Geometry of Classical Mechanics

Nikolai Nowaczyk
<mail@nikno.de> http://math.nikno.de/

2013-01-21

Contents

1 Physical Motivation

1.1 Definition (particle)**.** Let (Q, g) be a Riemannian manifold, $\gamma : I \to Q$ be a smooth curve and $m > 0$ be any number. Then (γ, m) is a *particle of mass* m *in* Q. We also say γ is the *trajectory*, and m is its *mass* and (Q, g) its *configuration space*. Denote by $D_t : \mathcal{T}(\gamma) \to \mathcal{T}(\gamma)$ the covariant derivative induced by g, let $|\_| := |\_|_g$ be the norm in each tagent space of Q induced by g, let $E^{\text{kin}} := \frac{1}{2} m |\_|^2 \in \mathcal{C}^\infty(TM)$ be the *kinetic energy*. For any such particle, we define the following quantities:

- its *velocity* $\mathbf{v} := v_\gamma := \dot\gamma \in \mathcal{T}(\gamma)$,
- its *speed* $|\dot\gamma| \in \mathcal{C}^0(I, \mathbb{R}_{\geq 0})$,
- its *acceleration* $\mathbf{a} := a_\gamma := D_t \dot\gamma \in \mathcal{T}(\gamma)$,
- its *momentum* $\mathbf{p} := m \dot\gamma^\flat \in \mathcal{T}^*(\gamma)$,
- its *kinetic energy* $\mathbf{E}^{\text{kin}} := E^{\text{kin}}_\gamma := \frac{1}{2} m |\dot\gamma|^2 \in \mathcal{C}^0(I, \mathbb{R}_{\geq 0})$.

Any function $V \in \mathcal{C}^\infty(Q)$ is a *potential*. We set $E^{\text{pot}} := V \in \mathcal{C}^\infty(M)$, $\mathbf{E}^{\text{pot}} := E^{\text{pot}} \circ \gamma \in \mathcal{C}^\infty(I, \mathbb{R})$, the induced *kinetic energy*. The field

$$F := F_V := -\operatorname{grad}^g(V) \in \mathcal{T}(Q)$$

is the induced *force field*. The particle is *Newtonian with respect to* V, if it satisfies *Newton's Second Law*

$$\boxed{\mathbf{F} = m\mathbf{a},}$$

where $\mathbf{F} := F \circ \gamma$. This is to be understood as an equality in $\mathcal{T}(\gamma)$, where $m = m \operatorname{id} \in \operatorname{End}(\gamma)$ is thought of as a field of endomorphisms.

1.1 Mechanics in Euclidan space

1.2 Theorem (Hamilton vs. Newton)**.** Consider $\mathbb{R}^n$ with the Euclidean metric $\bar g$ as a configuration manifold and let $T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$ be the associated phase space. Label the coordinates by $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ and let $\omega = \sum_i dq^i \wedge dp^i$ be the canonical symplectic structure, cf. Lemma 2.4. Let $V \in \mathcal{C}^\infty(\mathbb{R}^n)$ be any potential and $m > 0$ be any mass. Let $F := F_V := -\operatorname{grad} V$ be the induced force field and

$$\begin{aligned} H : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (q, p) &\mapsto \tfrac{1}{2m}|p|^2 + V(q). \end{aligned}$$

be the induced *energy Hamiltonian*. Then $(\mathbb{R}^{2n}, \omega, H)$ is a Hamiltonian system in the sense of Definition 3.12 and any curve $\gamma : I \to \mathbb{R}^n$ satisfies *Newton's second law*

$$\forall t \in I : m\ddot\gamma(t) = F(\gamma(t)),$$

if and only if $\tau := (\tau_q, \tau_p) := (\gamma, m\dot{\gamma}) : I \to \mathbb{R}^{2n}$ satisfies Hamilons equations (see (3.3))

$$\begin{aligned} \dot{\tau}_q^i(t) &= \dot{\gamma}^i(t) = \frac{\partial H}{\partial p^i}(\tau(t)), \\ \dot{\tau}_p^i(t) &= \ddot{\gamma}^i(t) = -\frac{\partial H}{\partial q^i}(\tau(t)), \end{aligned} \tag{1.1}$$

Proof. Calculating the gradient of H, we obtain

$$\operatorname{grad} H = \sum_{i=1}^n \frac{2}{2m}|p|\frac{p_i}{|p|}\frac{\partial}{\partial p_i} + \frac{\partial V}{\partial q^i}\frac{\partial}{\partial q^i} = \sum_{i=1}^n \frac{p_i}{m}\frac{\partial}{\partial p_i} - F^i \frac{\partial}{\partial q^i}$$

Consequently, Hamiltons equations (1.1) are equivalent in this case to

$$\begin{aligned} \dot{\tau}_q^i(t) &= \tfrac{1}{m}\tau_p^i(t), \\ \dot{\tau}_p^i(t) &= F^i(\tau_q(t)). \end{aligned}$$

The first equation is automatically satisfied by definition of τ. Also by definition of τ, the second equation is equivalent to

$$\dot{\tau}_p = F(\tau_q) \Longleftrightarrow m\ddot{\gamma} = F(\gamma),$$

which is Newton's second law. □

1.3 Remark. The Hamiltonian is traditionally denoted by $E := H$, called *energy*, and decomposed into $E = E^{\mathrm{kin}} + E^{\mathrm{pot}}$, the kinetic respectively potential energy. Therefore, we obtain

$$\mathbf{E} = \mathbf{E}^{\mathrm{kin}} + \mathbf{E}^{\mathrm{pot}}.$$

It follows from Lemma 3.14(i) that the energy is conserved, i.e. $\dot{\mathbf{E}} = 0$.

1.4 Theorem (Newton vs. Lagrange)**.** Let $V \in \mathcal{C}^\infty(\mathbb{R}^n)$ be any potential with induced force field $F = -\nabla V$ and define

$$\begin{aligned} L : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (x, v) &\mapsto \tfrac{1}{2}mv^2 - V(x) \end{aligned}$$

A curve $\gamma : I \to \mathbb{R}^n$ satisfies Newtons second law

$$\forall t \in I : m\ddot{\gamma}(t) = F(\gamma(t))$$

if and only if γ satisfies the *Euler-Lagrange equation*

$$\forall t \in I : \frac{\partial L}{\partial x^i}(\gamma(t), \dot{\gamma}(t)) = \frac{d}{dt}\frac{\partial L}{\partial v^i}(\gamma(t), \dot{\gamma}(t)).$$

Proof. This follows from the simple fact that

$$\begin{aligned} \frac{\partial L}{\partial x}(x, v) &= -\nabla V(x) = F(x), \\ \frac{\partial L}{\partial v}(x, v) &= mv. \end{aligned}$$

□

2 Symplectic Geometry Basics

Let M be a smooth manifold.

