$ as a manifold, we will put a Poisson structure as follows:

First we identify $(\mathfrak{g}^*)^* \cong \mathfrak{g}$ and

$$df(\mu) \in T^*_\mu \mathfrak{g}^* \cong (\mathfrak{g}^*)^* \cong \mathfrak{g}$$

for $\mu \in \mathfrak{g}^*$ and $f \in C^\infty(\mathfrak{g}^*)$. By this identification we can define a bracket on $C^\infty(\mathfrak{g}^*)$ by setting

$$\{f, g\}(\mu) = \langle \mu, [df(\mu), dg(\mu)] \rangle.$$

We can provide the coordinate description as follows: Let $\{e_1, \cdots, e_n\}$ be a basis of $\mathfrak{g} \subset C^\infty(\mathfrak{g}^*)$ and $x_1, \cdots, x_n$ the corresponding linear functions $\mathfrak{g}^*$. Then

$$\{x_i, x_j\}(\mu) = \langle \mu, [e_i, e_j] \rangle = \langle \mu, C_{ijk} x_k \rangle = C_{ijk} x_k(\mu)$$

where C_{ijk} are the structure constants of the Lie algebra $\mathfrak{g}$ for the basis $\{e_1, \cdots, e_n\}$. Therefore the corresponding Poisson tensor has the coordinate expression

$$\pi_{ij} = \{x_i, x_j\} = C_{ijk} x_k.$$

The structure equation of the Lie algebra implies that the Lie-Poisson bracket satisfies the Jacobi identity. The derivation rule holds from the fact that exterior differential satisfies derivation rule.

Definition 13.10. Let $\mathfrak{g}$ be any Lie algebra. The Lie-Poisson space is the pair $(\mathfrak{g}^*, \{\cdot, \cdot\})$.

In this way, for any Lie algebra $\mathfrak{g}$, its dual space $\mathfrak{g}^*$ carries a canonical Poisson structure which we call the *Lie Poisson structure*. This Poisson structure is special in that it admits a coordinate with respect to which the functions π_{ij} are all linear functions. For a general Poisson structure, such a coordinate does not exist in general.

14. Symplectic forms and the Jacobi identity

From now on, we assume that the bracket is nondegenerate but it may not satisfy the Jacobi identity.

One might wonder what if we drop the condition ω being closed as the skew-symmetric analog to the Riemannian metric which is symmetric positive-definite (and so nondegenerate) bilinear form.

Definition 14.1. We call a nondegenerate two-form ω a *quasi-symplectic form* on M, which is not necessarily closed. Consider $h \in C^\infty(M)$.

(1) The *quasi-Hamiltonian vector field*, associated to h, denoted by X_h, is the vector field defined by

$$X_h = \widetilde{\omega}^{-1}(dh). \tag{14.1}$$

(2) The *quasi-Poisson bracket*, denoted by $\{f, h\}$, is defined by

$$\{f, h\} = \omega(X_f, X_h). \tag{14.2}$$

We first mention that we can equivalently write

$$\{f, h\} = df(X_h) = X_h[f]. \tag{14.3}$$

It immediately follows from the definition that the quasi-Poisson bracket associated to any nondegenerate two-form satisfies skew-symmetry, bilinearity and the Leibnitz rule. The remaining question is on what condition of ω the associated quasi-Poisson bracket satisfies the Jacobi identity. The following is a consequence of nondegeneracy.

Exercise 14.2. Prove that the set of quasi-Hamiltonian vector fields is ample in that the following holds: Let $x \in M$ be any given point. Then we have

$$\{X(x) \mid X = X_h,\, h \in C^\infty(M)\} = T_x M.$$

Here comes a fundamental relationship between the closedness of the nondegenerate two form and the Jacobi identity of the associated quasi-Poisson bracket.

Theorem 14.3. *Let ω be a nondegenerate two-form and $\{\cdot, \cdot\}$ be its associated quasi-Poisson bracket. Then ω is closed if and only if the quasi-Poisson bracket satisfies the Jacobi-identity.*

Proof. We first derive the following general identity for any nondegenerate two form.

Lemma 14.4. *Let ω be as above. Then*

$$d\omega(X_f, X_g, X_h) = -(\{\{g,h\},f\} + \{\{h,f\},g\} + \{\{f,g\},h\}). \tag{14.4}$$

Proof. From the definition of the exterior derivative, we have

$$\begin{aligned} d\omega(X,Y,Z) &= X[\omega(Y,Z)] - Y[\omega(X,Z)] + Z[\omega(X,Y)] \\ &\quad -\omega([X,Y],Z) + \omega([X,Z],Y) - \omega([Y,Z],X). \end{aligned} \tag{14.5}$$

Substituting $X = X_f$, $Y = X_g$, $Z = X_h$, we derive

$$\begin{aligned} d\omega(X_f, X_g, X_h) &= X_f[\omega(X_g, X_h)] - X_g[\omega(X_f, X_h)] + X_h[\omega(X_f, X_g)] \\ &\quad -\omega([X_f, X_g], X_h) + \omega([X_f, X_h], X_g) - \omega([X_g, X_h], X_f). \end{aligned}$$

The first line becomes

$$\{\{g,h\},f\} + \{\{h,f\},g\} + \{\{f,g\},h\}$$

from the definition of the bracket. On the other hand, we compute

$$\begin{aligned} -\omega([X_f, X_g], X_h) &= dh([X_f, X_g]) = [X_f, X_g](h) = (L_{X_f}L_{X_g} - L_{X_g}L_{X_f})(h) \\ &= \{\{h,g\},f\} - \{\{h,f\},g\} \end{aligned}$$

and similarly

$$\begin{aligned} \omega([X_f, X_h], X_g) &= -\{\{g,h\},f\} + \{\{g,f\},h\} \\ -\omega([X_g, X_h], X_f) &= \{\{f,h\},g\} - \{\{f,g\},h\}. \end{aligned}$$

By adding them up, we have derived (14.4). □

It immediately follows from Lemma 14.4 that closedness of ω implies the Jacobi identity.

The converse also follows from (14.4) together with the ampleness of the set of quasi-Hamiltonian vector fields (Exercise 14.2) whose detail is in order. We first get

$$d\omega(X_f, X_g, X_h) = -\{\{f,g\},h\} - \{\{h,f\},g\} - \{\{g,h\},f\} = 0$$

for all f, g, h from (14.5). At any point $x \in M$, we evaluate $d\omega(u,v,w)$ against the three quasi-Hamiltonian vector fields X_f, X_g, X_h satisfying $X_f(x) = u$, $X_g(x) = v$, $X_h(x) = w$. This proves that the Jacobi-identity implies the closedness of ω and finishes the proof of the theorem. □

In symplectic canonical coordinates $(q_1, \cdots, q_n, p_1, \cdots, p_n)$ i.e., one for which

$$\{q_i, p_i\} = \delta_{ij},\ \{q_i, q_j\} = 0 = \{p_i, p_j\}$$

which is equivalent to saying

$$\omega = \sum_{i=1}^n dq_i \wedge p_i$$

, the Hamiltonian vector field X_H is given by

$$X_H = \sum_{i=1} \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i}.$$

Question 14.5. For given symplectic form ω, can we find a canonical coordinates?

Theorem 14.6 (Darboux Theorem)**.** *For any symplectic manifold (M, ω), there is a (symplectic) canonical coordinates $(q_1, \cdots, q_n, p_1, \cdots, p_n)$ under which we have*

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

Proof. We postpone its proof until the next section. □

Remark 14.7. One may ask similar question for a Riemannian manifold (M, g): Can we find a coordinate $(x^1, \cdots, x^n)$ such that $g = \sum_{i=1} dx^i dx^i$? There is an obstruction to the existence of such a coordinate. Riemann proved that such a coordinate exists if and only if g is flat. For the quasi-symplectic form ω, closedness thereof is the only condition for such an existence result. In this regard, symplectic geometry is much more topological than Riemannian geometry in that all symplectic forms are locally isomorphic if there ranks are the same.

Definition 14.8. Let M be a smooth manifold. A *symplectic structure* on M is a differential two-form ω which is closed and nondegenerate on M. The pair (M, ω) is called a symplectic manifold.

Example 14.9.
(1) Any cotangent bundle T^*N with the canonical symplectic form,
(2) Any surface equipped with an area form,
(3) Complex projective space $\mathbb{C}P^n$ equipped with the Fubini-Study form
$$\Omega_n = \frac{i}{2\pi} \frac{\sum_{0 \leq k,\, \ell \leq n} (z_k dw_\ell - z_\ell dw_k)(\overline{z}_k d\overline{z}_\ell - \overline{z}_\ell d\overline{z}_k)}{\sum_{k=0}^{n} |z_k|^2}$$
for the homogeneous coordinates $[z_0; z_1; \cdots; z_n]$ of $\mathbb{C}P^n$.
(4) Any complex submanifold $M \subset \mathbb{C}P^N$ for some N equipped with $\omega = i^*\Omega_N$.
(5) (Gompf, 1995) For any finitely presented group G, there exists a compact symplectic four-manifold M such that $G = \pi_1(M)$ where $\pi_1(M)$ is the fundamental group.

15. Proof of Darboux' Theorem

In this section, we provide the proof of Darboux theorem. It consists of two steps:

- Some linear algebra,
- Moser's deformation argument.

15.1. **Symplectic linear algebra.** Let (S, Ω) be a vector space with a symplectic bilinear form, i.e., a nondegenerate skew-symmetric bilinear form.

Definition 15.1. We call a linear map $A : S \to S$ a symplectic linear map if $A^*\Omega = \Omega$. Denote by $\mathrm{Sp}(S, \Omega)$ the set of symplectic automorphisms of (S, Ω).

It follows that each element $A \in \mathrm{Sp}(S)$ is invertible and so $\mathrm{Sp}(S) \subset \mathrm{GL}(S)$.

Lemma 15.2. *There exists a basis $\{e_1, \ldots, e_n, f_1, \ldots, f_n\}$ such that*

$$\Omega(e_i, e_j) = 0 = \Omega(f_i, f_j), \quad \Omega(e_i, f_j) = \delta_{ij}.$$

We call any such basis a symplectic basis or a Darboux basis.

This lemma says that any (S, Ω) is isomorphic to the canonical symplectic vector space

$$S_V := V \oplus V^* (\cong T^*V)$$

with the canonical symplectic inner product ω_V defined by

$$\Omega_V((v_1, \alpha_1), (v_2, \alpha_2)) = \alpha_2(v_1) - \alpha_1(v_2).$$

For S_V, any choice of basis $\{u_1, \ldots, u_n\}$ of V and its dual basis $\{u_1^*, \ldots, u_n^*\}$ for V^*,

$$e_i = (u_i, 0), \quad f_j = (0, u_j^*)$$

form a symplectic basis. This choice of the basis provides an isomorphism

$$(V \oplus V^*, \Omega_V) \cong \left(\mathbb{R}^n \oplus (\mathbb{R}^n)^*, \sum_{i=1}^n dq_i \wedge dp_i =: \omega_0 \right)$$

by considering the linear coordinates $(q_1, \ldots, q_n, p_1, \ldots, p_n =: \omega_0)$ associated to the basis $\{e_1, \ldots, e_n, f_1, \ldots, f_n\}$. We denote by

$$\mathrm{Sp}(\mathbb{R}^{2n})$$

the automorphism group of $(\mathbb{R}^{2n}, \omega_0)$.

Another consequence of the above lemma is that any two symplectic forms (V_1, Ω_1), (V_2, Ω_2) are isomorphic to each other.

15.2. **Moser's deformation method.** Let $p \in M$ be given and consider a coordinate chart $\varphi : U \to \mathbb{R}^{2n}$ centered at p. We consider the symplectic form

$$\omega' = \varphi^* \omega_0.$$

We are given two symplectic forms ω' and ω. After composing a linear transformation $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ with φ we can achieve

$$\omega'|_p = \omega|_p$$

on T_pM as a symplectic inner product. We still denote the resulting composition φ by re-choosing the coordinate chart φ.

Now we apply *Moser's deformation method* to construct a local diffeomorphism $\psi : V \to U$ such that

$$\psi^* \omega' = \omega, \quad \psi(p) = p$$

on a possibly smaller neighborhood $V \subset U$.

For this purpose, we consider the forms

$$\omega_t := (1-t)\omega + t\omega', \quad t \in [0,1] \tag{15.1}$$

Since $\omega_t|_p = \omega|_p$ for all $t \in [0,1]$, we may assume that ω_t are all nondegenerate on V' by shrinking V' if necessary.

We then will construct a family of local diffeomorphisms ψ_t with

$$\psi_0 = id, \quad \psi_t(p) = p$$

such that

$$\psi_t^* \omega_t = \omega, \quad \omega_0 = \omega, \, \omega_1 = \omega' \tag{15.2}$$

for all $t \in [0,1]$ on some even smaller neighborhood V' such that $\overline{V}' \subset V$. It will be enough to find the vector field X_t whose flow is given by ψ_t. To find such a vector field, we differentiate (15.2) to find the necessary condition for X_t to satisfy

$$\psi_t^* (\mathcal{L}_{X_t} \omega_t + \frac{\partial \omega_t}{dt}) = 0.$$

Since ψ_t are diffeomorphisms fixing the point p and $[0,1]$ is compact, we can find a neighborhood V' of p such that $\overline{\psi_t(V')} \subset V$ and so

$$\mathcal{L}_{X_t}\omega_t + \frac{\partial \omega_t}{dt} = 0$$

on V'. But we have $\frac{\partial \omega_t}{dt} = \omega' - \omega$ and hence the equation is reduced to

$$d(X_t \rfloor \omega_t) = \omega - \omega'$$

since ω_t are closed forms. Therefore to determine X_t, we have only to solve the following system of equations

$$\begin{cases} \beta = X_t \rfloor \omega_t, \\ d\beta = \omega - \omega' =: \Omega. \end{cases}$$

We now apply Poincaré's lemma to solve $d\beta = \omega - \omega'$ for β: Let $\eta_s : V' \to V'$ be the deformation retraction of V' to the point $p \in V'$. Then we have

$$\Omega = \eta_1^*\Omega = \eta_0^*\Omega + \int_0^1 \frac{d}{ds}\eta_s^*\Omega\, ds = \int_0^1 \eta_s^*(d(Y_s \rfloor \Omega))\, dt = d\left(\int_0^1 \eta_s^*(Y_s \rfloor \Omega)\, ds\right).$$

From this, we obtain

$$\beta = \int_0^1 \eta_s^*(Y_s \rfloor \Omega)\, ds.$$

Since η_s fix p and $\Omega|_p = 0$, it follows that $\beta|_p = 0$. We solve the equation

$$X_t \rfloor \omega_t = \beta$$

using the nondegeneracy of ω_t on V'. Since $\beta|_p = 0$, $X_t(p) = 0$ for all t.

It remains to prove that the ODE $\dot{x} = X_t(x)$ has its domain $\mathcal{D} \subset \mathbb{R} \times V$ of existence contains the tube

$$[0,1] \times V''$$

for some sufficiently small neighborhood of p such that $\overline{V''} \subset V$ so that ψ_t are defined on V'' for $t \in [0,1]$. But since $X_t(p) = 0$ for all $t \in [0,1]$, $\mathcal{D}$ contains $[0,1] \times \{p\}$. Since $\mathcal{D}$ is open in $\mathbb{R} \times V$ and $[0,1] \times \{p\}$ is compact, $\mathcal{D}$ contains an open set $(-\epsilon, 1+\epsilon) \times V''$, which finishes the construction of ψ satisfying

$$\psi^*\omega' = \omega$$

which in turn implies

$$\psi^*(\varphi^*\omega_0) = \omega$$

on V''. By considering a new coordinate chart

$$\widetilde{\varphi} = \varphi \circ \psi : V'' \to \mathbb{R}^{2n}$$

we have finished the proof of Darboux Theorem. □

Corollary 15.3. *For any symplectic manifold (M,ω), there exists some $r > 0$ such that there exists a symplectic embedding $B^{2n}(r) \to M$.*

Definition 15.4 (Gromov capacity)**.** The *Gromov capacity* denoted by $c_G(M,\omega)$ is defined to be

$$c_G(M,\omega) = \sup\{\pi r^2 \mid \exists \text{ a symplectic embedding } \phi : B^{2n}(r) \to M\}.$$

Definition 15.5. Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds. We call a differentiable map $f : M_1 \to M_2$ a symplectic map if $f^*\omega_2 = \omega_1$, and a symplectic diffeomorphism if f is a diffeomorphism in addition. We say that two symplectic forms are diffeomorphic to each other if there exists a symplectic diffeomorphism between them.

Note that any symplectic map must be an immersion and so $\dim M_1 \leq \dim M_2$.

Obviously, we have $c_G(M_1, \omega_1) \leq c_G(M_2, \omega_2)$ if there exists a symplectic embedding $(M_1, \omega_1) \to (M_2, \omega_2)$.

Theorem 15.6 (Gromov's nonsqueezing theorem)**.** *Let $B^{2n}(R)$ be a standard closed ball of radius R and $Z^{2n}(r) = D^2(r) \times \mathbb{C}^{n-1}$ the cylinder over $D^2(r) \subset \mathbb{C}$ as subsets of $\mathbb{C}^n \cong (\mathbb{R}^2 n, \omega_0)$. Then there exists a symplectic embedding $B^{2n}(R) \to \operatorname{Int} Z^{2n}(r)$ if and only if $r > R$. In particular $c_G(Z^{2n}(r)) = \pi r^2$.*

16. Hamiltonian vector fields and diffeomorhpisms

We denote by $\operatorname{Symp}(M, \omega)$ the group of symplectic self-diffeomorphisms.

Exercise 16.1. Let $\phi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be a diffeomorphism. We call ϕ a *canonical transformation* if the new coordinate system $(Q_1, \ldots, Q_n, P_1, \cdots, P_n)$ defined by

$$Q_i = q_i \circ \phi,\ P_j = p_j \circ \phi$$

is also canonical. Prove that ϕ is canonical if and only if ϕ is symplectic with respect to the canonical symplectic structure $\omega_0 = \sum_{i=1}^n dq_i \wedge dp_i$.

Definition 16.2. Two diffeomorphisms ϕ, $\psi : M_1 \to M_2$ are called symplectically isotopic if there exists an isotopy ϕ_t such that $\phi_t^*\omega_2 = \omega_1$ for all $t \in [0, 1]$ and $\phi_0 = \phi$ and $\phi_1 = \psi$. We call such an isotopy a symplectic isotopy.

We denote by $\operatorname{Symp}_0(M, \omega)$ the set of symplectomorphisms of (M, ω) that is symplectically isotopic to the identity.

Definition 16.3. A vector field X on M is called a symplectic (or locally Hamiltonian) vector field if its flow ϕ_t is a symplectic isotopy.

A straightforward calculation shows the following.

Proposition 16.4. *Let ϕ_t with $t \in [0, 1]$ is an isotopy with ϕ_0 symplectic. Denote by X_t the associated vector field*

$$X_t = \frac{\partial \phi_t}{\partial t} \circ \phi_t^{-1}.$$

Then ϕ_t is symplectic for all t if and only if $X_t \rfloor \omega$ is a closed one-form.

Proof. Suppose that ϕ_t is symplectic and so $\phi_t^*\omega = \omega$ for all t. We have the Lie derivative formula

$$\frac{d}{dt}\phi_t^*\omega = \phi_t^*(\mathcal{L}_{X_t}\omega).$$

Therefore by differentiating $\omega = \phi_t^*\omega$ in t, we obtain $0 = \phi_t^*(\mathcal{L}_{X_t}\omega)$. Since ϕ_t is a diffeomorphism, this implies

$$\mathcal{L}_{X_t}\omega = 0$$

for all t. On the other hand, by Cartan's formula, we obtain

$$\mathcal{L}_{X_t}\omega = d(X_t \rfloor \omega) + X_t \rfloor d\omega = d(X_t \rfloor \omega)$$

Combining the two, we have obtained $d(X_t\rfloor\omega) = 0$, i.e., $X_t\rfloor\omega$ is closed.

Conversely suppose that $X_t\rfloor\omega$ is closed. By reading the above proof backwards, we have obtained

$$\frac{d}{dt}\phi_t^*\omega = 0.$$

Integrating this out over $t \in [0,1]$ and applying the fundamental theorem of calculus, we obtain

$$\phi_1^*\omega = \phi_0^*\omega + \int_0^1 \left(\frac{d}{dt}\phi_t^*\omega\right) dt = \phi_0^*\omega = \omega$$

where the last equality holds by the assumption that ϕ_0 is symplectic. This finishes the proof. □

Definition 16.5. We call a vector field X *Hamiltonian* if the one-form $X\rfloor\omega$ is exact and call its primitive h a Hamiltonian function of X, i.e., $dH = X\rfloor\omega$. We denote by $X = X_H$ the Hamiltonian vector of H.

An isotopy ϕ^t is called a *Hamiltonian isotopy* if its associated vector field X_t is Hamiltonian for all t. We call a function $H = H(t,x)$ satisfying $X_t = X_{H_t}$ its time-dependent Hamiltonian.

By a slight abuse of notation, we still denote by X_H the time dependent vector field

$$X_H(t,x) := X_{H_t}(x)$$

and its generated isotopy by ϕ_H the isotopy given by $t \mapsto \phi_H^t$ which is the time-dependent flow of the non-autonomous first order ODE $\dot{x} = X_H(t,x)$ which we call *Hamilton's equation.*

Definition 16.6. A symplectic diffeomorphism ϕ is called Hamiltonian if $\phi = \phi_H^1$ for some Hamiltonian $H = H(t,x)$ for $t \in [0,1]$. We denote by

$$\mathrm{Ham}(M,\omega) \subset \mathrm{Symp}(M,\omega)$$

the set of Hamiltonian diffeomorphisms.

Proposition 16.7. *The set* $\mathrm{Ham}(M,\omega)$ *is a subgroup of* $\mathrm{Symp}(M,\omega)$.

Proof. Clearly $id \in \mathrm{Ham}(M,\omega)$ as the constant function generates constant flow. Now let ϕ, $\psi \in \mathrm{Ham}(M,\omega)$. By definition, there are Hamiltonians H, K such that $\phi = \phi_H^1$, $\psi = \phi_K^1$. Obviously the flow $t \mapsto \phi_H^t\phi_K^t$ generates the product $\phi\psi$. Therefore it is is enough to show that this flow is Hamiltonian. For this purpose we compute its generating vector field

$$\frac{d}{dt}(\phi_H^t\phi_K^t) \circ (\phi_H^t\phi_K^t)^{-1}$$

explicitly. We show that this is the Hamiltonian vector field X_L^t with the Hamiltonian $L = L(t,x)$ given by

$$L(t,x) = H(t,x) + K(t, (\phi_H^t)^{-1}(x)).$$

We denote this Hamiltonian by $H\#K$.

Similarly we show that the flow $t \mapsto (\phi_H^t)^{-1}$ is generated by $\overline{H}$ defined by

$$\overline{H}(t,x) = -H(t, \phi_H^t(x)).$$

This finishes the proof. □

The remarkable Hofer's pseudo-norm of Hamiltonian diffeomorphisms is defined by

$$\|\phi\| = \inf_{H \mapsto \phi} \|H\| = \text{leng}(\phi_H) \tag{16.1}$$

where $\|H\|$ is the $L^{1,\infty}$-norm of $H = H(t,x)$

$$\|H\| = \int_0^1 (\max H_t - \min H_t)\, dt.$$

Proposition 16.8. *Let ϕ, $\psi \in Ham^c(M,\omega)$. Then the following holds:*

(1) *(Symmetry)* $\|\phi\| = \|\phi^{-1}\|$
(2) *(Triangle inequality)* $\|\phi\psi\| \leq \|\phi\| + \|\psi\|$
(3) *(Symplectic invariance)* $\|h\phi h^{-1}\| = \|\phi\|$ *for all* $h \in Symp(M,\omega)$

Proof. We recall the formula for the inverse Hamiltonian $\overline{H}(t,x) = -H(t, \phi_H^t(x))$ generating $(\phi_H^t)^{-1}$. Obviously we have

$$\max_x \overline{H}_t = -\min_x H_t, \quad \min_x \overline{H}_t = -\max_x H_t.$$

It then follows $\|\overline{H}\| = \|H\|$. Furthermore we know $\overline{(\overline{H})} = H$. Combining these we derive (1) by taking the infimum of $\|\overline{H}\| = \|H\|$ over all $H \mapsto \phi$.

For (2), we recall the composition Hamiltonian $H\#F$ defined by

$$H\#F(t,x) = H(t,x) + F(t, (\phi_H^t)^{-1}(x))$$

generates $\phi\psi$ if $H \mapsto \phi$ and $F \mapsto \psi$. The latter fact in particular implies

$$\|\phi\psi\| \leq \inf_{H\mapsto\phi, F\mapsto\psi} \|H\#F\|. \tag{16.2}$$

On the other hand, we have

$$\begin{aligned} \max_x (H\#F)_t &\leq \max_x H_t + \max_x F_t \\ -\min_x (H\#F)_t &\geq -\min_x H_t - \min_x F_t. \end{aligned}$$

Then we derive

$$\begin{aligned} \|H\#F\| &= \int_0^1 \Big(\max_x (H\#F)_t - \min_x (H\#F)_t\Big)\, dt \\ &\leq \int_0^1 (\max_x H_t + \max_x F_t - \min_x H_t - \min_x F_t)\, dt \\ &= \|H\| + \|F\|. \end{aligned} \tag{16.3}$$

Combining (16.2) and (16.3), we have proved (2).

For the proof of (3), we first recall that for given Hamiltonian H the pull-back $h^*H = H \circ h$ generates $h^{-1} \circ \phi_H^t \circ h$ for any symplectic diffeomorphism $h : M \to M$. Obviously we have $\|H\| = \|H \circ h\| = \|h^*H\|$. Taking the infimum over all $H \mapsto \phi$, we have proved (3). ☐

Theorem 16.9 (Hofer, Polterovich, Lalonde-McDuff)**.** *The norm $\|\cdot\|$ is nondegenerate, i.e.,*

(4) *(Nondegendracy)* $\|\phi\| = 0$ *if and only if* $\phi = id$.

This norm, called the Hofer norm, then gives rise to a natural metric topology on $Ham(M,\omega)$, which we call the *Hofer topology* of $Ham(M,\omega)$.

A crucial ingredient in the proof of this nondegeneracy that Hofer introduced is the following

Definition 16.10 (Displacement energy)**.** Let $A \subset (M, \omega)$ be a relatively compact closed subset. The *displacement energy* of A, denoted by $e(A)$, is defined to be

$$e(A) := \inf\{\|H\| \mid A \cap \phi_H^1(\overline{A}) = \emptyset\}.$$

Outline of the proof. Enough to show that if $\phi \neq id$, then $\|\phi\| > 0$. Since $\phi \neq id$, there is a point $p \in M$ such that $\phi(p) \neq p$. By continuity of ϕ, there exists a small symplectic ball B such that $p \in \operatorname{Int} B$ such that $B \cap \phi(B) = \emptyset$. Therefore it will be enough to prove $e(B) > 0$. This is an immediate consequence of the following theorem whose proof relies on the machinery of pseudoholomorphic curves and Hamiltonian dynamics.

Theorem 16.11 (Hofer, Lalonde-McDuff, Usher)**.** *Let B be a symplectic ball of radius R, i.e., $B = \phi(B^{2n}(R))$ in (M, ω). Then $e(B) = c_G(B)$.*

Corollary 16.12. *Provided $c_G(B) < c_G(M, \omega)$, then $e(B) > 0$. In particular nondegeneracy of $\| \cdot \|$ holds.*

This finishes the proof. □

17. Autonomous Hamiltonians and conservation law

We specialize to the case of autonomous Hamiltonian $H = H(x)$, i.e., time-independent function. The assignment $H \mapsto X_H$ for autonomous Hamiltonians induces the following exact sequence: Suppose M is compact connected without boundary. Then we have an exact sequence

$$0 \to C^\infty(M) \to \operatorname{symp}(M, \omega) \to H^1(M, \mathbb{R}) \to 0$$

where the first map in the middle is $H \to X_H$ and the second map is $X \to [X \rfloor \omega]$. The image of the first map is precisely the set of Hamiltonian vector fields $\operatorname{ham}(M, \omega)$ and we have

$$\frac{\operatorname{symp}(M, \omega)}{\operatorname{ham}(M, \omega)} \cong H^1(M, \omega).$$

Furthermore we have the following fundamental conservation law.

Proposition 17.1 (Conservation law)**.** *Let $H = H(x)$ be time-independent and ϕ_H^t be its flow. Then $H(\phi_H^t(x)) = H(x)$ for all $(t, x) \in \mathbb{R} \times M$.*

Definition 17.2 (Characteristic flow)**.** Let H be an autonomous Hamiltonian and c be its regular value. We call the induced flow of X_H on $H^{-1}(c)$ the characteristic flow and each trajectory of X_H a characteristic (curve) of H in $H^{-1}(c)$.

Example 17.3 (Geodesic flow)**.** Let (N, g) be a Riemannian manifold and $E : TN \to \mathbb{R}$ be the kinetic energy density function defined by

$$E(q, v) = \frac{1}{2} g(v, v) = \frac{1}{2} |v|_g^2.$$

Consider the bundle isomorphism $\widetilde{g} : TN \to T^*N$ given by

$$v \mapsto g(v, \cdot).$$

Equip $M = T^*N$ with the canonical symplectic form $\omega_0 = \sum_{i=1} dq_i \wedge dp_i$. Consider the (autonomous) Hamiltonian $H : T^*N \to \mathbb{R}$ given by

$$H_g = E \circ \widetilde{g}^{-1}$$

and its associated Hamiltonian vector field X_{H_g}. Then we have

$$\widetilde{g}_* G = X_{H_g}$$

where G is the geodesic vector field of g on TN.

More generally, we consider another autonomous function G that is invariant under the Hamiltonian flow of X_H. Such a conserved quantity is called a *first integral* of H. This conservation law enables us to reduce the number of variables in solving the Hamilton's equation $\dot{x} = X_H(x)$ by restricting the system to the level set

$$G^{-1}(c).$$

Theorem 17.4 (Reduction of variables). *Let (M, ω) be a symplectic manifold. Let c be its regular value of H. Denote by $M_{H,c}$ the set of leaves of $H^{-1}(c)$. Suppose that $M_{H,c}$ is Hausdorff. Then $M_{H,c}$ carries a natural symplectic form ω_c uniquely determined by the equality*

$$\pi_c^* \omega_c = i_c^* \omega$$

where $i_c : H^{-1}(c) \hookrightarrow M$ and $\pi_c : H^{-1}(c) \to M_{H,c}$ are the canonical injection and projection respectively.

Proof. Since c is a regular value $H^{-1}(c)$ is a smooth manifold. Furthermore it carries a nowhere vanishing vector field induced by X_H and so defines a one-dimensional foliation thereon.

Let $p \in H^{-1}(c) \subset M$ and consider $i_c^*\omega|_p$ thereon. This is a closed two form and has one-dimensional kernel. The Hamiltonian vector field X_H is tangent to $H^{-1}(c)$ and

$$T_{[p]} M_{H,c} \cong T_p G^{-1}(c)/\mathbb{R}\langle X_H(p)\rangle.$$

We define a two-form

$$\omega_{H,c}(\widetilde{v}, \widetilde{w}) = \omega(v, w) \tag{17.1}$$

for any lifts v, $w \in T_p G^{-1}(c)$ of $\widetilde{v}$, $\widetilde{w} \in T_{[p]} M_{H,c}$. Check that this form is well-defined and satisfies $\pi_c^*\omega_c = i_c^*\omega$. Once this is satisfied, we have $\pi_c^* d\omega_c = i_c^* d\omega = 0$. Since π_c is a submersion, we obtain $d\omega_c = 0$. Nondegeneracy is obvious.

The proof of Statement (2) is easy to prove. □

Exercise 17.5. Prove that the definition (17.1) is well-defined by showing that the right hand side does not depend on the choice $p \in [p]$ and of the lifts v, $w \in T_p G^{-1}(c)$.

We call the above reduction process a *symplectic reduction.*

Remark 17.6. We would like to point out that the leaf space $M_{H,c}$ may not be Hausdorff, but the reduced symplectic form still make sense.

Proposition 17.7. *Suppose that F, G are two first integrals of the Hamiltonian H. Then $\{F, G\}$ is also a first integral of H.*

In this way, we can construct more independent conserved quantities *as long as* $\{F, G\} = 0$.

Definition 17.8. We say two functions F, G are *in involution* if $\{F, G\} = 0$.

If F, G are first integrals of H that are in involution, we can repeat the above symplectic reduction to further reduce the variables. This can be repeated at most n times, the number of freedoms in configuration coordinates.

Definition 17.9. A Hamiltonian system of $H = H(x)$ is called *completely integrable* if it admits first integrals $F_1, \cdots, F_n$ in involution such that the map

$$x \mapsto (F_1(x), \cdots, F_n(x)) =: \Phi(x)$$

is a submersion on a dense open subset of M.