2.1 Definition. A *symplectic form* is an element $\omega \in \Omega^2(M)$, which is closed and non-degenerate. A tuple (M, ω) such that ω is a symplectic form, is a *symplectic manifold*.

2.2 Definition. A diffeomorphism $f : (M, \omega) \to (N, \eta)$ between symplectic manifolds is a *symplectomorphism*, if $f^*\eta = \omega$.

2.3 Definition (tautological 1-form)**.** Let Q be a smooth manifold, $M := T^*Q$ and $\pi : M \to Q$ be the canonical projection. Then $\tau \in \Omega^1(M)$ defined by

$$\forall \alpha \in M : \forall X \in T_\alpha M : \tau|_\alpha(X) := \pi^*\alpha(X) = \alpha(\pi_* X). \tag{2.1}$$

is the *tautological* 1*-form*. The form $\omega := -d\tau \in \Omega^2(M)$ is the *canonical symplectic form* of Q. In this situation, we also say that Q is the *configuration space* and M is the *phase space*.

2.4 Lemma. Let Q be a smooth manifold and ω be the canonical symplectic form. Then (T^*Q, ω) is a symplectic manifold. Let $\pi : T^*Q \to Q$ be the canonical projection, (x^i) be local coordinates for Q, $\tilde{x}^i := \pi^* x^i = x^i \circ \pi$ and (ξ^i) be the local coframe on T^*Q induced by (x^i). Then ω can be expressed locally by

$$\boxed{\omega = \sum_{i=1}^n d\tilde{x}^i \wedge d\xi^i.}$$

2.5 Theorem (Darboux)**.** Let (M, ω) be a $2n$-dimensional symplectic manifold. Near every point $p \in M$, there are smooth coordinates $(x^1, y^1, \ldots, x^n, y^n)$ such that

$$\boxed{\omega = \sum_{i=1}^n dx^i \wedge dy^i.} \tag{2.2}$$

Those coordinates are called *Darboux coordinates* or *symplectic coordinates* or *canonical coordinates*.

3 Hamiltonian Geometry

In this section let (M, ω) be a symplectic manifold.

3.1 Hamiltonian Vector fields

3.1 Definition. Let $X \in \mathcal{T}(M)$ be any vector field.

(i) X is *symplectic*, if $\iota_X \omega$ is closed.

(ii) X is *Hamiltonian*, if $\iota_X\omega$ is exact.

(iii) X is a *Hamiltonian vector field for* $H \in \mathcal{C}^\infty(M)$, if $\iota_X\omega = dH$.

We denote by $\mathcal{T}^{\mathrm{ham}}(M)$ the Hamiltonian vector fields and by $\mathcal{T}^{\mathrm{symp}}(M)$ the symplectic vector fields on M.

3.2 Lemma. Let $X \in \mathcal{T}(M)$ and $H \in \mathcal{C}^\infty(M)$.

(i) X is symplectic if and only if X is locally Hamiltonian.

(ii) If $H^1_{\mathrm{dR}}(M) = \{0\}$, then X is symplectic if and only if X is Hamiltonian.

(iii) A Hamiltonian vector field X for any function H is always Hamiltonian.

(iv) Any Hamiltonian vector field is symplectic.

Proof.

(i) This follows from the general fact that a differential form is closed if and only if it is locally exact.

(ii) This follows from the definition of de Rahm cohomology.

(iii) By definition of X_H.

(iv) Any exact form is closed.

□

3.3 Lemma. Let $X \in \mathcal{T}(M)$ and $\theta : \mathcal{D} \to M$ be the induced flow. Then the following are equivalent.

(i) X is symplectic.

(ii) All diffeomorphisms θ_t are symplectomorphisms.

(iii) $\mathcal{L}_X\omega = 0$.

Proof.

"(i)⟹(ii)"
Since $\theta_0 = \mathrm{id}$, clearly $\theta_0^*\omega = \omega$ and for any $p \in M$ and $t_0 \in \mathcal{D}^{(p)}$

$$\tfrac{d}{dt}\theta_t^*\omega|_{t=t_0} = \theta_{t_0}^*\tfrac{d}{dt}\theta_t^*\omega|_{t=0} = \theta_{t_0}^*\mathcal{L}_X\omega \overset{\text{A.1}}{=} \theta_{t_0}^*(d\underbrace{\iota_{X_H}\omega}_{=0} + \iota_{X_H}\underbrace{d\omega}_{=0}) = 0,$$

since $\iota_X\omega$ and ω are both closed.

"(ii)⟹(iii)" By definition

$$\mathcal{L}_X\omega = \tfrac{d}{dt}\theta_t^*\omega = \tfrac{d}{dt}\omega = 0.$$

"(iii)⟹(i)" By (A.1)

$$0 = \mathcal{L}_X\omega = d\iota_X\omega + \iota_X d\omega = d\iota_X\omega.$$

□

3.4 Lemma. Let $H \in \mathcal{C}^\infty(M)$.

(i) There exists a unique Hamiltonian vector field $X_H \in \mathcal{T}(M)$ of H.

(ii) In any local Darboux coordinates (x^i, y^i) the Hamiltonian field X_H is given by

$$\boxed{X_H = \sum_{i=1}^n \frac{\partial H}{\partial y^i}\frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial y^i}.} \tag{3.1}$$

(iii) H is constant along the flow θ of X_H, thus each integral curve of X_H is contained in a level set of H.

(iv) At each regular point $p \in M$ of H, $X_H(p)$ is tangent to the level set $H^{-1}(\{H(p)\})$ of H.

(v) Each member of the family of diffeomorphisms θ_t defined by the flow θ of X_H is a symplectomorphism.

Proof.

(i) To see uniqueness, assume $X, Y \in \mathcal{T}(M)$ satisfy $\iota_X\omega = dH = \iota_Y\omega$. This implies for any $V \in \mathcal{T}(M)$

$$\omega(X,V) = \iota_X\omega(V) = \iota_Y\omega(V) = \omega(Y,V) \Longrightarrow \omega(X-Y,V) = 0.$$

Since ω is non-degenerate, this implies $X = Y$. To show existence it suffices to check the local formula (3.1).

(ii) We calculate

$$\begin{aligned}\iota_{X_H}\omega &\overset{(2.2)}{=} \sum_{i=1}^n \iota_{X_H}(dx^i \wedge dy^i)\\ &\overset{(A.2)}{=} \sum_{i=1}^n \iota_{X_H}(dx^i)\wedge dy^i - dx^i \wedge \iota_{X_H}(dy^i)\\ &\overset{(3.1)}{=} \sum_{i=1}^n \frac{\partial H}{\partial y^i}dy^i + \frac{\partial H}{\partial x^i}dx^i\\ &= dH.\end{aligned}$$

(iii) For any $(t_0, p) \in \mathcal{D}$

$$\tfrac{d}{dt}H(\theta^{(p)}(t))|_{t=t_0} = dH(\theta^{(p)}(t_0))(\dot\theta^{(p)}(t_0)) = \iota_{X_H}\omega|_{\theta(t_0,p)}(X_H|_{\theta(t_0,p)}) = 0.$$

(iv) Since

$$X_H(H) = \mathcal{L}_{X_H}(H) = \iota_{X_H}(dH) + d\iota_X H = \iota_{X_H}^2\omega = 0,$$

we obtain

(v) This follows from Lemma 3.3.

□

3.2 Lie Brackets

For this subsection we do not need the symplectic structure ω on M.

3.5 Definition (Lie algebra)**.** A *Lie algebra* is a real vector space $\mathfrak{g}$ together with map

$$[\_,\_] : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$$

called *Lie bracket* such that the following holds:

(i) Bilinearity: $[\_,\_]$ is bilinear over $\mathbb{R}$.

(ii) Anti-symmetry: $\forall X, Y \in \mathfrak{g} : [X, Y] = -[Y, X]$.

(iii) Jacobi Identity: $\forall X, Y, Z \in \mathfrak{g} : [X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$.

3.6 Definition ((anti)-homomorphisms of Lie algebras)**.** A *homomorphism of Lie algebras* $f : (\mathfrak{g}, [\_,\_]) \to (\mathfrak{h}, [\![\_,\_]\!])$ is a linear map such that

$$\boxed{\forall X, Y \in \mathfrak{g} : [\![f(X), f(Y)]\!] = f([X, Y]).}$$

We say f is an *anti-homomorphism* if f is a linear map such that

$$\boxed{\forall X, Y \in \mathfrak{g} : [\![f(X), f(Y)]\!] = -f([X, Y]).}$$

3.7 Theorem (Lie bracket on vector fields)**.** For any two vector fields $X, Y \in \mathcal{T}(M)$ define

$$\boxed{\forall f \in \mathcal{C}^\infty(M) : [X, Y](f) := X(Y(f)) - Y(X(f))}$$

Then $[X, Y] \in \mathcal{C}^\infty(M)$ is a vector field, called *Lie bracket* of X and Y. In addition the following hold:

(i) $[X, Y] \in \mathcal{T}(M)$ and in any local coordinates (x^i)

$$[X, Y] = (X^i \frac{dY^j}{\partial x^i} - Y^i \frac{dX^j}{\partial x^i}) \frac{\partial}{\partial x^i} = (X(Y^j) - Y(X^j)) \frac{\partial}{\partial x^j}$$

(ii) $(\mathcal{T}(M), [\_,\_])$ is a real Lie algebra.

(iii) If $F : M \to N$ is a smooth map and $X_1, X_2 \in \mathcal{T}(M)$, $Y_1, Y_2 \in \mathcal{T}(N)$ such that X_i is F-related to Y_i, $i = 1, 2$, then $[X_1, X_2]$ is F-related to $[Y_1, Y_2]$.

(iv) For any $f, g \in \mathcal{C}^\infty(M)$

$$[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X.$$

Proof. A proof of this can be found in [1, Chpt. 4] □

3.3 Poisson brackets

3.8 Definition (Poisson algebra)**.** A Poisson algebra $(\mathcal{P}, \{\_,\_\})$ is a commutative associative $\mathbb{R}$ - algebra $\mathcal{P}$ together with a Lie bracket $\{\_,\_\}$, which satisfies the *Leibniz rule*

$$\boxed{\forall f, g, h \in \mathcal{P} : \{f, gh\} = \{f, g\}h + g\{f, h\}.}$$

Let (M, ω) be a symplectic manifold again.

3.9 Definition (Poisson bracket)**.** Let $f, g \in \mathcal{C}^\infty(M)$. Then

$$\boxed{\{f, g\} := \omega(X_f, X_g) = \iota_{X_f}\omega(X_g) = df(X_g) = X_g(f)}$$

is the *Poisson bracket of* f *and* g. We say f and g *Poisson commute*, if $\{f, g\} = 0$.

3.10 Lemma. If $X, Y \in \mathcal{T}^{\mathrm{symp}}(M)$, then $[X, Y] \in \mathcal{T}^{\mathrm{ham}}(M)$ and

$$[X, Y] = d(\omega(Y, X)) \tag{3.2}$$

Proof. We calculate

$$\begin{aligned}
\iota_{[X,Y]}\omega &\overset{(A.3)}{=} \mathcal{L}_X\iota_Y\omega - \iota_Y\mathcal{L}_X\omega \\
&\overset{(A.1)}{=} \iota_X \underbrace{d\iota_Y\omega}_{=0} + d\iota_X\iota_Y\omega - \iota_Y\iota_X \underbrace{d\omega}_{=0} - \iota_Y \underbrace{d\iota_X\omega}_{=0} \\
&= d(\omega(Y, X)),
\end{aligned}$$

since X and Y are symplectic and ω is closed. □

3.11 Theorem. $(\mathcal{C}^\infty(M), \{\_,\_\})$ is a Poisson algebra and the map

$$\begin{array}{rcl}
(\mathcal{C}^\infty(M), \{_,_\}) & \to & (\mathcal{T}(M), [_,_]) \\
f & \mapsto & X_f
\end{array}$$

is a Lie algebra anti-homomorphism.