When a given mechanical system is completely integrable, there are two ways of obtaining the first integrals:

- Find a canonical transformation $(q, p) \to (Q, P)$ such that Q_i are the first integrals of the given Hamiltonian H – Hamilton-Jacobi method,
- Find maximal possible number of independent (commuting) symmetries – Lie group action.

We start with the Hamilton-Jacobi method in the next section.

18. Completely integrable systems and action-angle variables

(Reference: Arnold's book "Mathematical Methods of Classical Mechanics, 2nd edition, Section 49 - 50)

In this section, we explain a general theorem of completely integrable system, sometimes called the *Liouville theorem* or *Arnold-Avez theorem*.

We start with the definition of the following fundamental geometric objects, Lagrangian submanifolds in symplectic geometry.

Definition 18.1. Let (M, ω) be a symplectic manifold.

(1) A submanifold $L \subset M$ is called *isotropic* (resp. coisotropic) if $T_x L$ is isotropic (resp. if $(T_x L)^\omega$ is isotropic) in $T_x M$.
(2) A submanifold $L \subset (M, \omega)$ is called *Lagrangian* if L is maximally isotropic, i.e., if $\omega|_L = 0$ and $\dim L = n$. More generally, a map $i : N \to M$ is called a Lagrangian immersion if $\dim N = n$ and $i^*\omega = 0$.

Theorem 18.2 (Darboux-Weinstein theorem)**.** *Let $L \subset (M, \omega)$ be any Lagrangian submanifold. There exist neighborhoods $U \subset M$ of L and $V \subset T^*L$ of the zero section $Z = o_{T^*L}$, and a diffeomorphism*

$$\Phi : U \to V$$

such that $\omega = \Phi^*\omega_0$ on U. We call Φ a Darboux-Weinstein chart and U a Weinstein neighborhood of L.

We denote by $(q_1, \ldots, q_n, p_1, \cdots, p_n)$ the canonical coordinates of U associated to a coordinates $(q_1, \ldots, q_n)$ of L (with respect to the chart Φ).

The proof is a generalization of Moser's deformation method whose proof will be left as a homework.

Corollary 18.3. *There exists a neighborhood U of any Lagrangian submanifold $L \subset (M, \omega)$ the restriction of ω to U is exact.*

Liouville proved that any completely integrable system *can be solved by quadratures.* More precisely, we have the following whose proof will be occupied by this section.

Theorem 18.4 (Liouville, Arnold-Avez)**.** *Suppose that we are given n functions $\{F_1, \ldots, F_n\}$ in involution on a symplectic (M^{2n}, ω). Consider the map $\Phi : M \to \mathbb{R}^n$ defined by $\Phi = (F_1, \ldots, F_n)$ and a regular value $\mathbf{c} = (c_1, \ldots, c_n)$ of Φ. Consider the level set $M_\mathbf{c} := \Phi^{-1}(\mathbf{c})$. Then*

1. *$M_\mathbf{c}$ is a smooth manifold invariant under the flow of the Hamiltonian $H = F_1$.*
2. *If $M_\mathbf{c}$ is compact and connected, then it is diffeomorphic to the n-dimensional torus T^n.*
3. *The phase flow of the Hamiltonian H determines a conditionally periodic motion on $M_\mathbf{c}$, i.e., in angular coordinates $\varphi = (\varphi_1, \ldots, \varphi_n)$ we have*
$$\frac{d\varphi}{dt} = \omega, \quad \omega = \omega(\mathbf{c}).$$
4. *The canonical equations with the Hamiltonian H can be integrated by quadratures in the action-angle variable coordinates.*

We start with the following lemma

Lemma 18.5. *$TM_\mathbf{c}$ carries a global frame of vector fields $X_1, \cdots, X_n$ such that $[X_i, X_j] = 0$.*

Proof. Since $dF_i(x)$ are linearly independent at all $x \in M_\mathbf{c}$ and $F_i \equiv c_i$ constants, we have only to set $X_i = X_{F_i}|_{M_\mathbf{c}}$. □

In particular, the flows X_i on $M_\mathbf{c}$ commute one another, and $M_\mathbf{c}$ is a Lagrangian submanifold.

18.1. **Construction of angle coordinates.** If we denote by $g_i^t : M_\mathbf{c} \to M_\mathbf{c}$ the phase flow of X_i on $M_\mathbf{c}$, we have a well-defined smooth map
$$g^\mathbf{t} : \mathbb{R}^n \to M_\mathbf{c}; \quad (t_1, t_2, \ldots, t_n) \mapsto g_1^{t_1} \cdots g_n^{t_n}.$$
Then we have $g^\mathbf{t} g^\mathbf{x} = g^{\mathbf{t}+\mathbf{s}}$. Therefore we have a map
$$\phi_{x_0} : \mathbb{R}^n \to M_\mathbf{c}; \quad \phi_{x_0}(\mathbf{t}) = g^\mathbf{t}(x_0)$$
for each $x_0 \in M_\mathbf{c}$. It follows that ϕ_{x_0} is a submersion.

Lemma 18.6. *If $M_\mathbf{c}$ is compact and connected, then it is diffeomorphic to an n-torus T^n.*

Proof. Since ϕ_{x_0} is a submersion, it is an open map. Since $M_\mathbf{c}$ is compact, it is also a closed map. Therefore the map is surjective since $M_\mathbf{c}$ is connected.

We consider the isotropy group of x_0
$$\Gamma = \phi_{x_0}^{-1}(x_0) := \{\mathbf{t} \in \mathbb{R}^n \mid \phi_{x_0}(\mathbf{t}) = x_0\}.$$
It follows that this set does not depend on the choice of x_0 and that $0 \in \Gamma$.

Since $M_\mathbf{c}$ is compact, Γ is a discrete subgroup of $\mathbb{R}^n$ such that the map ϕ_{x_0} descends to
$$\pi_{x_0} : \mathbb{R}^n/\Gamma \to M_\mathbf{c}$$
which is a diffeomorphism. This finishes the proof. □

Exercise 18.7. Prove that Γ is indeed a discrete subgroup of $\mathbb{R}^n$ isomorphic to the integral lattice $\mathbb{Z}^n \subset \mathbb{R}^n$ by proving that there is an open neighborhood of the origin $0 \in \mathbb{R}^n$ such that $U \cap \Gamma = \{0\}$.

Next we take an integral basis $\{\mathbf{e_1}, \dots, \mathbf{e_n}\}$ of Γ which satisfies

$$g^{\sum_{i=1}^n k_i \mathbf{e}_i}(x_0) = x_0$$

where $\Gamma = \text{span}_{\mathbb{Z}}\{\mathbf{e_1}, \dots, \mathbf{e_n}\} \subset \mathbb{R}^n$. The associated angular coordinates $\varphi = (\varphi_1, \dots, \varphi_n)$ of $M_{\mathbf{c}}$ is uniquely determined ($\mod 2\pi$) by the requirement that the map

$$(\varphi_1, \cdots, \varphi_n) \mapsto g^{\sum_{i=1}^n \varphi_i \mathbf{e}_i}(x_0)$$

is an $\mathbb{R}^n$-equivariant submersion and is one-one on $[0, 2\pi)^n$.

By definition, under the action of the phase flow of $H = F_1$, we have

$$\dot{\varphi}_i \equiv \omega_i, \quad \omega_i = \omega_i(\mathbf{c})$$

and so $\varphi(t) = \varphi(0) + \omega t$ for some vector $\omega \in \mathbb{R}^n$, called the *angular velocity.*

So far we have solved the problem for a given value $\mathbf{c}$. To cover the whole phase space, we now vary $\mathbf{c}$ and perform the construction smoothly in $\mathbf{c}$. In terms of the map

$$\Phi : M \to \mathbb{R}^n,$$

we consider a local trivialization of Φ on a neighborhood $B \subset \mathbb{R}^n$ of the given regular value $\mathbf{c}$,

$$\Phi^{-1}(B) \to M_{\mathbf{c}} \times B \cong T^n \times B.$$

By identifying the neighborhood with Weinstein neighborhood of $M_{\mathbf{c}}$, we put a canonical coordinates $(q_1, \cdots, q_n, p_1, \cdots, p_n)$ on $T^n \times B \subset T^*T^n$. Denote by $\mathbf{f}$ a point in B and by $\omega(\mathbf{f}$ the associated angular velocity for $M_{\mathbf{f}}$.

Denoting by φ also the lift of the angle coordinates of $M_{\mathbf{c}}$ to $M_{\mathbf{c}} \times B \to B$ under the cotangent projection $\pi : T^*M_{\mathbf{c}} \to M_{\mathbf{c}}$, we have achieved a coordinate system

$$(\varphi, \mathbf{F}) = (\varphi_1, \dots, \varphi_n, F_1, \cdots, F_n).$$

In this coordinates, the phase flow of $H = F_1$ on $\Phi^{-1}(B)$ is given by

$$\frac{d\mathbf{F}}{dt} = 0, \quad \frac{d\varphi}{dt} = \omega(\mathbf{F}),$$

which can be easily integrated along the phase flow of F_i:

$$\mathbf{F}(t) = \mathbf{F}(0), \quad \varphi(t) = \varphi(0) + \omega(\mathbf{F}(0))t.$$

Therefore in order to explicitly integrate the original system $\dot{x} = X_H(x)$, it suffices to find an explicit formula of the angle coordinates in terms of (q, p) by quadratures. For this purpose, we will need to adjust the definition of angle coordinates while we solving the equation

$$\omega = \sum_{i=1}^n d\varphi_i \wedge dI_i,$$

so that the coordinate system (φ, I) becomes a (Hamiltonian) canonical coordinate, which we call the *action-angle coordinates*. We will solve this in two steps:

- First, we will construct the action variables $\mathbf{I} = \mathbf{I}(\mathbf{F})$, $\mathbf{I} = (I_1, \dots, I_n)$ by quadratures,
- Then we will define the final angle coordinates $(\varphi_1, \dots, \varphi_n)$ by differentiation.

18.2. **Construction of action coordinates.** Let us start with the case of one degree of freedom.

18.2.1. *The case of one degree of freedom.*

Example 18.8. We illustrate the construction of action-angle coordinates in the case $\mathbb{R}^2$ of one degree of freedom for the harmonic oscillator Hamiltonian

$$H = \frac{1}{2}(q^2 + p^2)$$

which governs the dynamics of harmonic oscillator. In this case, $\Phi = H$ and H is regular everywhere except on $H^{-1}(0) = \{0\}$. Let $c_0 > 0$ which is a regular value of H and consider a open neighborhood (a, b) containing c_0 in $\mathbb{R}$.

We are looking for a symplectic map

$$\phi : H^{-1}(a, b) \to H^{-1}(c_0) \times (a, b) \cong S^1 \times (a, b)$$

such that (φ, I) becomes an action-angle variable in that

- $\phi_* \omega_0 = d\varphi \wedge dI$,
- $I = I(h), \quad \oint_{H^{-1}(h)} d\varphi = 2\pi$.

Actually in this case, it is very easy to construct the action-angle coordinates by noting that $H = \frac{1}{2}r^2$ for the polar coordinates (r, θ) of $\mathbb{R}^2 \setminus 0$ by noting that

$$dq \wedge dp = d\theta \wedge d(r^2/2) :$$

The action-angle coordinate of the harmonic oscillator $H = (q^2 + p^2)/2$ is nothing but $(\varphi, I) = (\theta, H)$. We also note

$$\text{Area}(\{H \leq h\}) = \int_{\{H \leq h\}} \omega_0 = \int_{\{H \leq \}} d\theta \wedge dI = 2\pi h$$

and hence we have

$$I(h) = \text{Area}(H \leq h).$$

In other words, *the action coordinate measures the area of the disc bounded by the circle* $H^{-1}(h)$.

For the general case of one degree of freedom we consider a Weinstein neighborhood $M_c \times B$ of M_c. We note that the form ω_0 is exact $M_c \times B \subset T^*M_c$, we fix a primitive Θ with $\omega_0 = -d\Theta$. Then we have a one-form $i_h^*\Theta$ on $M_h = H^{-1}(h)$ (which is (automatically) closed in the case of one degree of freedom).

Definition 18.9 (Period)**.** For each $h \in B = (c - \delta, c + \delta)$, we define

$$\Pi(h) := \int_{M_h} i_h^* \Theta$$

and call it the *period* of M_h.

Denote by (q, p) a canonical coordinate of $T^*\mathbb{R}$ also regarded as a canonical coordinate of $M_c \times B \subset T^*M_c$ with $q \mod 2\pi$.

We will find the *generating function* of the form

$$S = S(q, I) : \mathbb{R} \times \mathbb{R} \cong T^*\mathbb{R} \to \mathbb{R}$$

such that

(1) On each fixed level set $\{H = h\} \subset T^*\mathbb{R}$, we have $i_h^*\Theta = d_q(S|_{I=h})$ where $\Theta = pdq$ is the Liouville one-form,

(2) the map
$$q \mapsto \frac{\partial S^h}{\partial I}; \quad S^h := S|_{I=h}$$
is 2π-periodic on each level set $H^{-1}(h) =: M_h$.

(3) The Hamiltonian H factorizes into $H(q,p) = h(I(q,p))$.

The generating function $S = S(q, I)$ is one that satisfies
$$p = \frac{\partial S}{\partial q}, \quad \varphi = \frac{\partial S}{\partial I}:$$
If so, we derive
$$\begin{aligned} dq \wedge dp &= dq \wedge \left(\frac{\partial S}{\partial q}\right) = dq \wedge \frac{\partial}{\partial I}\left(\frac{\partial S}{\partial q}\right) dI \\ &= \frac{\partial}{\partial q}\left(\frac{\partial S}{\partial I}\right) dq \wedge \frac{\partial}{\partial q}\left(\frac{\partial S}{\partial I}\right) dI = d\varphi \wedge dI \end{aligned}$$
which does our purpose.

To find such a function $S = S(q, Q)$, we solve a (stationary) *Hamilton-Jacobi equation*
$$H\left(q, \frac{\partial S}{\partial q}(q)\right) = h$$
for each h which gives rise to h-family of functions, i.e., represent $M_h \subset H^{-1}(h)$ by the image Image dS^h in T^*M_c for a solution of the equation by S^h. We write
$$\widetilde{S}(q,h) := S^h(q).$$
The solution S^h is nothing but the primitive of $i_h^*\Theta$
$$i_h^*\Theta = dS^h$$
where $i_h : M_c \to T^*M_c$ is the map $i_h = (\pi|_{M_h})^{-1} : M_c \to T^*M_c$. We can find S^h by integration
$$\widetilde{S}^h(q) := \int_{q_0}^{q} i_h^*\Theta \tag{18.1}$$
for S with $q \in M_h$ with a fixed reference point $q_0(h) \in M_h$ for each h. Since $i_h^*\Theta$ is closed but may not be exact, the function $\widetilde{S}^h(q)$ may not be single-valued. Then the above definition (18.1) is well-defined modulo

- the period $\Pi(h)$ of the closed one-form $i_h^*\Theta$, and
- the choice of reference points $q_0(h)$, i.e., a choice of section
$$q_0 : B \to M_c \times B, \quad h \mapsto (q_0(h), h).$$

To fulfill the condition
$$\oint_{M_h} d\varphi = 2\pi.$$
We re-define I to be
$$I(h) = \frac{\Pi(h)}{2\pi}.$$
Now we invert the function $I = I(h)$ so that $h = h(I)$, and define
$$S(q, I) = \widetilde{S}(q, h(I)).$$

It leads to a multi-valued function

$$\frac{\partial S}{\partial I} =: \varphi \tag{18.2}$$

on $M_h \cong S^1$ which is well-defined modulo $\Pi(h(I))$.

This will complete construction of action coordinates I and hence the construction of action-angle coordinate (φ, I) for the one degree of freedom by setting φ to be (18.2).

18.2.2. *The case of higher degree of freedom.* Now let

$$M_{\mathbf{h}} = \Phi^{-1}(\mathbf{h}) = \{x \in M \mid F_i(x) = h_i\}$$

for a regular value h of Φ. We fix a sufficiently convex neighborhood $B \subset \mathbb{R}^n$ of $\mathbf{h}$ so that Φ is submersive on $\Phi^{-1}(B)$.

Lemma 18.10. *The pushforward form $\Phi_*\omega =: \omega'$ is exact on*

$$\Phi^{-1}(B) \cong M_{\mathbf{h}} \times B$$

for a contractible neighborhood of $\mathbf{h}$ in $\mathbb{R}^n$.