Proof. The Poisson bracket $\{\_,\_\}$ is bilinear by (3.1) and anti-symmetric by construction. To see that this map is an anti-homorphism, we calculate

$$-[X_g, X_f] \overset{(3.2)}{=} d\omega(X_f, X_g) = X_{\omega(X_f, X_g)} = X_{\{f,g\}}.$$

To see that the Jacobi-identity is satisfied, we calculate

$$\begin{aligned}
\{f, \{g, h\}\} &= X_{\{g,h\}}f = -[X_g, X_h]f \\
&= -X_g(X_h(f)) + X_h(X_g(f)) = -X_g(\{f, h\}) + X_h(\{f, g\}) \\
&= -\{\{f, h\}, g\} + \{\{f, g\}, h\}.
\end{aligned}$$

Finally, the Leibniz rule follows from

$$\{fg, h\} = X_h(fg) = X_h(f)g + X_h(g)f = \{f, h\}g + \{g, h\}f.$$

□

3.4 Hamiltonian Systems

3.12 Definition (Hamiltonian system)**.** For any $H \in \mathcal{C}^\infty(M)$ we say the triple (M, ω, H) is a *Hamiltonian system* with *Hamiltonian* H. The flow of X_H is the *Hamiltonian flow* and its integral curves γ are the *trajectories* or *orbits* of the system. In any local Darboux coordinates (x^i, y^i) they satisfy *Hamilton's equations*

$$\boxed{\begin{aligned} \dot\gamma_x^i(t) &= \frac{\partial H}{\partial y^i}(\gamma(t)), \\ \dot\gamma_y^i(t) &= -\frac{\partial H}{\partial x^i}(\gamma(t)), \end{aligned}} \tag{3.3}$$

as can be easily derived from (3.1). For an important example of this, see Theorem 1.2.

3.13 Definition (conserved quantity, infinitesimal symmetry)**.** Let (M, ω, H) be a Hamiltonian system. A function $f \in \mathcal{C}^\infty(M)$ is a *conserved quantity*, if f is constant along any trajectory of the system. Sometimes f is also called a *integral of motion*, *constant of motion* or *integral of first kind*. A vector field $V \in \mathcal{T}(M)$ is an *infinitesimal symmetry* of that system, if ω and H are invariant under the flow $\vartheta : \mathcal{D} \to M$ of V, i.e. for any $(t,p) \in \mathcal{D}$

$$H(\vartheta(t,p)) = H(p), \qquad\qquad {\vartheta_t}^*\omega|_p = \omega|_{\vartheta(t,p)}.$$

3.14 Lemma. Let (M, ω, H) be a Hamiltonian system.

(i) A function $f \in \mathcal{C}^\infty(M)$ is a conserved quantity if and only if $\{H, f\} = 0$.

(ii) A vector field $V \in \mathcal{T}(M)$ is an infinitesimal symmetry if and only if it is symplectic and $VH = 0$.

(iii) If ϑ is the flow of an infinitesimal symmetry V and γ is a trajectory of the system, then for any $s \in \mathbb{R}$, the curve $\vartheta_s \circ \gamma$ is also a trajectory on its domain of definition.

Proof.

(i) Denote by θ the flow of X_H and calculate for any $p \in M$

$$\begin{aligned} \tfrac{d}{dt}(f \circ \theta^{(p)}) &= \theta^{(p)*}(\mathcal{L}_{X_H} f) = \theta^{(p)*}(\iota_{X_H} df) = \theta^{(p)*}(\iota_{X_H}\iota_{X_f}\omega) \\ &= \theta^{(p)*}(\omega(X_f, X_H)) = \theta^{(p)*}(\{f, H\}). \end{aligned}$$

This implies the claim.

(ii) By Lemma 3.3 V is symplectic, if and only if ω is invariant under ϑ. For any $(t,p) \in \mathcal{D}$

$$\tfrac{d}{dt}(H \circ \vartheta^{(p)})(t) = dH_{\vartheta^{(p)}(t)}(\dot\vartheta^{(p)(t)}) = dH_{\vartheta(t,p)}(V|_{\vartheta(t,p)}) = V(H)|_{\vartheta(t,p)},$$

thus $V(H) = 0$ if and only if H is invariant under ϑ.

(iii)

□

3.15 Corollary. For any Hamiltonian system (M, ω, H), the function H is a conserved quantity and X_H is an infinitesimal symmetry.

Proof. By Lemma 3.4(iii), H is a conserved quantity. By definition X_H is symplectic, and since $X_H(X) = 0$ by (3.1), we obtain that X_H is an infinitisemal symmetry by Lemma 3.14(ii). □

3.16 Theorem (Noether's Theorem)**.** Let (M, ω, H) be a Hamiltonian system.

(i) If $f \in \mathcal{C}^\infty(M)$ is a conserved quantity, then X_f is an infinitesimal symmetry.

(ii) If $H^1_{\mathrm{dR}}(M) = \{0\}$, then for every infinitesimal symmetry V there exists a function f such that $V = X_f$. In that case f is unique up to locally constant functions.

3.17 Definition (independent)**.** A system of functions $f_1, \ldots, f_n \in \mathcal{C}^\infty(M)$ is *independent*, if there exists an open dense subset $U \subset M$ such that for all $p \in U$ the covectors $(df_1|_p, \ldots, df_n|_p) \in T^*_p M$ are linearly independent.

3.18 Definition (completely integrable)**.** A Hamiltonian system (M^{2n}, ω, H) is *(completely) integrable*, if there are conserved quantities $f_1 = H, f_2 \ldots, f_n$ that are independent and satisfy $\{f_i, f_j\} = 0$ for all $1 \leq i, j \leq n$.

3.19 Lemma. Let (M^{2n}, ω, H) be an integrable system with conserved quantities $f_1 = H, f_2, \ldots, f_n$. Let $c \in \mathbb{R}^n$ be a regular value of $f := (f_1, \ldots, f_n)$. If the Hamiltonian vector fields $X_{f_1}, \ldots, X_{f_n}$ are complete on the level set $f^{-1}(c)$, then $f^{-1}(c)$ are homogenous spaces for $\mathbb{R}^n$, i.e. are diffeomorphic to $\mathbb{R}^{n-k} \times \mathbb{T}^k$, where $\mathbb{T}^k$ is the k-dimensional Torus.