Proof. Let η_s be the deformation retraction of $\Phi^{-1})(B)$ to $\Phi^{-1}(\mathbf{h})$ given by

$$\eta_s(x) = s\Phi(x) + (1-s)\mathbf{h}.$$

Then we have $\eta_0^*\omega = 0$ and $\eta_1 = id$. Therefore

$$\omega' = d\left(\int_0^1 \eta_s^*(Y_s \rfloor \omega')\, ds\right)$$

which explicitly constructed a primitive

$$\Theta = \int_0^1 \eta_s^*(Y_s \rfloor \omega')\, ds.$$

□

Since ω is exact on a neighborhood of any Lagrangian submanifold (e.g., see Darboux-Weinstein theorem) we can write

$$\omega = -d\Theta$$

for a choice of primitive on $\Phi^{-1}(B)$. We may take $\Theta = pdq$ in the canonical coordinates of Darboux-Weinstein chart.

Lemma 18.11. *For each $\mathbf{h} = (h_1, \ldots, h_n) \in B \subset \mathbb{R}^n$, the form $i_{\mathbf{h}}^*\Theta$ is closed with the map $i_{\mathbf{h}} : M_{\mathbf{h}} \to \Phi^{-1}(B)$ is the inclusion map.*

Proof. This follows from the property that $M_{\mathbf{h}}$ is Lagrangian. □

Definition 18.12 (Period group)**.** We define

$$\Gamma(\mathbf{h}) := \left\{ \oint_\gamma i_{\mathbf{h}}^*\Theta \,\middle|\, \gamma : S^1 \to M_{\mathbf{h}} \right\} \subset \mathbb{R}$$

and call it the period group of $M_{\mathbf{h}}$.

This is a subgroup of $\mathbb{R}$ which is either discrete or countable dense in $\mathbb{R}$.

We pick a basis $\{e_1, \ldots, e_n\}$ of

$$H_1(M_{\mathbf{h}}, \mathbb{Z}) \cong H_1(\Phi^{-1}(B), \mathbb{Z}).$$

Then for each $j = 1, \ldots, n$, we set

$$\widetilde{I}_j(\mathbf{h}) = \frac{1}{2\pi} \oint_{e_j} i_{\mathbf{h}}^* \Theta$$

which induces a base change map $\widetilde{\mathbf{I}} = (\widetilde{I}_1, \ldots, \widetilde{I}_n)\ B \to \mathbb{R}^n$.

Definition 18.13. The functions $\mathbf{I} = (I_1, \ldots, I_n$ with $I_j = \widetilde{I}_j \circ \mathbf{F}$ are called the *action variables* of H.

Then we define the generating function

$$S(q, \mathbf{I}) = \int_{q_0}^{q} i_{\mathbf{h}(\mathbf{I})}^* \Theta$$

where the integral path from $q_0 = q_0(\mathbf{h})$ to q is chosen insider $M_{\mathbf{h}(I)}$. We then consider the canonical transformation

$$(q, p) \mapsto (\varphi, \mathbf{I})$$

given by the generating function $S = S(q, \mathbf{I})$ of type I such that

$$\mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}, \quad \varphi = \frac{\partial S}{\partial \mathbf{I}}.$$

It follows that the angle coordinates $(\varphi_1, \ldots, \varphi_i)$ defined this way indeed satisfies the quantization condition

$$\oint_{e_j} d\varphi_i = 2\pi \delta_{ij} \Longleftrightarrow \langle d\varphi_i, e_j \rangle = 2\pi \delta_{ij}.$$

By the same calculation as in the case of one degree of freedom, we derive

$$\begin{aligned} \sum_{i=1}^n dq_i \wedge dp_i &= \sum_{i=1}^n dq_i \wedge \sum_{j=1} \frac{\partial^2 S}{\partial q_i \partial I_j} \partial q_i \ dI_j \\ &= \sum_{j,i} \frac{\partial}{\partial q_i} \left(\frac{\partial S}{\partial I_j} \right) dq_i \wedge dI_j \\ &= \sum_j \left(\sum_{i=1}^n \frac{\partial \varphi_j}{\partial q_i} dq_i \right) \wedge dI_j = \sum_{j=1}^n d\varphi_j \wedge dI_j. \end{aligned}$$

This finishes construction of the required action-angle coordinates $(\varphi, \mathbf{I})$.

We note that all our constructions involve only "algebraic operations" of inverting functions and "quadrature" of calculating of the integrals of known functions. This completes the proof of the Liouville-Arnold-Avez theorem.

18.3. **Underlying geometry of the Hamilton-Jacobi method.** We now describe the geometry governing the construction of action-angle variables. There are two important ingredients used in the construction, one is the generating function of canonical transformation of the type I $S = S(q, Q)$ and the other is a usage of the system of stationary Hamilton-Jacobi equations

$$F_i \left(q, \frac{\partial S}{\partial q} \right) = h_i, \quad i = 1, \ldots, n.$$

Definition 18.14. Let $i : L \to T^*N$ be an exact Lagrangian immersion. We call a function $f : L \to \mathbb{R}$ satisfying $i^*\theta = df$ a *Liouville primitive* of i. When i is embedding, we just say it a Liouville primitive of L.

Let $\phi : T^*\mathbb{R}^n \to T^*\mathbb{R}^n$ be a canonical transformation and consider its graph

$$\Gamma_\phi \subset T^*\mathbb{R}^n \times T^*\mathbb{R}^n.$$

We equip the product with the symplectic form

$$\Omega = \pi_2^*\omega_0 - \pi_1^*\omega_0.$$

We denote by (q, p, Q, P) the natural coordinates pulled-back from $T^*\mathbb{R}^n$ via π_1, π_2. Then Γ_ϕ is a Lagrangian submanifold of Ω. Furthermore

$$\Omega = dQ \wedge dP - dq \wedge dp = d(pdq - PdQ).$$

Set $\Theta = pdq - PdQ$. Therefore we know that Γ_ϕ is an exact Lagrangian submanifold of Ω and so there is a Liouville primitive of Γ_ϕ.

We consider the projection $\pi_{q,Q} : T^*\mathbb{R}^n \to T^*\mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ and *assume that its restriction to* Γ_ϕ *is invertible.* Then this inverse, denoted by ι_ϕ, defines a Lagrangian embedding of $\mathbb{R}^n \times \mathbb{R}^n$ into $(T^*\mathbb{R}^n \times T^*\mathbb{R}^n, \Omega)$. Then we can write the Liouville primitive as a function of $S_\phi = S_\phi(q, Q)$ such that

$$\iota_\phi^*\Theta = dS_\phi.$$

This S_ϕ is precisely the generating function of ϕ of type I.

Next we turn our attention to the method of solving the system of Hamilton-Jacobi equations

$$H_i\left(q, \frac{\partial S}{\partial q}\right) = h_i, \quad i = 1, \ldots, n. \tag{18.3}$$

We have already established

$$M_{\mathbf{h}} = \Phi^{-1}(\mathbf{h}) = \bigcap_{i=1}^n F_i^{-1}(h)$$

is a Lagrangian torus. Then we vary $\mathbf{f}$ in B, each $M_{\mathbf{f}}$ is a Lagrangian torus with action coordinates $(\varphi_1, \ldots, \varphi_n)$ arising from $M_{\mathbf{h}}$. Identifying $\Phi^{-1}(B)$ with a neighborhood of the zero section of $T^*M_{\mathbf{h}}$, we are given a foliation of $\Phi^{-1}(B)$ by exact Lagrangian tori. A solution to (18.3) is nothing but construction of a smooth family of Liouville primitives S^h so that the image of its differential $dS^{\mathbf{h}}$ parameterizes $M_{\mathbf{h}}$ which is equivalent to asking S^h to satisfy the system (18.3).

An additional point of Liouville's theorem and action-angle variables is that construction of solutions of Hamilton's equation of $H = F_1$ just involves "inverting maps" and "quadratures", *when it admits maximal number of explicitly given first integrals* $F_1, \cdots, F_n$.

19. Lie groups and Lie algebras

From now on, we will carry out a systematic study of Hamiltonian systems with continuous symmetry and a method of symplectic reductions and moment maps.

A topological group G is a group equipped with a topology with respect to which the group operations are continuous. More precisely, the product

$$G \times G \to G; \quad (g, h) \mapsto gh$$

and the inverse operation

$$G \to G; \quad g \mapsto g^{-1}$$

are continuous maps with respect to the given topology.

A *Lie group* G is a smooth manifold that carries a group structure whose group operations are smooth. More precisely, the product and the inverse operation are smooth with respect to the smooth structure of G.

Remark 19.1. (1) (Yamabe) Any arc-wise connected locally compactly topological group can be given a compatible smooth structure and so a Lie group.

(2) An important example is the general linear group $GL(n,\mathbb{R}) \subset M^{n\times n}(\mathbb{R}) \cong \mathbb{R}^{n^2}$ under the matrix product and inverse operations which are indeed polynomial functions and rational functions respectively.

There is a canonically associated Lie algebra $\mathfrak{g} := \mathrm{Lie}(G)$ to each Lie group G. To construct $\mathrm{Lie}(G)$, we first note that a multiplication by $g \in G$ defines two self-diffeomorphisms, called the left multiplication L_g and the right multiplication R_g defined by

$$L_g(h) = gh, \; R_g(h) = hg.$$

Definition 19.2. We call a vector field X on G, not necessarily continuous, *left-invariant* (resp. *right-invariant*) if $(L_g)_* X = X$ (resp. if $(R_g)_* X = X$.

Lemma 19.3. *Every left invariant vector field X on a Lie group G is smooth. The same holds for the right invariant vector fields.*

Proof. Enough to prove that X is smooth in a neighborhood V of the identity $e \in G$ which would imply $(L_g)_* X$ is smooth on $g(V)$ of g.

Then prove that $X[x_i]$ are smooth for any coordinate functions $(x_1, \ldots, x_n)$ at $e \in G$. □

Corollary 19.4. *A Lie group G always has trivial tangent bundle (and hence is also orientable).*

Proof. Let $V = T_e G$ and fix a basis $\{e_1, \ldots, e_n\}$ thereof. $\Phi : G \times \mathbb{R} \to TG$ defined by $\Phi(g, (s_1, \ldots, s_n)) = TL_g(\sum_{i=1} s_i e_i)$, which is clearly smooth.

Denote by X_i the left-invariant vector field defined by $X_i(g) = TL_g(e_i)$. Then The inverse $\Psi : TG \to G \times \mathbb{R}$ of Φ is given as follows: Any vector v in $T_g G$ can be written as a linear combination

$$v = \sum_{i=1} c_i X_i(g).$$

Then we define $\Psi(v) = (\pi(v), (c_1, \ldots, c_n))$. □

Lemma 19.5. *The Lie bracket $[X,Y]$ of two left-invariant vector fields X, Y are left-invariant.*

Proof. This is immediately follows from the property of Lie bracket of vector fields

$$\phi_*[X,Y] = [\phi_* X, \phi_* Y]$$

for any diffeomorphism ϕ. □

This lemma shows that the set of left-invariant vector fields on a Lie group G becomes a Lie algebra under the restriction of Lie bracket of vector fields.

Definition 19.6. The Lie algebra, denoted by $\mathfrak{g}$, is the tangent space T_eG equipped with the bracket defined as follows: Let a, $b \in \mathfrak{g}$ and X_a, X_b the left-invaraint vector field of G such that $X_a(w) = a$, $X_b(e) = b$. Then we defind

$$[a, b] := [X_a, X_b](e).$$

Example 19.7. Consider the general linear group $GL(n, \mathbb{R}) \subset M^{n\times n}(\mathbb{R})$. We have a natural Lie algebra structure on $M^{n\times n}(\mathbb{R})$ given by the commutator bracket

$$[A, B] = AB - BA, \quad A, B \in M^{n\times n}(\mathbb{R}).$$

Consider the natural group structure on $GL(n, \mathbb{R})$ induced by the matrix multiplication. Then its induced Lie algebra, denoted by $\mathfrak{gl}(n, \mathbb{R})$, is isomorphic to this Lie algebra $(M^{n\times n}(\mathbb{R}), [\cdot, \cdot])$.

Definition 19.8. Let $\mathfrak{g}$ be a Lie algebra. A subspace $\mathfrak{h} \subset \mathfrak{g}$ is called a Lie subalgebra, if $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$.

In this sense, the set of left-invariant (resp. right-invariant) vector fields on G is a Lie subalgebra of the set of vector fields on G.

Theorem 19.9. *Let G be a Lie group, and $\mathfrak{h} \subset \mathfrak{g}$ a subalgebra. Then there exists a unique connected Lie subgroup H of G whose Lie algebra is $\mathfrak{g}$.*

Proof. We have a natural distribution $\mathcal{H} \subset TG$ defined by

$$\mathcal{H}_g = \{TL_g(v) \mid v \in \mathfrak{h}\}.$$

This distribution is involutive since $\mathfrak{h}$ is a Lie subalgebra and so integrable. Let H be the maximal integral manifold of $\mathcal{H}$ that passes through $e \in G$. We claim that H is a Lie subgroup. □

Example 19.10. Consider $GL(n, \mathbb{R})$ and its Lie algebra $\mathfrak{gl}(n, \mathbb{R}) = M^{n\times n}(\mathbb{R})$. Consider the set of skew-symmetric matrices, denoted by $\mathfrak{o}(n)$. Then $\mathfrak{o}(n)$ is a Lie subalgebra of $\mathfrak{gl}(n, \mathbb{R})$ whose associated (connected) Lie subgroup is nothing but $\mathfrak{S}O(n)$.

Theorem 19.11 (Ado)**.** *Every Lie algebra is isomorphic to a subalgebra of $\mathfrak{gl}(N, \mathbb{R})$ for some N.*

Therefore any Lie algebra is the Lie algebra of a connected Lie group.

Definition 19.12. A Lie group homomorphism $\phi : G \to H$ is a group homomorphism that is smooth.

(In fact, any continuous group homomorphism between two Lie groups is automatically smooth.)

For any Lie group homomorphism $\phi : G \to H$, the derivative map

$$d_e\phi : T_eG \to T_eH$$

is a Lie algebra homomorphism. The converse also holds locally in the following sense.

Theorem 19.13. *Let G and H be Lie groups, and let $\ell : \mathfrak{g} \to \mathfrak{h}$ be a Lie algebra homomorphism. There there is a neighborhood U of $e \in G$ and a smooth map $\phi : U \to H$ such that*

$$\phi(gh) = \phi(g)\phi(h) \quad \textit{when } g, h, gh \in U$$

and such that for every $\xi \in \mathfrak{g}$, we have

$$d_e\phi(\xi) = \ell(\xi).$$

Moreover, if there are two smooth homomorphisms ϕ, $\psi : G \to H$ with $d_e\phi = d_e\psi$, and G is connected, then $\phi = \psi$.

Proof. Let $\mathfrak{k} \subset \mathfrak{g} \times \mathfrak{h}$ be the set of all $(\xi, \ell(\xi))$ for $\xi \in \mathfrak{g}$. Since ℓ is a Lie algebra homomorphism, $\mathfrak{k}$ is a subalgebra of $\mathfrak{g} \times \mathfrak{h}$. By Theorem 19.9, we have a unique connected Lie subgroup $K \subset G \times H$. Denote by

$$\pi_1 : G \times H \to G, \quad \pi_2 : G \times H \to H$$

the natural projections, and by $\omega = (\pi_1)|_K$ the restriction of π_1 to $K \subset G \times H$. It follows that ω is a homomorphism.

By construction, we have $d\omega(e)(\xi, \ell(\xi)) = \xi$ which is an isomorphism. By the inverse function theorem, we can find a neighborhood V of $(e, e) \in K$ such that ω takes V diffeomorphically onto an open neighborhood U of $e \in G$. Then the composition $\phi := \pi_2 \circ \omega^{-1} : U \to H$ is the required map.

Finally for given ϕ, $\psi : G \to H$, we define the one-one map $\theta : G \to G \times H$ by

$$\theta(g, h) = (g, \psi(g)).$$

The image $K' \subset G \times H$ is a Lie subgroup of $G \times H$ and $d_e\theta(\xi) = (\xi, \ell(\xi))$ and hence $T_eK' = T_eK$. By the uniqueness, we have $K = K'$ which implies $\psi(g) = \phi(g)$ for all $g \in G$. □

Corollary 19.14. *If two Lie groups G and H have isomorphic Lie algebras, then they are locally isomorphic.*

Exercise 19.15. Prove that two simply connected Lie groups whose Lie algebras are isomorphic are isomorphic, and that all connected Lie groups with a given Lie algebras are covered by the same simply connected Lie group.

Example 19.16. The following example shows that $SO(3)$ and $SU(2)$ have isomorphic Lie algebras but are not globally isomorphic as a Lie group:

(1) Let us consider the case $G = SO(3)$. Its associated Lie algebra is given by

$$\text{so}(3) = \{L \in M^{3\times 3}(\mathbb{R}) \mid L^t = -L\} \cong \mathbb{R}^3.$$

With respect to the basis

$$E = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, F = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

we have the simple structure equation

$$[E, F] = G, \; [F, G] = E, \; [G, E] = F$$

which is isomorphic to the cross product $\vec{i}$, $\vec{j}$, $\vec{k}$ for the standard basis of $\mathbb{R}^3$.

(2) Consider the $SU(2)$

$$SU(2) = \{U \in M^{2\times 2}(\mathbb{C}) \mid U^*U = UU^* = I, \det U = 1\}$$

where $U^* = \overline{U}^t$ is the Hermitian conjugate. It can be written as

$$U = \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix}$$

with $|a|^2 + |b|^2 = 1$. Its Lie algebra $su(2)$ is given by the traceless skew-Hermitian matrix

$$\begin{pmatrix} l\sqrt{-1} & m \\ -\overline{m}, & -l\sqrt{-1} \end{pmatrix}$$

with ℓ real. Therefore if we consider the basis

$$E = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},\ F = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ G = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

then they satisfy the same commutation relation

$$[E, F] = G,\ [F, G] = E,\ [G, E] = F$$

as that of $so(3) \cong \mathbb{R}^3$.

Corollary 19.17. *A connected Lie group G with an abelian Lie algebra is itself abelian.*

Proof. We recall that $\mathbb{R}^n$ is a abelian Lie group under the coordinatewise addition whose Lie algebra is also $\mathbb{R}^n$ with trivial Lie bracket. Therefore the Lie group G is locally abelian, i.e., that there exists a neighborhood U of o such that $gh = hg$ for all g, h, $gh \in U$. It follows that G is abelian since any neighborhood of G generates G. □

Exercise 19.18. Let G be a topological group and $H \subset G$ a subgroup. Prove

(1) If H is open, then so is every coset gH.
(2) If H is open, the H is closed.

Exercise 19.19. Let G be a connected topological group, and U a neighborhood of $e \in G$. Let U^n denote all products $g_1, \cdots, g_n$ for $g_i \in U$.

(1) Show that U^{n+1} is a neighborhood of U^n.
(2) Conclude $\cup_n U^n = G$.

An implication of this exercise is that the connected topological group is generated by any neighborhood of the identity.

Let us specialize to the case $\mathbb{R}$ and a Lie group G.

Proposition 19.20. *For every $\xi \in T_eG = \mathfrak{g}$, there is a unique homomorphism $\phi : \mathbb{R} \to G$ such that*

$$\left.\frac{d\phi}{dt}\right|_{t=0} = \xi.$$

Proof. Note that $\mathfrak{h} := \{t\xi\} \subset \mathfrak{g}$ is an abelian subalgebra and the map $t \mapsto t\xi$ defines a Lie algebra homomorphism onto $\mathfrak{h}$. Therefore there is an open neighborhood $(-\epsilon, \epsilon)$ of $0 \in \mathbb{R}$ and a map $\phi_\epsilon : (-\epsilon, \epsilon) \to G$ such that for all s, $t \in (-\epsilon, \epsilon)$ with $|s + t|, \epsilon$, $\phi(s + t) = \phi(s) \times \phi(t)$ and

$$\left.\frac{d\phi}{dt}\right|_{t=0} = \xi.$$

To extend ϕ to $\mathbb{R}$, we write t with $|t| \geq \epsilon$ uniquely as

$$t = k(\epsilon/2) + r, \quad k \text{ an integer, } |r| < \epsilon/2$$

and define

$$\phi(t) = \begin{cases} \phi(\epsilon/2) \cdots \phi(\epsilon/2) \cdot \phi(r) & (\phi(\epsilon/2) \text{ appears } k \text{ times})\, k \geq 0, \\ \phi(-\epsilon/2) \cdots \phi(-\epsilon/2) \cdot \phi(r) & (\phi(\epsilon/2) \text{ appears } k \text{ times})\, k < 0. \end{cases}$$

□

A homomorphism $\phi : \mathbb{R} \to G$ above is called a *one-parameter subgroup* of G.
We define

$$\exp \xi = \phi(1)$$

which defines a map $\exp : \mathfrak{g} \to G$ which we call the exponential map of G. Obviously we have $\phi(t) = \exp(t\xi)$.

Example 19.21. Consider the case $G = GL(n, \mathbb{R})$. In this case, we have the map

$$\exp : M^{n\times n}(R) \to GL(n, \mathbb{R})$$

defined by the convergent power series

$$\exp(A) = I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \cdots$$

which indeed satisfies $\exp((t+s)A) = \exp(tA)\exp(sA)$ and so coincides with the above defined exponential map.

In general $\exp(A+B) \neq \exp A \exp B$ unless $AB = BA$.

Proposition 19.22. *The map* $\exp : \mathfrak{g} \to G$ *is a local diffeomorphism. If* $\psi : G \to H$ *is a homomorphism, then we have*

$$\exp \circ d_e\psi = \psi \circ \exp .$$

Corollary 19.23.
(1) *Every one-one smooth homomorphism* $\psi : G \to H$ *is an immersion.*
(2) *Every continuous homomorphism* $\phi : \mathbb{R} \to \Gamma$ *is smooth.*
(3) *Every continuous homomorphicm* $\phi : G \to H$ *is smooth.*

20. Group actions and adjoint representations

Let G be a Lie group and M be a manifold. A (Lie) group left-action is a smooth map

$$\Phi : G \times M \to M$$

that satisfies

$$\Phi(g, \Phi(h, x)) = \Phi(gh, x), \quad \Phi(e, x) = x.$$

A *right-action* is one that satisfies

$$\Phi(g, \Phi(h, m)) = \Phi(hg, m).$$

We denote the left action by $g \cdot x$ or gx and the right action by xg. We call a manifold M with action of G a *G-manifold*.

From now on, we will focus on the left action unless otherwise said. We call the set $G \cdot x := \{gx \in M \mid g \in G\}$ the G-orbit of $x \in M$.

For each $g \in G$, the map $\Phi_g := \Phi(g, \cdot)$ defined a diffeomorphism whose inverse is given by Φ_g^{-1}. In this regard, each group action induces a group homomorphism

$$G \to \mathrm{Diff}(M); \quad g \mapsto \Phi_g.$$

Definition 20.1. A linear action of G on a vector space V, i.e., a group homomorphism of G to $GL(V)$ for a vector space V is called a *representation* of G. In that case, we say the vector space V a G-module. We say that the representation is *irreducible* if the G-module $\mathbb{R}^N$ has no invariant submodule other than $\{0\}$ and V itself.

By considering the natural curve $t \mapsto \exp(t\xi)x$ through x, each ξ defines a vector field ξ_M on M whose value is given by

$$\xi_M(x) = \frac{d}{dt}\Big|_{t=0} \exp(t\xi)x \tag{20.1}$$

at each $x \in M$. For any of G on (M, ω), the map $\xi \to \xi_M$ induces an infinitesimal action of $\mathfrak{g}$ on TM in that a Lie algebra homomorphism $\mathfrak{g} \to \mathcal{X}(M)$ whose proof will be given shortly.

Recall that G acts on G in two different ways and that the left invariant vector field is defined by

$$\widetilde{\xi}(g) = \frac{d}{dt}\Big|_{t=0} g\exp(t\xi) = dL_g(\xi).$$

Note that any conjugation action of $g \in G$ on G defined by

$$g \mapsto ghg^{-1} = L_g R_{g^{-1}}(h)$$

fixes the identity and so its derivative defines a linear map on $\mathfrak{g}$. We denote

$$\mathrm{Ad}_g(\xi) = d_e(L_g R_{g^{-1}})(\xi) = \frac{d}{dt}\Big|_{t=0} g\exp(t\xi)g^{-1}$$

Definition 20.2. The adjoint representation of G is the homomorphism

$$\mathrm{Ad} : G \to GL(\mathfrak{g}); \quad g \mapsto \mathrm{Ad}_g .$$

Proposition 20.3. $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$ *is a Lie algebra homomorphism, i.e.,*

$$\mathrm{Ad}_g([\xi, \eta]) = [\mathrm{Ad}_g \xi, \mathrm{Ad}_g \eta].$$

Proof. By definition,

$$\begin{aligned}
\mathrm{Ad}_g([\xi,\eta]) &= dL_g dR_{g^{-1}}(\widetilde{[\xi,\eta]}(e)) = dL_g dR_{g^{-1}}(dL_g)^{-1}(\widetilde{[\xi,\eta]}(g)) \\
&= dR_{g^{-1}}(\widetilde{[\xi,\eta]}(g)) = R_g^*\widetilde{[\xi,\eta]}(e) \\
&= [R_g^*\widetilde{\xi}, R_g^*\widetilde{\eta}](e) = [\widetilde{\mathrm{Ad}_g \xi}, \widetilde{\mathrm{Ad}_g \eta}](e) \\
&= [\mathrm{Ad}_g \widetilde{\xi, \mathrm{Ad}_g} \eta](e) = [\mathrm{Ad}_g \xi, \mathrm{Ad}_g \eta].
\end{aligned}$$

□

Example 20.4. In fact, using the isomorphism $SU(2) \cong S^3 =$ unit quarternions $= Spin(3)$, we can identify the Lie algebra $su(2)$ with the imaginary quarternion

$$\mathrm{span}_{\mathbb{R}}\{i,\, j,\, k\} \cong T_{(1,0,0,0)}S^3.$$

Then the conjugation by q

$$p \mapsto qpq^{-1} = qp\overline{q}$$

for the imaginary quarternion p induces an orientation preserving isometry on $\mathbb{R}^3$ and hence defines a surjective homomorphism

$$\mathrm{Ad} : SU(2) \to SO(\mathbb{R}^3) = SO(3)$$

which is a covering of degree 2.

We also define a linear map $\mathrm{ad}_\eta : \mathfrak{g} \to \mathfrak{g}$ by

$$\mathrm{ad}_\eta(\xi) := [\eta, \xi]$$

for each $\eta \in \mathfrak{g}$. Then we have

Proposition 20.5. *Consider the linear maps* $\mathrm{Ad}_{\exp(t\eta)}: \mathfrak{g} \to \mathfrak{g}$. *Then*

$$\frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp(t\eta)} = \mathrm{ad}_\eta\,.$$

Proof. Recall we denoted by $\widetilde{\eta}$ the left invariant vector field on G with $\widetilde{\eta}(e) = \eta$. Consider the map $\psi_t : G \to G$ defined by

$$\psi_t(g) = g\exp(t\eta) = R_{\exp(\eta)}(g).$$

Then by definition the one-parameter subgroup associated to the left-invariant vector field $\widetilde{\eta}$ is given by

$$\phi^t_{\widetilde{\eta}}(g) = g\exp(t\eta) = \psi_t(g).$$

Therefore we compute

$$\begin{aligned}
[\eta,\xi] &= [\widetilde{\eta},\widetilde{\xi}](e) = \mathcal{L}_{\widetilde{\eta}}(\widetilde{\xi})(e) = \frac{d}{dt}\Big|_{t=0}(\psi_t)^*\widetilde{\xi} \\
&= \frac{d}{dt}\Big|_{t=0} d\psi_{-t}\widetilde{\xi}(\psi_t(e)) \\
&= \frac{d}{dt}\Big|_{t=0} dR_{\exp(-t\eta)}\widetilde{\xi}(\exp(t\eta)) \\
&= \frac{d}{dt}\Big|_{t=0} dR_{\exp(-t\eta)}dL_{\exp(t\eta)}(\xi) = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp(t\eta)}(\xi).
\end{aligned}$$

In other words, we have proved $\mathrm{ad}_\eta = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp(t\eta)}$. □

Proposition 20.6. *For all* ξ, $\eta \in \mathfrak{g}$,

$$\widetilde{[\eta,\xi]} = [\widetilde{\eta},\widetilde{\xi}].$$

Equivalently, $\xi \mapsto -\widetilde{\xi}$ *is a Lie algebra homomorphism.*

Proof. We recall $\widetilde{\xi}(g) = dL_g(\xi)$ and

$$\begin{aligned}
[\widetilde{\eta},\widetilde{\xi}](g) &= \frac{d}{dt}\Big|_{t=0}(R_{\exp t\eta})^*\widetilde{\xi}(g) = \frac{d}{dt}\Big|_{t=0} dR_{\exp -t\eta}(\widetilde{\xi}(g\exp t\eta)) \\
&= \frac{d}{dt}\Big|_{t=0} dR_{\exp -t\eta}dL_{g\exp t\eta}\xi = \frac{d}{dt}\Big|_{t=0} dR_{\exp -t\eta}dL_g dL_{\exp t\eta}\xi \\
&= \frac{d}{dt}\Big|_{t=0} dL_g dR_{\exp -t\eta}dL_{\exp t\eta}\xi = dL_g\left(\frac{d}{dt}\Big|_{t=0}\mathrm{Ad}_{\exp t\eta}(\xi)\right) \\
&= dL_g([\eta,\xi]) = \widetilde{[\eta,\xi]}(g)
\end{aligned}$$

which finishes the proof. □

21. Lie Poisson space $\mathfrak{g}^*$ and the coadjoint action of G

We start with a general discussion on the Lie group actions on smooth manifolds. Consider the action of G on M for an arbitrary smooth manifold.

Proposition 21.1. *The vector field* ξ_M *satisfies*

$$g_*\xi_M = (\mathrm{Ad}_g\,\xi)_M$$

for all $g \in G$.

Proof. By definition, we have

$$\begin{aligned} g_*\xi_M(x) &= d\Phi_g(\xi_M(\Phi_g^{-1}(x))) = \frac{d}{dt}\Big|_{t=0} \Phi_g(\exp(t\xi))\Phi_g^{-1}(x) \\ &= \frac{d}{dt}\Big|_{t=0} \left(\Phi_g(\exp(t\xi))\Phi_g^{-1}\right)(x) = (\mathrm{Ad}_g\,\xi)_M(x). \end{aligned}$$

□

This gives rise to

Corollary 21.2. *The assignment $\xi \mapsto \xi_M$ is a Lie algebra homomorphism, i.e.,*

$$-[\xi_M, \eta_M] = [\xi, \eta]_M.$$

Proof. We just differentiate $(\exp t\eta)_*\xi_M = (\mathrm{Ad}_{\exp t\eta}\,\xi)_M$ in t at $t = 0$. □

Now we specialize to the case of the Lie-Poisson space $(\mathfrak{g}^*, \{\cdot,\cdot\}_{LP})$. By taking the dual of the adjoint representation Ad of G on $\mathfrak{g}$

$$\mathrm{Ad}^* : G \to GL(\mathfrak{g}^*); \quad g \mapsto \mathrm{Ad}^*_{g^{-1}},$$

we obtain the coadjoint (left) action of G on defined by

$$\langle \mathrm{Ad}^*_{g^{-1}}(\mu), \xi\rangle := \langle \mu, \mathrm{Ad}_{g^{-1}}(\xi)\rangle.$$

We call Ad^* the *coadjoint representation* of G.

Definition 21.3 (Coadjoint orbit)**.** Let $\mu \in \mathfrak{g}^*$ and consider its orbit

$$\mathcal{O}_\mu = \{\mathrm{Ad}^*_{g^{-1}}(\mu) \mid g \in G\}$$

under the coadjoint action Ad^* of G. We call the orbit the *coadjoint orbit* of μ.

Definition 21.4. Let G be a Lie group. Let P be a Poisson manifold. We say the action is Poisson if Φ_g is a Poisson map for all $g \in G$, i.e., if

$$\{f, h\} \circ \Phi_g = \{f \circ \Phi_g, h \circ \Phi_g\}, \quad f, h \in C^\infty(P)$$

for all $g \in G$. Equivalently, if $(\Phi_g)_*\Pi = \Pi$.

Lemma 21.5. *Let $(\mathfrak{g}^*, \{\cdot,\cdot\})$ be a Lie-Poisson space.*

1. *Each $\mathrm{Ad}^*_{g^{-1}}$ is a Poisson map and $\mathrm{Ad}^* : G \to \mathrm{Diff}_\Pi(\mathfrak{g}^*)$ is a homomorphism into the group of Poisson maps of $\mathfrak{g}^*$.*
2. *For any $\xi \in \mathfrak{g}$, its associated vector field $\widetilde{\xi} = \{\cdot, \xi\}$ on $\mathfrak{g}^*$ regarded ξ as a linear function on $\mathfrak{g}^*$ is tangent to the coadjoint leaves. In particular, the Lie Poisson bracket induces a canonical Poisson bracket on each orbit $\mathcal{O}_\mu$.*

We recall that any tangent vector v of $\mathcal{O}_\mu$ at $\nu \in \mathcal{O}_\mu$ has the form

$$v = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}^*_{\exp(t\xi)^{-1}}(\nu) =: \xi_{\mathfrak{g}^*}(\nu)$$

for some $\xi \in \mathfrak{g}$.