3.20 Theorem (Arnold-Liouville)**.** Let (M^{2n}, ω, H) be an integrable system with conserved quantites $f_1 = H, \ldots, f_n$. Let $c \in \mathbb{R}^n$ be a regular value of $f := (f_1, \ldots, f_n)$. Then the level set $f^{-1}(c)$ is a Lagrangian submanifold of M. In addition the following holds:

(i) If the flows of $X_{f_1}, \ldots, X_{f_n}$ starting at a point $p \in f^{-1}(c)$ are complete, then the conneted component of $f^{-1}(c)$ containing p is a homogenous space for $\mathbb{R}^n$. With respect to this affine structure, that component has coordinates $\varphi_1, \ldots, \varphi_n$, called *angle coordinates*, in which the flows of the vector fields $X_{f_1}, \ldots, X_{f_n}$ are linear.

(ii) There are coordinates $\psi_1, \ldots, \psi_n$, known as *action coordinates*, complementary to the angle coordinates such that the ψ_i's are conserved quantities and $(\varphi_1, \ldots, \varphi_n, \psi_1, \ldots, \psi_n)$ form a Darboux chart.

4 Variational Principles

4.1 The Euler-Lagrange Equation

4.1 Definition (action)**.** Let M be any smooth manifold and $F \in \mathcal{C}^\infty(TM, \mathbb{R})$. Let $\gamma : [a,b] \to M$ be a smooth curve. Then

$$\mathcal{A}^F_\gamma := \int_a^b F(\dot\gamma(t))dt$$

is the *action* of γ with respect to F. For any fixed $p, q \in M$ we set

$$\mathcal{P}(a,b,p,q) := \{\gamma \in \mathcal{C}^\infty([a,b], M) \mid \gamma(a) = p,\ \gamma(b) = q\}.$$

The action defines a map

$$\begin{aligned} \mathcal{A}^F : \mathcal{P}(a,b,p,q) &\to \mathbb{R} \\ \gamma &\mapsto \mathcal{A}^F_\gamma. \end{aligned}$$

A curve $\gamma_0 \in \mathcal{P}$ is *minimizing* (in $\mathcal{P}(a,b,p,q)$), if

$$\mathcal{A}^F_{\gamma_0} = \min_{\gamma \in \mathcal{P}(a,b,p,q)} \mathcal{A}^F_\gamma.$$

We fix an action F and drop its superscript in notation.

4.2 Lemma. Let $\gamma_0 : [a,b] \to M$ be minimizing. Let $[a_1,b_1] \subset [a,b]$ be a subinterval and let $p_1 := \gamma_0(a_1)$, $b_1 := \gamma_0(b_1)$. Then $\gamma_1 := \gamma_0|_{[a_1,b_1]}$ is minimizing among $\mathcal{P}(a_1,b_1,p_1,q_1)$.

4.3 Theorem. Let $F \in \mathcal{C}^\infty(TM, \mathbb{R})$, $\mathcal{A} = \mathcal{A}^F$ be the induced action and let $\gamma \in \mathcal{P}(a,b,p,q)$ be a minimizer. Let $t \in [a,b]$, (x^i) be local coordinates near $\gamma(t)$ and let (v^i) be the induced local frame for TM. Then γ satisfies the *Euler-Langrange equation*

$$\boxed{\frac{\partial F}{\partial x^i}(\gamma(t), \dot\gamma(t)) = \frac{d}{dt}\frac{\partial F}{\partial v^i}(\gamma(t), \dot\gamma(t)).} \tag{EL}$$

Proof. Choose $t \in I$ and local coordinates $(x^i) : U \to V$, $(v^i) : TU \to TV$ near $\gamma(t)$. By continuity there exist $a_1, b_1 \in \mathbb{R}$ such that $t \in [a_1,b_1] \subset [a,b]$ and $\gamma([a_1,b_1]) \subset U$. By Lemma 4.2, the restriction $\gamma|_{[a_1,b_1]}$ is also minimizing among $\mathcal{P}(a_1,b_1,\gamma(a_1),\gamma(b_1))$. Therefore we can assume that the endpoints p and q lie in the same coordinate domain U. Consequently, we can assume that U is a subset of $\mathbb{R}^n$. Let $c = (c_1, \ldots, c_n) : [a,b] \to U$ be any smooth curve satisfying $c(a) = 0 = c(b)$ For all sufficiently small $\varepsilon > 0$ the curve

$$\gamma_\varepsilon := \gamma + \varepsilon c$$

is smooth and belongs to $\mathcal{P}(a,b,p,q)$. Setting $\mathcal{A}_\varepsilon := \mathcal{A}_{\gamma_\varepsilon}$, we obtain for the derivative

$$\frac{d\mathcal{A}_\varepsilon}{d\varepsilon} = \int_a^b \frac{d}{d\varepsilon} F(\gamma_\varepsilon(t), \dot\gamma_\varepsilon(t))dt = \int_a^b \sum_{i=1}^n \frac{\partial F}{\partial x_i}(\gamma_\varepsilon, \dot\gamma_\varepsilon)c_i + \frac{\partial F}{\partial v^i}(\gamma_\varepsilon, \dot\gamma_\varepsilon)\dot c_i dt \tag{4.1}$$

Since γ is minimizing

$$0 = \frac{d\mathcal{A}_{\gamma_\varepsilon}}{d\varepsilon}|_{\varepsilon=0} = \int_a^b \sum_{i=1}^n \frac{\partial F}{\partial x_i}(\gamma(t),\dot\gamma(t))c_i(t) + \frac{\partial F}{\partial v^i}(\gamma(t),\dot\gamma(t))\dot c_i(t)dt$$
$$= \sum_{i=1}^n \int_a^b \left(\frac{\partial F}{\partial x_i}(\gamma(t),\dot\gamma(t)) - \frac{d}{dt}\frac{\partial F}{\partial v^i}(\gamma(t),\dot\gamma(t))\right)c_i(t)dt,$$

where we used partial integration and the fact that $c(a) = 0 = c(b)$. Since this holds for all such curves c, we obtain the claim. □

4.4 Theorem (convexity and minimization)**.** Assume that $F \in \mathcal{C}^\infty(TM, \mathbb{R})$ satisfies

$$\forall (x,v) \in TM : \det\left(\frac{d^2F}{\partial v^i \partial v^j}(x,v)\right) > 0, \tag{4.2}$$

i.e. for any fixed x, $v \mapsto F(x,v)$, is strictly convex. Let $\gamma \in \mathcal{P}(a,b,p,q)$ be a solution of the Euler-Lagrange equation (EL). Then for every sufficiently small subinterval $[a_1,b_1] \subset [a,b]$, the curve $\gamma|_{[a_1,b_1]}$ is minimizing in $\mathcal{P}(a_1,b_1,p_1,q_1)$, $p_1 := \gamma(a_1)$, $q_1 := \gamma(b_1)$.