Lemma 21.6. *We have*

$$\xi_{\mathfrak{g}^*}(\nu) = -\mathrm{ad}^*_\xi \nu = -\nu \circ \mathrm{ad}_\xi.$$

Proof. We compute

$$\xi_{\mathfrak{g}^*}(\nu) = \frac{d}{dt}\Big|_{t=0}(\nu \circ \mathrm{Ad}_{\exp(t\xi)^{-1}}) = \nu \circ \left(\frac{d}{dt}\Big|_{t=0}\mathrm{Ad}_{\exp(t\xi)^{-1}}\right) = -\nu \circ \mathrm{ad}_\xi \quad (21.1)$$

as an element $T_\nu \mathcal{O}_\mu \subset T_\nu \mathfrak{g}^* \cong \mathfrak{g}^*$. This is also the same as $-\mathrm{ad}_\xi^* \nu$. This finishes the proof. □

Theorem 21.7. *The induced Poisson bracket on each coadjoint orbit $\mathcal{O}_\mu$ is non-degenerate, i.e., carries a natural G-invariant symplectic form ω_μ defined by*

$$\omega_\mu(\xi_{\mathfrak{g}^*}(\nu), \eta_{\mathfrak{g}^*}(\nu)) = \langle \nu, [\xi, \eta]\rangle$$

at any $\nu \in \mathcal{O}_\mu$.

Proof. Let $\mu \in \mathfrak{g}^*$. Regarding ξ, $\eta \in \mathfrak{g}$ as linear functions on $\mathfrak{g}^*$, their Poisson bracket $\{\xi, \eta\}$ is defined by

$$\{\xi, \eta\}(\nu) = \langle \nu, [\xi, \eta]\rangle$$

for $\nu \in \mathfrak{g}^*$. If $\nu = \mathrm{Ad}^*_{g^{-1}}(\mu)$, we have

$$\{\xi, \eta\}(\mathrm{Ad}^*_{g^{-1}}(\mu)) = \{\xi \circ \mathrm{Ad}_{g^{-1}}, \eta \circ \mathrm{Ad}_{g^{-1}}\}(\mu). \quad (21.2)$$

Now we define a two-form ω_μ on $\mathcal{O}_\mu$ by

$$\omega_\mu(\xi_{\mathfrak{g}^*}(\nu), \eta_{\mathfrak{g}^*}(\nu)) = \langle \nu, [\xi, \eta]\rangle.$$

(21.2) proves the Ad^*-invariance of ω_μ on $\mathcal{O}_\mu$. It remains to show closedness and nondegeneracy of ω_μ.

For the nondegeneracy of ω_μ, suppose $\omega_\mu(v_1, v_2) = 0$ with v_1, $v_2 \in T_\nu\mathcal{O}_\mu$ for all v_2. We know $v_i = -\,\mathrm{ad}^*_{\xi_i} \nu$ for some $\xi_i \in \mathfrak{g}$, $i = 1, 2$. Then it is equivalent to saying $\langle \nu, [\xi_1, \xi_2]\rangle = 0$ for all ξ_2. This then is equivalent to saying $\mathrm{ad}^*_{\xi_1}\nu = 0$. This proves $v_1 = 0$ by (21.1) which finishes the proof of nondegeneracy.

Finally the closedness follows from the Jacobi identity of the Poisson bracket on $\mathcal{O}_\mu$.

□

Corollary 21.8. *All coadjoint orbits are even dimensional and $\mathfrak{g}$ is the union of $\mathcal{O}_\mu$.*

Definition 21.9 (Kirillov-Kostant-Souriou symplectic form)**.** We call the above defined ω_μ the KKS symplectic form on the coadjoint orbit $\mathcal{O}_\mu$. We call each coadjoint orbit a symplectic leaf of the Lie-Poisson space $\mathfrak{g}^*$.

In this regard, $\mathfrak{g}^*$ carries a natural singular foliations whose leaves are all even dimensional of varying dimension.

It is easier to visualize the adjoint orbit than the coadjoint orbit, if the Lie algebra $\mathfrak{g}$ admits an invariant inner product. One distinguished class of Lie algebras is *semi-simple* Lie algebras. We recall the linearized adjoint representation

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{gl}(\mathfrak{g});\ \mathrm{ad}(\eta) = \mathrm{ad}_\eta\,.$$

Definition 21.10 (Killing form)**.** Let $\mathfrak{g}$ be a Lie algebra. The bilinear form

$$B(\xi, \eta) := \mathbf{Tr}(\mathrm{ad}(\xi)\,\mathrm{ad}(\eta))$$

is called the Killing form of $\mathfrak{g}$. $\mathfrak{g}$ is called *semi-simple* if the Killing form B is nondegenerate.

Therefore if $\mathfrak{g}$ is semi-simple, we can pull-back all discussion on $\mathfrak{g}^*$ to $\mathfrak{g}$ via the isomorphism

$$\xi \to B(\xi, \cdot);\ \mathfrak{g} \to \mathfrak{g}^*.$$

For example, the adjoint orbits of $\mathfrak{g}$ carry a symplectic form pulled-back from the KKS form on the coadjoint orbits.

Example 21.11. Let us consider the case $G = SO(3)$. Its associated Lie algebra is given by

$$\mathrm{so}(3) \cong \mathbb{R}^3 = \{L \in M^{3\times 3}(\mathbb{R}) \mid L^t = -L\}.$$

With respect to the basis

$$E = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, F = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

we have the simple structure equation

$$[E, F] = G,\ [F, G] = E,\ [G, E] = F$$

which is isomorphic to the cross product $\vec{i}$, $\vec{j}$, $\vec{k}$ for the standard basis of $\mathbb{R}^3$.

We have a natural bilinear from on $M^{3\times 3}(\mathbb{R})$ defined by

$$\langle L_1, L_2 \rangle = \mathbf{Tr}(L_1 L_2).$$

Lemma 21.12. *This restricts to a nondegenerate pairing on* so(3)*, which is the same as the Killing form B.*

Proof. A general element A of so(3) is given by

$$L = \begin{pmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{pmatrix}.$$

A direct calculation shows

$$\langle L_1, L_2 \rangle = -2 v_1 \cdot v_2$$

where $v_i = (a_i, b_i, c_i)$ and $v_1 \cdot v_2$ is the dot product on $\mathbb{R}^3$.

On the other hand, a straightforward computation proves

$$B(L_1, L_2) = -2 v_1 \cdot v_2 = \mathbf{Tr}(L_1 L_2).$$

□

Note that the above bilinear form is invariant under the adjoint action of SO(3) on so(3). The inner product induces identification of $\mathfrak{g}^*$ with $\mathfrak{g}$, and the coadjoint action on $\mathfrak{g}^*$ with the adjoint action on $\mathfrak{g}$ and with $\mathbb{R}^3$ with the standard action of SO(3) on $\mathbb{R}^3$.

Therefore each coadjoint orbit $\mathcal{O}_\mu$ equipped with its KKS form is symplectomorphic to a two-sphere $S^2(r)$ of some radius r equipped with the standard area form.

Exercise 21.13. Make precise the discussion in the last paragraph of the above example by quantifying all the relationships between μ and r and others.

22. Symplectic action and the moment map

We start with the definition of symplectic action of a Lie group G on (M, ω). Let $G \times M \to M$ be an action of G on M. Denote by Φ_g the diffeomorphism associated to the group element $g \in M$.

Definition 22.1. We say a group G acts on M symplectically if it preserves the given symplectic form, i.e., $\Phi_g^* \omega = \omega$ for all $g \in G$.

Recall that for the simplicity of notation, we just denote $\Phi_g(x) = gx$.

Proposition 22.2. *G acts on M symplectically if and only if $d(\xi_M \rfloor \omega) = 0$ for all ξ.*

Therefore the assignment $\xi \mapsto \xi_M$ actually induces a Lie algebra homomorphism

$$\mathfrak{g} \to \mathrm{symp}(M, \omega).$$

Definition 22.3. We say an action of G on M is pre-Hamiltonian if $\xi_M \rfloor \Omega$ is exact for all ξ.

Recall that the choice of Hamiltonian for the given Hamiltonian vector field is unique up to addition by constants.

Suppose M is compact connected without boundary. Then we have a commutative diagram of exact sequences

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{R} & \longrightarrow & C^\infty(M) & \xrightarrow{d} & Z^1(M) & \longrightarrow & H^1(M, \mathbb{R}) & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow{\scriptstyle \widetilde{\omega}^{-1}} & & \downarrow & & \\ 0 & \longrightarrow & \mathbb{R} & \longrightarrow & C^\infty(M) & \xrightarrow[X_{\{\cdot\}}]{} & \mathrm{symp}(M, \omega) & \longrightarrow & H^1(M, \mathbb{R}) & \longrightarrow & 0 \end{array}$$

where the second map in the lower exact sequence is $H \to X_H$ and the third map is $X \to [X \rfloor \omega]$. The image of the first map is precisely the set of Hamiltonian vector fields $\mathrm{ham}(M, \omega)$ and we have

$$\frac{\mathrm{symp}(M, \omega)}{\mathrm{ham}(M, \omega)} \cong H^1(M, \omega).$$

For given pre-Hamiltonian action of G on M, for each $\xi \in \mathfrak{g}$, there is associated to an element $J_\xi \in C^\infty(M)$ satisfying

$$dJ_\xi = \xi_M \rfloor \omega.$$

This is equivalent to saying that $X_{J_\xi} = \xi_M$ for all ξ, $\eta \in \mathfrak{g}$. But the choice of J_ξ for ξ_M is unique only up to addition by constant.

We collect them all to introduce the notion of moment map.

Definition 22.4. Suppose that the action of G on (M, ω) is pre-Hamiltonian. A *moment map* (or *moment mapping*) is a map $J : M \to \mathfrak{g}^*$ that satisfies

$$d\langle J, \xi \rangle = \xi_M \rfloor \omega \tag{22.1}$$

for all $\xi \in \mathfrak{g}$ where $\langle J, \xi \rangle = J_\xi$ with J regarded as a vector-valued function on M valued in $\mathfrak{g}^*$.

The moment map provides a canonical way of producing first integrals of any G-invariant Hamiltonian H in the following sense.

Theorem 22.5. *Let Φ be a symplectic action of G on (M,ω) with a moment map J. Suppose $H : M \to \mathbb{R}$ is invariant under the action, i.e., $H(x) = H(\Phi_g(x))$ for all (g,x). Then J is preserved under the flow, i.e., $J \circ \phi_H^t = J$. In particular, J_ξ are first integrals of H for all $\xi \in \mathfrak{g}$.*

We next ask the questions whether for a given pre-Hamiltonian action of G such an assignment $\xi \mapsto J_\xi$ actually can be made so that the following

- (Equivariance) $J \circ \Phi_g = \operatorname{Ad}^*_{g^{-1}} J$ for $g \in G$,
- (Lie algebra homomorphism) $J_{[\xi,\eta]} = \{J_\xi, J_\eta\}$ for $\xi\,\eta \in \mathfrak{g}$.

We will see that the equivariance also implies the second property and so focus on the question on the equivariance property. In general there is some cohomological obstruction to achieve such an equivariance property.

Definition 22.6. A moment map J is called Ad^*-*equivariant* provided the associated cocycle σ becomes 0, i.e, provided

$$J \circ \Phi_g = \operatorname{Ad}^*_{g^{-1}} J$$

for every $g \in G$.

Definition 22.7 (Hamiltonian G-space)**.** We say (M,ω) is a Hamiltonian G-manifold if there is given a pre-Hamiltonian G-action on M with Ad^*-equivariant moment map J. In this case, we say a pre-Hamiltonian action of G on (M,ω) is *Hamiltonian* and call the quadruple (M,ω,Φ,J) a *Hamiltonian G-space* or a *Hamiltonian G-manifold*.

Now we describe the obstruction to the Ad^*-equivariance of the moment map. For this purpose, we recall the following standard definition of group cohomology. (See [Wb94] for example.)

Definition 22.8 (Group cohomology)**.** Let G be a Lie group linearly acting on a vector space V. (Such a vector space is called a G-module.) Let V be a left G-module. Let $(C^*(G,V),\delta)$ be the complex where $C_n(G,V)$ is the module of functions $\varphi : G^n \to V$ and the boundary map $\delta : C^n(G,V) \to C^{n+1}(G,V)$ is given by

$$\begin{aligned}(\delta\varphi)\,(g_1,\ldots,g_{n+1}) &= g_1 \cdot \varphi(g_2,\ldots,g_{n+1})\\ &+\sum_{i=1}^{n}(-1)^i\varphi(g_1,\ldots,g_{i-1},g_ig_{i+1},g_{i+2},\ldots,g_{n+1}) + (-1)^{n+1}\varphi(g_1,\ldots,g_n).\end{aligned} \tag{22.2}$$

A function $\varphi : G^k \to V$ is called a group k-cocycle, if $\delta\varphi = 0$ and a coboundary if $\varphi = \delta\psi$ for some cochain $\psi : G^{k-1} \to V$. The cohomology groups of G with coefficients in the left G-module V are the homology groups, denoted by $H^*(G,V)$, of the complex $(C^*(G,V),\delta)$.

Using the G-action Φ on M and the coadjoint action on $\mathfrak{g}^*$, it induces a natural G representation on $V = C^\infty(M,\mathfrak{g}^*)$ by

$$g \cdot f = \operatorname{Ad}^*_{g^{-1}} f \circ \Phi_g^{-1}; \quad f \in C^\infty(M,\mathfrak{g}^*).$$

Lemma 22.9. *Define a group 1-cochain*

$$\psi : G \to C^\infty(M,\mathfrak{g}^*)$$

by setting

$$\langle \psi(g), \xi\rangle(x) := J_\xi - \mathrm{Ad}^*_{g^{-1}} J_\xi(\Phi_g^{-1}(x))$$

for each $\xi \in \mathfrak{g}$. Then ψ is 1-cocycle.

Proof. We need to prove $\delta\psi = 0$ which is equivalent to saying

$$0 = \delta\psi(g,h) \Longleftrightarrow \psi(gh) = g\psi(h) + \psi(g). \tag{22.3}$$

For each given ξ and $x \in M$, we compute

$$\begin{aligned}
& \langle \psi(gh), \xi\rangle(x) \\
= & \ (\langle J - \mathrm{Ad}^*_{(gh)^{-1}} J \circ \Phi_{gh}^{-1}, \xi\rangle(x) \\
= & \ \langle J - \mathrm{Ad}^*_{g^{-1}} J \circ \Phi_g^{-1}, \xi\rangle(x) + \langle \mathrm{Ad}^*_{g^{-1}} J \circ \Phi_g^{-1}(x) - \mathrm{Ad}^*_{(gh)^{-1}} J \circ \Phi_{gh}^{-1}(x), \xi\rangle \\
= & \ \langle J - \mathrm{Ad}^*_{g^{-1}} J \circ \Phi_g^{-1}, \xi\rangle(x) + \langle \mathrm{Ad}^*_{g^{-1}}(J - \mathrm{Ad}^*_{h^{-1}} J \circ \Phi_h^{-1}) \circ \Phi_g^{-1}, \xi\rangle(x) \\
= & \ \langle \psi(g), \xi\rangle(x) + \langle \mathrm{Ad}^*_{g^{-1}} \psi(h) \circ \Phi_g^{-1}, \xi\rangle(x).
\end{aligned}$$

Therefore we have obtained

$$\psi(gh) = \psi(g) + \mathrm{Ad}^*_{g^{-1}} \psi(h) \circ \Phi_g^{-1} = \psi(g) + g\psi(h)$$

which finishes the proof of (22.3). □

Proposition 22.10. *Let (M, ω, Φ, J) be as above and M is connected. The function $\psi_{g,\xi} := \langle \psi(g), \xi\rangle$ is a constant function on M for all g, ξ. We let $\sigma_J : G \to \mathfrak{g}^*$ be defined by*

$$\langle \sigma_J(g), \xi\rangle = \text{the constant value of } \psi_{g,\xi}.$$

In particular σ_J is a 1-cocycle valued in $\mathfrak{g}^*$, i.e., satisfies

$$\sigma_J(gh) = \sigma_J(g) + \mathrm{Ad}^*_{g^{-1}} \sigma_J(h). \tag{22.4}$$

We call it the coadjoint cocycle *associated to J.*

Proof. We just take the differential of the function $\psi_{g,\xi}$:

$$\begin{aligned}
d\psi_{g,\xi}(x) &= d\langle J, \xi\rangle - d\langle J \circ \Phi_g^{-1}, \mathrm{Ad}_{g^{-1}} \xi\rangle \\
&= dJ_\xi - \langle d(J \circ d\Phi_g^{-1}, \mathrm{Ad}_{g^{-1}} \xi\rangle \\
&= X_{J_\xi} \rfloor \omega - (X_{J_{\mathrm{Ad}_{g^{-1}} \xi}} \rfloor \omega) \circ d\Phi_g^{-1} \\
&= (\xi_M \rfloor \omega) - ((\mathrm{Ad}_{g^{-1}} \xi)_M \rfloor \omega) \circ d\Phi_g^{-1} \\
&= (\xi_M \rfloor \omega) - ((\Phi_g^* \xi_M) \rfloor \omega) \circ d\Phi_g^{-1}.
\end{aligned}$$

But we evaluate

$$\begin{aligned}
(\Phi_g^* \xi_M \rfloor \omega)(Y(x)) &= \omega(\Phi_g^* \xi_M(x), Y(x)) = \omega(d\Phi_g^{-1} \xi_M(\Phi_g(x)), Y(x)) \\
&= \omega(\xi_M(\Phi_g(x)), d\Phi_g Y(x)) = (\xi_M \rfloor \omega)(d\Phi_g(Y(x))
\end{aligned}$$

where we use the symplectic property of $d\Phi_g$ for the first equality in the second line. Therefore $d\psi_{g,\xi}(x) = 0$ for all x and hence $\psi_{g,\xi}$ is a constant function.

For the last statement, we evaluate $\psi(gh) = g\psi(h) + \psi(g)$ at any point $x \in M$ and get $\psi(gh)(x) = (g\psi(h))(x) + \psi(g)(x)$. But

$$(g\psi(h))(x) = \mathrm{Ad}^*_{g^{-1}} \psi(h)(\Phi_g^{-1}(x)) = \mathrm{Ad}^*_{g^{-1}} \psi(h)(x)$$

since $\psi(h)$ is a constant function. By definition of σ_J, (22.4) follows. □

In general the cocycle on G associated to a G-module V measures the equivariance of the action Φ. We denote by $[\sigma_J]$ the associated cohomology class of the 1-cocycle σ_J.

Proposition 22.11. *Let Φ be a pre-Hamiltonian action of G on M. If J, J' be two moment maps associated to Φ with cocycles σ_J, $\sigma_{J'}$, then $[\sigma_J] = [\sigma_{J'}]$. We denote the common cohomology class by $[\Phi] \in H^1(G, \mathfrak{g}^*)$ and call it the* coadjoint cohomology class *of the G-action Φ on M.*

Proof. By definition, we have

$$d\langle J - J', \xi\rangle = 0$$

for all $\xi \in \mathfrak{g}$. In particular, $\langle J - J', \xi\rangle$ is a constant function, say ν, and hence

$$\sigma_J(g) - \sigma_{J'}(g) = J(\Phi_g(x)) - J'(\Phi_g(x)) - \mathrm{Ad}^*_{g^{-1}}(J(x) - J'(x)) = \nu - \mathrm{Ad}^*_{g^{-1}}\nu.$$

This proves $\sigma_J - \sigma_{J'} = \delta\nu$ and so finishes the proof. □

Therefore the cohomology class $[\sigma_J]$ does not depend on the choice of moment maps J but depends only on the pre-Hamiltonian action of G on (M, ω).

Theorem 22.12. *Suppose that the action of G on (M, ω) is pre-Hamiltonian. Let J be a moment map of the action and σ_J be its coadjoint cycle. If $[\sigma_J] = 0$, we can translate J to $J' := J - \nu$ for some $\nu \in \mathfrak{g}^*$ so that J' is Ad^*-equivariant.*

Proof. By the hypothesis $[\sigma_J] = 0$, we can express

$$\sigma_J = \delta\nu$$

for some zero-chain ν which is nothing but a constant function valued at $\nu \in \}^*$. Then by definition, for any $x \in M$, we have

$$J(x) - \mathrm{Ad}^*_{g^{-1}} J(\Phi_g(^{-1}(x)) = \sigma_J = \nu - \mathrm{Ad}^*_{g^{-1}}\nu$$

which is equivalent to

$$(J - \nu)(x) = J(\xi) - \nu = \mathrm{Ad}^*_{g^{-1}} J(\Phi_g^{-1}(x)) - \mathrm{Ad}^*_{g^{-1}}\nu.$$

But the latter is the same as

$$\mathrm{Ad}^*_{g^{-1}}(J(\Phi_g^{-1}(x)) - \nu) = \mathrm{Ad}^*_{g^{-1}}(J - \nu)(\Phi_g^{-1}(x)) = \mathrm{Ad}^*_{g^{-1}} J'(\Phi_g^{-1}(x)).$$

Combining the two, we have proved $J'(x) = \mathrm{Ad}^*_{g^{-1}} J'(\Phi_g^{-1}(x))$ which is equivalent to $J'(\Phi_g(x)) = \mathrm{Ad}^*_{g^{-1}} J'(x)$ and hence the proof. □

In other words, one can rephrase the definition of Hamiltonian G-spaces as a symplectic manifold (M, ω) equipped with a pre-Hamiltonian action Φ such that $[\Phi] = 0$.

Corollary 22.13. *Let (M, ω, Φ, J) be a pre-Hamiltonian G-space. Then*

$$J_{[\xi,\eta]} = \{J_\xi, J_\eta\}.$$

In particular the homomorphism $\mathfrak{g} \to \mathrm{symp}(M, \omega)$; $\xi \mapsto \xi_M$ can be lifted to a Lie algebra homomorphism to $\mathfrak{g} \mapsto C^\infty(M)$; $\xi \mapsto J_\xi$ under the Poisson bracket $\{\cdot, \cdot\})$ of $C^\infty(M)$ associated to ω.

Proof. A straightforward calculation. □

Example 22.14 (Abelian Lie group)**.** When G is a commutative Lie group, such as $G = T^n$, then the adjoint action and hence the coadjoint action is trivial. Therefore the equivariance is equivalent to the G-invariance

$$J(gx) = J(x)$$

for all $g \in G$ and in particular, J also satisfies

$$\{J_\xi, J_\eta\} = 0$$

for all $\xi\ \eta \in \mathfrak{g}$, i.e., the functions J_ξ are in involution. Finally $\sigma_J : G \to \mathfrak{g}^*$ being a cocycle means its G-invariance $\sigma_J(gh) = \sigma_J(g)$ for all h, and $[\sigma_J] = 0$ implies $\sigma_J = 0$.

Corollary 22.15. *Let $\Phi : T^n \times M \to M$ be a pre-Hamiltonian torus action. Then it is Hamiltonian.*

Proof. Let J be a moment map Φ. Then we have the coadjoint cycle is given by

$$\sigma_J(g) = \Phi_g^* J - J$$

which we know is a constant vector in $\mathfrak{g}^*$ when J is regarded as a $\mathfrak{g}^*$-valued function. We then average $\Phi_g^* J$ and consider the new map

$$J' = \int_{T^n} \Phi_g^* J \, d\mu$$

for the Haar measure $d\mu$ of the torus T^n with $\int_G d\mu = 1$. By the T^n-invariance of dJ, it follows J' is also a moment map of the action Φ. But obviously J' satisfies $\Phi_g^* J' = J'$ by definition. This finishes the proof. □

Many of important mechanical examples arise from the following theorem.

Theorem 22.16. *Let (M, ω) be an exact symplectic manifold with $\omega = d\lambda$. Suppose ω admits a G-invariant primitive λ with $\omega = d\lambda$. Then the map $J : M \to \mathfrak{g}^*$ defined by*

$$\langle J(x), \xi \rangle = -\lambda(\xi_M)(x)$$

is an Ad^**-equivariant moment map for the action.*

Proof. By the G-invariance $\phi_g^* \lambda = \lambda$ of λ, we obtain

$$d(\xi_M \rfloor \lambda) + \xi_M \rfloor d\lambda = 0,$$

which is equivalent to $-d(\xi_G \rfloor \lambda) = \xi \rfloor d\lambda$ for all $\xi \in \mathfrak{g}$. Therefore if we define $J : M \to \mathfrak{g}^*$ so that

$$J_\xi = -\xi_M \rfloor \lambda = -\lambda(\xi_M),$$

it satisfies the defining equation of the moment map $d\langle J, \xi \rangle = \xi_M \rfloor \omega$. To see the Ad^*-equivariance, we recall

$$\Phi_g^* \xi_M = (\mathrm{Ad}_{g^{-1}}^* \xi)_M.$$

We then compute

$$\begin{aligned} \langle \mathrm{Ad}_{g^{-1}}^* J, \xi \rangle(x) &= \langle J(x), \mathrm{Ad}_{g^{-1}} \xi \rangle = -\lambda\left((\mathrm{Ad}_{g^{-1}} \xi)_M(x)\right) \\ &= -\lambda(\Phi_g^* \xi_M(x)) = -\lambda(d\Phi_g^{-1}(\xi_M(\Phi_g(x))) = -\lambda(\xi_M(\Phi_g(x))) \\ &= \langle J(\Phi_g(x), \xi \rangle \end{aligned}$$

for all $\xi \in \mathfrak{g}$. This is equivalent to $\mathrm{Ad}_{g^{-1}}^* J = \Phi_g^* J$. This finishes the proof. □

Example 22.17 (Mechanical examples)**.** Let N be a configuration space and Φ be an action of G on N. Let X be a vector field on N.

We consider the phase space T^*N and the function $P_X : T^*N \to \mathbb{R}$ defined by

$$P_X(\alpha_q) = \langle \alpha_q, X(q) \rangle.$$

We call P_X the *momentum corresponding to* X. For example, if X is the generating vector field of a circular symmetry, then P_X is the associated angular momentum.

This leads to the algorithm to produce a conserved quantity when there is a continuous symmetry in the (Lagrangian) mechanical system.

(Nöther's principle) Let Φ be an action of G on N and let Φ^{T^*} be its cotangent lift on $M = T^*N$. Then this action is symplectic and admits an Ad^*-equivariant moment map given by

$$J_\xi(\alpha_q) = \langle \alpha_q, \xi_N(q) \rangle = P_{\xi_N}(\alpha_q),$$

i.e., $J_\xi = P_{\xi_N}$.

Proof. By Theorem 22.16, we have

$$J_\xi(\alpha_q) = -(-\theta)(\xi_{T^*G}(\alpha_q)) = \theta(\xi_{T^*G}(\alpha_q)) = \alpha_q(d\pi(\xi_{T^*G}(\alpha_q))) = \alpha_q(\xi_N(q)).$$

By definition of P_X with $X = \xi_N$ above, we have finished the proof. □

23. Marsden-Weinstein symplectic reduction theorem

Let us examine some general properties of the moment map.

We first rephrase Corollary 22.13 in the following Poisson property of the map when the symplectic manifold (M, ω) is regarded as a Poisson manifold.

Proposition 23.1. *Let (M, ω, Φ, J) be a Hamiltonian G-space. Then the moment map $J : M \to \mathfrak{g}^*$ is a Poisson map.*

Proof. This is just a translation of Corollary 22.13. □

One of the most fundamental property of the moment map is that it provides a canonical procedure of reducing the number of degrees of freedom for any G-invariant Hamiltonian system even when G is not abelian. When the group G is abelian, this reduction is a special case of coisotropic reduction. However when the group is non-abelian, this does not belong to the coisotropic reduction and its proof is more subtle. It was formulated by Marsden and Weinstein [MW] and is often called the Marsden-Weinstein reduction or the symplectic reduction.

Similarly as for the discussion on the completely integrable system, we consider a regular value μ of $J : M \to \mathfrak{g}^*$. Then $J^{-1}(\mu)$ becomes a submanifold of M. Denote by G_μ the isotropy group of the coadjoint action of G on $\mathfrak{g}^*$. Then G_μ acts on $J^{-1}(\mu)$.

Theorem 23.2 (Marsden-Weinstein reduction)**.** *Suppose the group G_μ acts freely on $J^{-1}(\mu)$. Denote by $i_\mu : J^{-1}(\mu) \to M$ and $\pi_\mu : J^{-1}(\mu) \to M_\mu$ for the quotient*

$$J^{-1}(\mu)/G_\mu =: M_\mu.$$

Then M_μ carries a canonical symplectic form ω_μ that is characterized by the identity

$$\pi_\mu^* \omega_\mu = i_\mu^* \omega.$$

If the action is proper, then M_μ is a Hausdorff manifold.

Proof. We start with proving the following lemma.

Lemma 23.3. *For $x \in J^{-1}(\mu)$, we have*

(1) *$T_y(G_\mu x) = T_y(G \cdot x) \cap T_y J^{-1}(\mu)$ for each $y \in G_\mu x$, and*
(2) *$T_y J^{-1}(\mu) = (T_y(G \cdot x))^\omega$.*

In particular for each $y \in G_\mu x$ with $x \in J^{-1}(\mu)$,

$$T_y(G_\mu x) = (T_y J^{-1}(\mu))^\omega \cap T_y J^{-1}(\mu). \tag{23.1}$$

Proof. Let $v \in T_y(G \cdot x)$. Then we can write $v = \xi_M(y)$ for some $\xi \in \mathfrak{g}$. By the Ad^*-equivariance, we derive

$$-\operatorname{ad}^*_\xi(\mu) = dJ(\xi_M(y))$$

for all $\xi \in \mathfrak{g}$. In particular, if $y \in G_\mu x$ and $\xi \in \mathfrak{g}_\mu = \operatorname{Lie} G_\mu$, then

$$dJ(\xi_M(y)) = -\operatorname{ad}^*_\xi(\mu) = 0 : \tag{23.2}$$

The vanishing follows by differentiating $\mathrm{Ad}^*_{\exp(-t\xi)}\mu = \mu$ in t at 0, since $\exp(-t\xi) \in G_\mu$ for all $t \in \mathbb{R}$ if $\xi \in \mathfrak{g}_\mu$. This proves $T_y(G_\mu x) \subset \ker d\pi|_y = T_y J^{-1}(\mu)$ where the equality comes from the regularity of μ. This proves $T_y(G_\mu x) \subset T_y(Gx) \cap T_y J^{-1}(\mu)$ for $y \in G_\mu x$. Conversely if $\xi_M(y) \in T_y J^{-1}(\mu)$, then $dJ(\xi_M(y)) = 0$. Combining this with (23.2), we obtain

$$0 = dJ(\xi_M(y)) = -\operatorname{ad}^*_\xi(\mu)$$

which proves $\xi \in \mathfrak{g}_\mu = \operatorname{Lie} G_\mu$. This finishes the proof of (1).

For (2), we note that if $\xi \in \mathfrak{g}$ and $w \in T_y M$ at $y \in Gx$,

$$\omega(\xi_M(y), w) = d\langle J(y), \xi\rangle(w)$$

by definition of J. Therefore $w \in \ker dJ(y) = T_y J^{-1}(\mu)$ if and only if $\omega(\xi_M(y), w) = 0$ for all $\xi \in \mathfrak{g}$. Since $T_y(G \cdot x)$ is spanned by $\xi_M(y)$, $\xi \in \mathfrak{g}$, this is equivalent to $w \in (T_y(G \cdot x))^\omega$. □

For $x \in J^{-1}(\mu)$, we denote by $[x]$ the coset $G_\mu x$ in M_μ. We know by definition that for each given $\widetilde{v} \in T_{[x]}(J^{-1}(\mu)/G_\mu)$, two lifts v, $v' \in T_y J^{-1}(\mu)$ with $y \in G_\mu x$ of $\widetilde{v}$ must satisfy

$$v - v' \in T_y(G_\mu x)$$

By (23.1), $v - v' \in (T_y J^{-1}(\mu))^\omega \cap T_y J^{-1}(\mu)$.

Therefore if we define

$$\omega_\mu(\widetilde{v}_1, \widetilde{v}_2) := \omega(v_1, v_2)$$

for $v_i \in T_y J^{-1}(\mu)$ with $y \in G_\mu x$, $i = 1, 2$, the right hand side does not depend on the representatives v_i but depend only on $\widetilde{v}_i$. By definition, the two form ω_μ satisfies

$$\pi_\mu^* \omega_\mu = i_\mu^* \omega.$$

In particular, $\pi_\mu^* d\omega_\mu = 0$. Since π_μ is a submersion, this implies $d\omega_\mu = 0$.