Proof. Let $c_1,\dots,c_n \in \mathcal{C}^\infty([a,b])$, $c_i(a) = c_i(b) = 0$, $c := (c_1,\dots,c_n)$ and set $\gamma_\varepsilon := \gamma + \varepsilon c \in \mathcal{P}(a,b,p,q)$ and $\mathcal{A}_\varepsilon := \mathcal{A}_{\gamma_\varepsilon}$. Then the Euler-Lagrange equation is satisfied if and only if $\frac{d\mathcal{A}_\varepsilon}{d\varepsilon}(0) = 0$, so 0 is a critical point for $\mathcal{A}_\varepsilon$. To see that it is minimizing, we analyse the second derivative:

$$\frac{d^2\mathcal{A}_\varepsilon}{d\varepsilon^2}(0) \overset{(4.1)}{=} \int_a^b \frac{d}{d\varepsilon}\sum_{i=1}^n \frac{\partial F}{\partial x_i}(\gamma_\varepsilon,\dot\gamma_\varepsilon)c_i + \frac{\partial F}{\partial v^i}(\gamma_\varepsilon,\dot\gamma_\varepsilon)\dot c_i|_{\varepsilon=0}dt$$
$$= \int_a^b \sum_{i,j=1}^n \frac{\partial^2 F}{\partial x_i \partial x_j}(\gamma,\dot\gamma)c_i c_j dt \tag{I}$$
$$+2\int_a^b \sum_{i,j=1}^n \frac{\partial^2 F}{\partial x_i \partial v_j}(\gamma,\dot\gamma)c_i \dot c_j dt \tag{II}$$
$$+\int_a^b \sum_{i,j=1}^n \frac{\partial^2 F}{\partial v^i \partial v_j}(\gamma,\dot\gamma)\dot c_i \dot c_j dt \tag{III}$$

We estimate There exist constants $K_I, K_{II} > 0$ such that

$$|(\mathrm{I})| = \Big|\int_a^b \underbrace{\sum_{i,j=1}^n \frac{\partial^2 F}{\partial x_i \partial x_j}(\gamma,\dot\gamma)}_{=:\partial_x^2 F} c_i c_j dt\Big| \le \int_a^b |\langle c, \partial_x^2 F c\rangle| dt \le K_I |c|^2_{L^2([a,b])},$$

for some constant $K_I > 0$. By the same argument and the Hölder inequality

$$|(\mathrm{II})| \le K_{II}\int_a^b |\langle c, \dot c\rangle| dt \le K_{II}|c|_{L^2([a,b])}|\dot c|_{L^2([a,b])},$$

for another constant $K_{II} > 0$. By (4.2), there exists (III) > 0 such that

$$|(\mathrm{III})| \ge K_{III}|\dot c|^2_{L^2([a,b])}$$

Altogether this implies

$$\frac{d^2\mathcal{A}_\varepsilon}{d\varepsilon^2}(0) \geq K_{III}|\dot{c}|^2_{L^2([a,b])} - 2K_{II}|c|_{L^2([a,b])}|\dot{c}|_{L^2([a,b])} - K_I|c|^2_{L^2([a,b])}$$

By the Wirtinger inequality, c.f. (4.3), if b is very close to a, this expression is positive. □

4.5 Lemma (Wirtinger inequality)**.** For any $f \in \mathcal{C}^1([a,b])$ satisfying $f(a) = f(b) = 0$, we have

$$\int_a^b |f'(t)|dt \geq \frac{\pi^2}{(b-a)^2}\int_a^b |f(t)|^2 dt \tag{4.3}$$

4.2 Legendre Transform

4.6 Definition (Hessian)**.** Let V be an n-dimensional real vector space, $e_1, \ldots, e_n$ be a basis of V and $v_1, \ldots, v_n$ be the associated coordinates. For any function $F \in \mathcal{C}^\infty(V, \mathbb{R})$ and $p \in V$, $u = u^i e_i \in T_pV = V$ we set

$$d^2F|_p(u) := \sum_{i=1}^n \frac{\partial^2 F}{\partial v^i \partial v^j}(p)u_i u_j = \frac{d^2}{dt^2}F(p+tu)|_{t=0}.$$

The associated quadratic form d^2F on V is the *Hessian of F*.

4.7 Definition. A function $F \in \mathcal{C}^\infty(V, \mathbb{R})$ is *strictly convex*, if

$$\boxed{\forall p \in V : d^2F_p > 0,}$$

i.e. the quadratic form d^2F_p is positive definite.

4.8 Lemma. Let $F \in \mathcal{C}^\infty(V, \mathbb{R})$ be strictly convex. Then the following are equivalent.

(i) F has a critical point.

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

(iii) F has a unique critical point.

(iv) F has a global minimum.

(v) F is proper, i.e. $\lim_{|p|\to\infty} F(p) = \infty$.

Proof. □

4.9 Definition (stable)**.** A strictly convex function $F \in \mathcal{C}^\infty(M, \mathbb{R})$ is *stable*, if one (hence all) of the conditions in Lemma 4.8 is satisfied.

4.10 Example. For any $a \in \mathbb{R}$, the function $\mathbb{R} \to \mathbb{R}$, $x \mapsto e^x + ax$, is strictly convex, but it is stable only for $a < 0$ (see Figure 1).

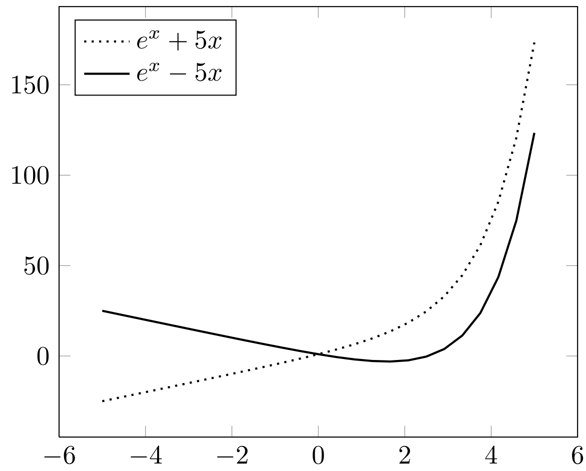

Figure 1: The function $x \mapsto e^x + ax$ is stable only for $a < 0$.

4.11 Definition (stability set)**.** Let $F \in \mathcal{C}^\infty(V, \mathbb{R})$ and $l \in V^*$. Define

$$\begin{aligned} F_l : V &\to \mathbb{R} \\ v &\mapsto F(v) - l(v) \end{aligned}$$

Clearly $d^2F = d^2F_l$, so in particular, F is strictly convex if and only if F_l is. If F is strictly convex, then

$$S_F := \{l \in V^* \mid F_l \text{ is stable}\}.$$

is the *stability set.*

4.12 Definition (Legendre transform)**.** The *Legendre transform* associated to a map $F \in \mathcal{C}^\infty(V, \mathbb{R})$ is

$$\begin{aligned} L_F : V &\to V^* \\ p &\mapsto dF_p \in T_p^*V = V^* \end{aligned}$$

4.13 Theorem. Let F be strictly convex. Then

$$L_F : V \to S_F$$

is a diffeomorphism.

4.14 Definition (dual function)**.** For any strictly convex $F \in \mathcal{C}^\infty(V, \mathbb{R})$ the function

$$\begin{aligned} F^* : S_F &\to \mathbb{R} \\ l &\mapsto -\min\nolimits_{p \in V} F_l(p) \end{aligned}$$

is the *dual function*.

4.15 Remark. In particular, we have

$$\forall v \in V : F^*(dF_v) = dF|_v - F(v) \tag{4.4}$$

4.16 Theorem.

$$L_F^{-1} = L_{F^*} \tag{4.5}$$

Proof. □

4.17 Theorem. Let M be a smooth manifold, $F \in \mathcal{C}^\infty(TM, \mathbb{R})$ be a function such that for all $p \in M$, $F_p := F|_{T_pM}$ is strictly convex and $S_{F_p} = T_p^*M$. Define the associated Legendre transform and Hamiltonian by

$$\begin{aligned} L : TM \to T^*M, \qquad & L|_{T_pM} := L_{F_p} \\ H : T^*M \to \mathbb{R}, \qquad & H|_{T_p^*M} := F_p^* \end{aligned}$$

A curve $\gamma : [a, b] \to M$ satisfies the Euler-Lagrange equations (EL) if and only if $\gamma^* := L \circ \dot\gamma$ is an integral curve for $X_H \in \mathcal{T}(T^*M)$, where T^*M carries the canonical symplectic structure.

Proof. We choose coordinates on some neighbourhood U and obtain induced coordinates for TU and T^*U, which will be labeled by

$$\begin{aligned} &(x_1, \dots, x_n), && \text{on } U, \\ &(x_1, \dots, x_n, v_1, \dots, v_n), && \text{on } TU, \\ &(x_1, \dots, x_n, \xi_1, \dots, x_n), && \text{on } T^*U. \end{aligned}$$

Locally, the curves $\dot\gamma$ and γ^* can be written in these coordinates as

$$\begin{aligned} \dot\gamma &= (\gamma_x, \gamma_v) = (\gamma_x, \dot\gamma_x) \\ \gamma^* &= (\gamma_x^*, \gamma_\xi^*) = (\gamma_x, L(\gamma_x, \dot\gamma_x)) \end{aligned}$$

Recall that γ satisfies the Euler-Lagrange equation if and only if

$$\frac{\partial F}{\partial x}(\gamma_x, \gamma_v) = \frac{d}{dt}\frac{\partial F}{\partial v}(\gamma_x, \gamma_v) \tag{4.6}$$

and γ^* is an integral curve of X_H if and only if it satisfies the Hamilton equations

$$\dot\gamma_x^* = \frac{\partial H}{\partial \xi}(\gamma_x^*, \gamma_\xi^*), \tag{4.7}$$

$$\dot\gamma_\xi^* = -\frac{\partial H}{\partial x}(\gamma_x^*, \gamma_\xi^*). \tag{4.8}$$

Let $t_0 \in [a, b]$ such that $\gamma(t_0) \in U$, set $(x_0, v_0) = \dot\gamma(t_0)$ and $(x_0, \xi_0) = \gamma^*(t_0)$. Since

$$\frac{\partial H}{\partial \xi}(x_0, \xi_0) = \frac{\partial F_x^*}{\partial \xi}(\xi_0) = L_{F_x^*}(\xi_0) \overset{(4.5)}{=} L_{F_x}^{-1}(\xi_0), \tag{4.9}$$

we conclude from the definition of γ^*

$$(4.7) \iff \dot\gamma_x^* \overset{(4.9)}{=} L_{F_x}^{-1}(\gamma_\xi^*) \iff \gamma_\xi^* = L_{F_x}(\dot\gamma_x^*) = L_{F_x}(\gamma_v),$$

so (4.7) holds by definition of γ^*. Consequently, we have to show that the Euler-Lagrange equation (4.6) is equivalent to the second Hamilton equation (4.8) in this case. To that end, we recall that for all (x,v), $\xi := L(x,v)$

$$H(x, L(x,v)) = F_x^*(\xi) \overset{(4.4)}{=} \xi v - F(x,v) = L(x,v)v - F(x,v).$$

If we differentiate this equation with respect to x, we obtain

$$\frac{\partial H}{\partial x}|_{(x,\xi)} + \frac{\partial H}{\partial \xi}|_{(x,\xi)} \frac{\partial L}{\partial x}|_{(x,v)} = \frac{\partial L}{\partial x}|_{(x,v)} v - \frac{\partial F}{\partial x}|_{(x,v)} \tag{4.10}$$

Using the fact that if $\xi = L(x,v)$, we obtain

$$\frac{\partial H}{\partial \xi}|_{(x,\xi)} = \frac{\partial F_x^*}{\partial \xi}(\xi) = L_{F_x^*}(\xi) = L_{F_x^{-1}}(F_x(v)) = v,$$

thus we can simplify (4.10) to

$$\frac{\partial H}{\partial x}|_{(x,\xi)} = -\frac{\partial F}{\partial x}|_{(x,v)}. \tag{4.11}$$