Finally let $\widetilde{v} \in T_{[x]} M_\mu$ and $\omega_\mu(\widetilde{v}, \widetilde{w}) = 0$ for all $\widetilde{w} \in T_{[x]} M_\mu$. Here the lifts v, w are unique modulo

$$T_y(G_\mu x) = T_y J^{-1}(\mu) \cap (T_y J^{-1}(\mu))^\omega.$$

This implies that for a (and so any) lift v of $\widetilde{v}$,

$$\omega(v, w) = 0$$

for all $w \in T_y J^{-1}(\mu)$. Therefore again by (23.1)

$$v \in (T_y J^{-1}(\mu))^\omega \cap T_y J^{-1}(\mu) = T_y(G_\mu \cdot x).$$

This proves $0 = d\pi_\mu(v) = \widetilde{v}$ which proves nondegeneracy of ω_μ.

The last statement follows because the given condition guarantees that the quotient

$$J^{-1}(\mu)/G_\mu$$

becomes Hausdorff. This finishes the proof of the theorem. □

Obviously M_μ is even dimensional, which is manifest from its dimension formula:

$$\begin{aligned} \dim M_\mu &= \dim J^{-1}(\mu) - \dim G_\mu = \dim M - \dim \mathfrak{g}^* - \dim G_\mu \\ &= \dim M - 2\dim \mathfrak{g}^* + (\dim \mathfrak{g}^* - \dim G_\mu) \\ &= \dim M - 2\dim \mathfrak{g}^* + \dim \mathcal{O}_\mu \end{aligned}$$

where $\mathcal{O}_\mu$ is the coadjoint orbit whose dimension we already know is even dimensional. Here the first equality is by definition of $M_\mu = J^{-1}/G_\mu$ and from the free action of G_μ on $J^{-1}(\mu)$, the second from the regularity of μ for J, and the last from the definition of the G-orbit $\mathcal{O}_\mu$ and that of G_μ.

Remark 23.4.
(1) If the group G is compact, the properness hypothesis in the theorem is automatic and so M_μ becomes a Hausdorff manifold.
(2) An important generalization of the theorem is the case where the isotropy group G_μ acts on $J^{-1}(\mu)$ not freely but with a finite isotropy group. In this case, the reduced phase space M_μ becomes a symplectic smooth orbifold, instead of a smooth manifold.

Example 23.5 (Symplectic quotient)**.** Let (M, ω, Φ, J) be a Hamiltonian G-space. If $\mu = 0$, then $G_\mu = G$. We often call the corresponding reduced space the *symplectic quotient* of M by G, and denote it by

$$M /\!\!/ G := J^{-1}(0)/G.$$

Example 23.6 (Coadjoint orbit as a reduced space)**.** Let $M = T^*G$ and consider the left action of G on M by

$$g \cdot (h, \alpha_h) = (gh, dL^*_{g^{-1}}(\alpha_h)).$$

Let ξ_{T^*G} be the vector field generated by $\xi \in \mathfrak{g}$. Then by definition, we have

$$d\pi(\xi_{T^*G}(h, \alpha_h)) = \xi_G(h).$$

Denote by $J : T^*G \to \mathfrak{g}^*$ be the moment map of this action of G on T^*G. Then by Theorem 22.16 applied to $\lambda = -\theta$, we have

$$\langle J(h, \alpha_h), \xi\rangle = \theta(\xi_{T^*G}(h, \alpha_h)) = \alpha_h(\xi_G(h)) = \alpha_h(dR_h(\xi)) = (dR_h)^*\alpha_h(\xi).$$

Therefore $J(h, \alpha_h) = (dR_h)^*\alpha_h$.

Lemma 23.7. *Let $J(h, \alpha_h) = \mu$ and β^μ be the right invariant one-form with its value $\beta^\mu(e) = \mu$, i.e., $\beta^\mu(h) = (dR_{h^{-1}})^*\mu$ for $h \in G$.*

(1) *Then $J^{-1}(\mu)$ is the graph of β^μ, i.e.,*

$$J^{-1}(\mu) = \{(h, (dR_{h^{-1}})^*\mu) \mid h \in g\} \tag{23.3}$$

(2) *Let G_μ be the isotropy group of* Ad^**-action on $\mathfrak{g}^*$ and $\beta = \beta^\mu$ be as above. Then*

$$G_\mu = \{g \in G \mid dL_g^*\beta = \beta\}.$$

Proof. It remains to show Statement ((2). If $g \in G_\mu$, then $(dL_g)^* dR^*_{g^{-1}}\mu = \mu$ by definition of Ad^*-action. By definition and the right invariance of β, we have $\beta(e) = \mu$ and

$$\beta(g) = dR^*_{g^{-1}}\mu.$$

Therefore $g \in G_\mu$ is equivalent to saying $(dL_g^*\beta)(e) = \beta(e)$. Then by the right invariance, this is also equivalent to global equality $dL_g^*\beta = \beta$. □

Now we consider a map $\phi : J^{-1}(\mu) \to \mathcal{O}_\mu$ defined by

$$\phi(h, \alpha_h) = dL_h^*\alpha_h = \mathrm{Ad}_h^*(\mu).$$

Obviously, this map descends to $[\phi] : J^{-1}(\mu)/G_\mu \to \mathcal{O}_\mu$ and is surjective. Since the action of G on T^*G is free, so is that of G_μ on $J^{-1}(\mu)$. We denote by $\pi_\mu : J^{-1}(\mu) \to J^{-1}(\mu)/G_\mu$ and $i_\mu : J^{-1}(\mu) \to T^*G$ the canonical projection and inclusion respectively.

Finally we will prove

Proposition 23.8. *Let ω_μ^{KKS} be the KKS symplectic form on the coadjoint orbit $\mathcal{O}_\mu$. Consider the two-form $\omega' := -\phi^*\omega_\mu^{KKS}$. Then it satisfies*

$$\pi_\mu^*\omega' = i_\mu^*\omega_0 \tag{23.4}$$

and ω' is G_μ-invariant. In particular the reduced symplectic form $(\omega_0)_\mu$ on the reduced space $\mathcal{O}_\mu \cong J^{-1}(\mu)/G_\mu$ is given by

$$(\omega_0)_\mu = -[\phi]^*\omega_\mu^{KKS}.$$

Proof. Let $\phi(h, \alpha_h) = \nu$ with $\nu = \mathrm{Ad}_h(\mu)$. It follows from the expression of $J^{-1}(\mu)$ in (23.3) any tangent vector X at (h, α_h) can be expressed as

$$X = \xi_{T^*G}(h, \alpha_h)$$

for some $\xi \in \mathfrak{g}$. Let

$$X = \xi_{T^*G}(h, \alpha_h),\ Y = \eta_{T^*G}(h, \alpha_h)$$

for some ξ, $\eta \in \mathfrak{g}$. From the expression of ϕ, we compute

$$d\phi(X) = \mathrm{ad}^*_\xi \nu.$$

Therefore we obtain

$$\omega_\mu(d\phi(X), d\phi(Y)) = \omega_\mu(\mathrm{ad}^*_\xi \nu, \mathrm{ad}^*_\eta \nu) = \langle \nu, [\xi, \eta]\rangle.$$

By definition, we have

$$\xi = \widetilde{\xi}(e) = \xi_G(e) = d\pi\xi_{T^*G}(e)$$

since $\widetilde{\xi}(e) = \xi_G(e)$ for the left-invariant vector field $\widetilde{\xi}$ on G. Then we can rewrite

$$\begin{aligned}
\langle \nu, [\xi, \eta]\rangle &= \nu([\xi_{T^*G}(e,\nu), \eta_{T^*G}(e,\nu)]) = \theta_{(e,\nu)}\left(-[\xi_{T^*G}, \eta_{T^*G}](e,\nu)\right) \\
&= d\theta\left(\xi_{T^*G}(e,\nu), \eta_{T^*G}(e,\nu)\right) = -\omega_0\left(\xi_{T^*G}(e,\nu), \eta_{T^*G}(e,\nu)\right) \\
&= -\omega_0\left(\xi_{T^*G}(dR^*_{h^{-1}}\mu), \eta_{T^*G}(h, dR^*_{h^{-1}}\mu)\right) = -\omega_0\left(\xi_{T^*G}(\alpha_h), \eta_{T^*G}(\alpha_h)\right)
\end{aligned}$$

where we use the invariance of ω_0 under the induced action on T^*G from the right action $R_h : G \to G$. This proves

$$\omega_0(X, Y) = -\omega_\mu(d\phi(X), d\phi(Y))$$

for all X, $Y \in T_{(h,\alpha_h)}J^{-1}(\mu)$. This finishes the proof. □

24. Functorial properties of moment map

In this section, we gather some functorial properties of the moment map.

Proposition 24.1. *Let* $(M_i, \omega_i, \Phi_i, J_i)$ *be Hamiltonian* G_i *spaces for* $i = 1, 2$. *Then their product*

$$(M_1 \times M_2, \omega_1 \times \omega_2, \Phi_1 \times \Phi_2, J_1 \times J_2)$$

is a Hamiltonian $(G_1 \times G_2)$*-space.*

Obviously when G acts on (M, ω) as an Hamiltonian action, so does any subgroup $H \subset G$ and so we have the associated moment map. We have a natural inclusion $i : \mathfrak{h} \subset \mathfrak{g}$ and projection $\pi := i^* : \mathfrak{g}^* \to \mathfrak{h}^*$.

Proposition 24.2. *Let* (M, ω, Φ, J) *be a Hamiltonian* G*-space. Then* $(M, \omega, \Phi|_H, \pi \circ J)$ *is a Hamiltonian* G*-space such that the following diagram commutes:*

$$\begin{array}{ccc} M & \xrightarrow{J} & \mathfrak{g}^* \\ {\scriptstyle =}\downarrow & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow{\pi \circ J} & \mathfrak{h}^* \end{array} .$$

Proof. We have only to check the defining equation of the moment map

$$d\langle J', \xi\rangle = \xi_M \rfloor \omega$$

for all $\xi \in \mathfrak{h} \subset \mathfrak{g}$ with $J' = \pi \circ J$.

But we have

$$\langle J', \xi\rangle = \langle \pi \circ J, \xi\rangle = \langle i^* \circ J, \xi\rangle = \langle J, i(\xi)\rangle = \langle J, \xi\rangle.$$

Therefore we derive

$$d\langle J', \xi\rangle = d\langle J, \xi\rangle = \xi_M \rfloor \omega$$

for all $\xi \in \mathfrak{h}$. This proves that $\pi \circ J$ is the moment map of the subgroup $H \subset G$. □

In particular, we have

Corollary 24.3. *Let* $(M_i, \omega_i, \Phi_i, J_i)$ *be two Hamiltonian* G*-spaces, i.e., for* $G_i = G$ *for both* $i = 1, 2$ *in Proposition 24.1. Then the moment map for this action is given by*

$$J(x_1, x_2) = J_1(x_1) + J_2(x_2).$$

Proof. We first note that the moment map J for the diagonal action is the composition

$$M_1 \times M_2 \to \mathfrak{g}^* \times \mathfrak{g}^* \to \mathfrak{g}^*$$

where the first map is the production map $J_1 \times J_2$ and the second map is the adjoint map Δ^* of the diagonal inclusion

$$\Delta : \xi \to (\xi, \xi); \quad \mathfrak{g} \to \mathfrak{g} \times \mathfrak{g}.$$

But the latter map $\Delta^* : \mathfrak{g}^* \times \mathfrak{g}^* \to \mathfrak{g}^*$ is nothing but the sum $(\mu, \nu) \mapsto \mu + \nu$ which finishes the proof.

□

So far we have reduced symplectic manifold alone. Now we want to reduce the Hamiltonian systems.

Theorem 24.4. *Let (M, ω, Φ, J) be a Hamiltonian G-space and let $H : M \to \mathbb{R}$ be a G-invariant Hamiltonian. Then the following hold:*

1. *The flow of X_H leaves $J^{-1}(\mu)$ for all $\mu \in \mathfrak{g}^*$, and hence canonically induces a flow on the reduced space $M_\mu := J^{-1}(\mu)/G_\mu$ at every $\mu \in \mathfrak{g}^*$.*
2. *Furthermore the resulting flow is also a Hamiltonian flow generated by a unique function, denoted by $H_\mu : M_\mu \to \mathbb{R}$, that satisfies*
$$H_\mu \circ \pi_\mu = H \circ i_\mu$$
on $J^{-1}(\mu)$. In particular, under this reduction of the flow, the level of Hamiltonian is unchanged.

Proof. The G-invariance of H means $H \circ \Phi_g = H$. By differentiating $\Phi_g^* H = H$ at $t = 0$ for $g = \exp t\xi$, we obtain $dH(\xi_M) = 0$. We can rewrite

$$dH(\xi_M) = \omega(X_H, \xi_M) = -(\xi_M \rfloor \omega)(X_H) = -d\langle J, \xi\rangle(X_H) = -X_H[J_\xi].$$

This proves $X_H[J_\xi] = 0$ and so the flow of H leaves the function J_ξ invariant for all $\xi \in \mathfrak{g}$. It is equivalent to saying that it leaves the value of J unchanged under the flow of X_H. This proves Statement (1).

For (2), we also see that H is invariant under the action of G_μ and so the function $H|_{J^{-1}(\mu)}$ descends to a function H_μ on the quotient $M_\mu = J^{-1}(\mu)/G_\mu$. By definition, we have

$$\pi_\mu^* dH_\mu = i_\mu^* dH$$

which is equivalent to

$$\pi_\mu^*(X_{H_\mu} \rfloor \omega_\mu) = i_\mu^*(X_H \rfloor \omega) \tag{24.1}$$

on $J^{-1}(\mu)$. It remains to show $\pi \circ \phi_H^t$ is the flow of X_{H_μ}. This is a consequence of the following lemma by considering the derivative

$$d\pi_\mu X_H(\phi_H^t(x))$$

of $\pi \circ \phi_H^t$ in t at $x \in J^{-1}(\mu)$.

Lemma 24.5. *$X_H|_{J^{-1}(\mu)}$ is projectible to the quotient $M_\mu = J^{-1}(\mu)/G_\mu$ and its projection is the Hamiltonian vector field X_{H_μ} on M_μ, i.e.,*

$$d\pi X_H(x) = X_{H_\mu}(\pi(x))$$

at every $x \in J^{-1}(\mu)$.

Proof. Denote $y = \pi_\mu(x)$. We evaluate

$$(d\pi_\mu X_H(x) \rfloor \omega_\mu)(v) = \omega_\mu(d\pi_\mu X_H(x), v)$$

for $v \in T_y M_\mu$. By considering $v = d\pi(u)$ for some $u \in T_x J^{-1}$, we arrive at

$$\begin{aligned} (d\pi_\mu X_H(y) \rfloor \omega_\mu)(v) &= (d\pi_\mu X_H(y) \rfloor \omega_\mu)(d\pi(u)) = (\pi^* \omega_\mu)(X_H(x), u) \\ &= i_\mu^* \omega(X_H(x), u) = \pi_\mu^*(X_{H_\mu} \rfloor \omega_\mu)(u) \\ &= X_{H_\mu} \rfloor \omega_\mu(d\pi(u)) = (X_{H_\mu} \rfloor \omega_\mu)(v). \end{aligned}$$

This proves

$$d\pi_\mu X_H(y) \rfloor \omega_\mu = X_{H_\mu} \rfloor \omega_\mu$$

at each $y \in M_\mu$. By nondegeneracy of ω_μ, we obtain

$$d\pi_\mu X_H(x) = X_{H_\mu}(\pi(x))$$

for all $x \in J^{-1}(\mu)$. This proves that $X_H|_{J^{-1}(\mu)}$ is projectible and its projection is X_{H_μ}. ∎

This lemma proves that the projected flow $\pi \circ \phi_H^t$ is again a Hamiltonian flow generated by the Hamiltonian vector field associated to the reduced Hamiltonian H_μ and hence the proof of (2). $\square$

References

[dC] Do Carmo, M.-P., Riemannian Geometry, Birkhäuser, Boston, 1992.

[MW] Marsden, J., Weinstein, A., *Reduction of symplectic manifolds wiht symmetry*, Rep. Math. Phys. 5 (1974), 121-130.

[Sp1] Spivak, M., Differential Geometry II, Publish & Perish Inc. Berkeley, 1979.

[Sp2] Spivak, M., Differential Geometry IV, Publish & Perish Inc. Berkeley, 1979.

[Wb94] Weibel, C., An Introduction to Homological Algebra, Cambridge Studies in Adv. Math. 38, Cambridge University Press, 1994. Press

Center for Geometry and Physics, Institute for Basic Sciences (IBS), Pohang, Korea & Department of Mathematics, POSTECH, Pohang, KOREA

E-mail address: yongoh1@postech.ac.kr