In particular if γ satisfies the Euler-Lagrange Equation, then by

$$\begin{aligned}
\dot{\gamma}_\xi^*(t_0) &= \frac{d}{dt} L_{F_{\gamma_x(t)}}(\dot{\gamma}_x(t))|_{t=t_0} \\
&= \frac{d}{dt}\frac{\partial F}{\partial v}(\gamma_x(t), \dot{\gamma}_x(t))|_{t=t_0} \\
&\overset{(4.6)}{=} \frac{\partial F}{\partial x}(\gamma_x(t), \gamma_v(t))|_{t=t_0} \\
&\overset{(4.11)}{=} -\frac{\partial H}{\partial x}|_{(x,\xi)}(\gamma_x^*(t), \gamma_\xi^*(t))
\end{aligned}$$

γ^* satisfies the second Hamilton equation. Conversely, if γ^* satisfies the second Hamilton equation, then by the same reasoning

$$\frac{\partial F}{\partial x}(\gamma_x, \gamma_v) \overset{(4.11)}{=} -\frac{\partial H}{\partial x}(\gamma_x^*, \gamma_\xi^*) \overset{(4.8)}{=} \dot{\gamma}_\xi^*(t_0) = \frac{d}{dt}\frac{\partial F}{\partial v}(\gamma_x(t), \dot{\gamma}_x(t))|_{t=t_0}$$

γ satisfies the Euler Lagrange equation. □

A Flows

A.1 Definition.

(i) A *flow domain* is an open subset $\mathcal{D} \subset \mathbb{R} \times M$ such that for each $p \in M$ the set $\mathcal{D}^{(p)} := \{t \in \mathbb{R} \mid (t,p) \in \mathcal{D}\}$ is an open interval containing 0.

(ii) A *flow* is a smooth map $\theta : \mathcal{D} \to M$, where $\mathcal{D}$ is a flow domain, such that for any $p \in M$ $\theta(0,p) = p$, and for any $s \in \mathcal{D}^{(p)}$ and any $t \in \mathcal{D}^{\theta(s,p)}$ such that $s + t \in \mathcal{D}^{(p)}$ we have $\theta(t, \theta(s,p)) = \theta(s+t, p)$. In this case, we set

$$\theta_t(p) := \theta(t,p) =: \theta^{(p)}(t).$$

(iii) For any flow $\theta : \mathcal{D} \to M$ the field $X \in \mathcal{T}(M)$ defined by

$$\forall p \in M : X_p = \tfrac{d}{dt}\theta^{(p)}(t)|_{t=0}$$

is the *infinitesimal generator of* X.

(iv) A flow $\theta : \mathcal{D} \to M$ is *maximal*, if it admits no extension to a larger flow domain.

A.2 Theorem (fundamental theorem on flows)**.** Let $X \in \mathcal{T}(M)$ be any smooth vector field. There exists a unique maximal flow $\theta : \mathcal{D} \to M$ whose infinitesimal generator is X. This flow has the following properties:

(i) For any $p \in M$m the curve $\theta^{(p)} : \mathcal{D}^{(p)} \to M$ is the unique maximal integral curve of X starting at $p \in M$.

(ii) If $s \in \mathcal{D}^{(p)}$, then $\mathcal{D}^{(\theta(s,p))} = \mathcal{D}^{(p)} - s$.

(iii) For each $t \in \mathbb{R}$, the set $M_t := \{p \in M \mid (t,p) \in \mathcal{D}\}$ is open in M and $\theta_t : M_t \to M_t$ is a diffeomorphism satisfying $\theta_t^{-1} = \theta_{-t}$.

(iv) For any $(t,p) \in \mathcal{D}$, $(\theta_t)_* X_p = X_{\theta_t(p)}$.

We say θ is *the flow of* X.

A.3 Definition (complete vector field)**.** A vector field $X \in \mathcal{T}(M)$ is *complete*, if its maximal flow $\theta : \mathcal{D} \to M$ satisfies $\mathcal{D} = \mathbb{R} \times M$.

A.4 Definition (Lie derivative)**.** Let $X \in \mathcal{T}(M)$ and θ be the flow of X. For any tensor field $\tau \in \mathcal{T}^k(M)$ define

$$\forall p \in M : (\mathcal{L}_X \tau)|_p := \tfrac{d}{dt}(\theta_t^* \tau)|_{t=0} = \lim_{t \to 0} \tfrac{\theta_t^*(\tau_{\theta_t(p)}) - \tau_p}{t}.$$

Then $\mathcal{L}_X \tau \in \mathcal{T}^k(M)$ is the *Lie derivative of* τ.

A.5 Definition (interior multiplication)**.** For any $\tau \in \Omega^k(M)$ and $X \in \mathcal{T}(M)$, let $\iota_X \omega \in \Omega^{k-1}(M)$ be defined by

$$\forall Y_1, \ldots, Y_{k-1} \in \mathcal{T}(M) : \iota_X \omega(Y_1, \ldots, Y_k) := \omega(X, Y_1, \ldots, Y_{k-1}).$$

The map $\iota_X : \Omega(M) \to \Omega(M)$ is called *interior multiplication with* X.

A.6 Theorem (properties of interior multiplication)**.** Let M be a smooth manifold, $X \in \mathcal{T}(M)$ and $\alpha \in \mathcal{T}^k(M)$.

(i) For any $X \in \mathcal{T}(M)$ and any *Cartan's formula* holds:

$$\boxed{\mathcal{L}_X \alpha = \iota_X d\alpha + d\iota_X \alpha.} \tag{A.1}$$

(ii) Interior multiplication is an *anti-derivation*, i.e. for any $\beta \in \Omega^l(M)$

$$\iota_X(\alpha \wedge \beta) = \iota_X(\alpha) \wedge \beta + (-1)^l \alpha \wedge \iota_X(\beta). \tag{A.2}$$

(iii) Interior multiplication satisfies

$$\iota_{[X,Y]}\alpha = \mathcal{L}_X \iota_Y \alpha - \iota_Y \mathcal{L}_X \alpha \tag{A.3}$$

References

[1] John M. Lee. *Introduction to Smooth Manifolds*. Springer, 2002.
[2] Ana Cannas da Silva. *Lectures on Symplectic Geometry*. Springer, 2008.

Index