Symplectic Geometry

Kai Cieliebak

October 29, 2010

Contents

Part I

Classical Mechanics

Chapter 1

Newtonian mechanics

1.1 Newton's law

In classical mechanics, the motion of a particle in a force field is governed by *Newton's law*

$$\mathbf{F} = m\mathbf{a}.$$

Here $\mathbf{r}(t) \in \mathbb{R}^3$ is the position of the particle at time t, $\mathbf{v} = \dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt}$ is its velocity and $\mathbf{a} = \ddot{\mathbf{r}} = \frac{d\mathbf{v}}{dt}$ its acceleration, m is the particle's mass, and $\mathbf{F} = \mathbf{F}(t, \mathbf{r}, \dot{\mathbf{r}})$ is the force acting upon the particle. Since Newton's equations are of second order, the motion of the particle is determined uniquely by its initial position and velocity.

Of particular importance are *conservative forces* which can be written as a gradient $\mathbf{F} = -\nabla V$ for a *potential* $V = V(t, \mathbf{r})$. If $\mathbf{F}$ is conservative with time-independent potential V the *total energy*

$$E := \frac{1}{2}m\|\mathbf{v}\|^2 + V(\mathbf{r})$$

is conserved:

$$\frac{dE}{dt} = m\dot{\mathbf{r}} \cdot \ddot{\mathbf{r}} + \nabla V(\mathbf{r}) \cdot \dot{\mathbf{r}} = 0.$$

The term $T := \frac{1}{2}m\|\mathbf{v}\|^2$ is called *kinetic energy*, and $V(\mathbf{r})$ *potential energy.*

Basic forces

The basic forces we encounter in classical mechanics are:

Gravitation. The gravitational force exerted by a particle of mass m_2 at position $\mathbf{r}_2$ upon a particle of mass m_1 at position $\mathbf{r}_1$ equals

$$\mathbf{F}_{grav} = -\gamma m_1 m_2 \frac{\mathbf{r}_1 - \mathbf{r}_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3},$$

where $\gamma = 6.67 \times 10^{-11} \frac{m^3}{kg \times s^2}$ is the gravitational constant. This is called *Newton's law of gravitation.* The gravitational force is conservative with potential

$$V_{grav} = -\frac{\gamma m_1 m_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|}.$$

Electricity. The electric force exerted by a particle of charge e_2 at position $\mathbf{r}_2$ upon a particle of charge e_1 at position $\mathbf{r}_1$ equals

$$\mathbf{F}_{el} = \frac{e_1 e_2}{4\pi\varepsilon_0} \frac{\mathbf{r}_1 - \mathbf{r}_2}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3},$$

where $\varepsilon_0 = \frac{10^7}{4\pi c^2} \frac{C^2}{kg \cdot m} = 8.85 \times 10^{-12} \frac{C^2 \times s^2}{kg^2 \times m^3}$ is the vacuum dielectric constant (in charge units $C = Coulomb$, $c \approx 3 \cdot 10^8 \frac{m}{s}$ velocity of light). This is called *Coulomb's law.* The electric force is conservative with potential

$$V_{el} = \frac{e_1 e_2}{4\pi\varepsilon_0 \|\mathbf{r}_1 - \mathbf{r}_2\|}.$$

Magnetism. The magnetic force exerted by a particle of charge e_2 and velocity $\mathbf{v}_2$ at position $\mathbf{r}_2$ upon a particle of charge e_1 and velocity $\mathbf{v}_1$ at position $\mathbf{r}_1$ equals

$$\mathbf{F}_{mag} = \frac{\mu_0 e_1 e_2}{4\pi} \frac{\mathbf{v}_1 \times \big(\mathbf{v}_2 \times (\mathbf{r}_1 - \mathbf{r}_2)\big)}{\|\mathbf{r}_1 - \mathbf{r}_2\|^3},$$

where $\mu_0 = 4\pi \cdot 10^{-7} \frac{kg \cdot m}{C^2} = 1.26 \times 10^{-6} \frac{kg \times m}{C^2}$ is the vacuum magnetic permeability. The magnetic force is not conservative. It is usually written as the *Lorentz force*

$$\mathbf{F}_{mag} = e_1 \mathbf{v}_1 \times \mathbf{B}(\mathbf{r}_1),$$

where the *magnetic induction* $\mathbf{B}$ is given by the *Biot-Savart law*

$$\mathbf{B}(\mathbf{r}_1) = \frac{\mu_0 e_2}{4\pi} \mathbf{v}_2 \times (\mathbf{r}_1 - \mathbf{r}_2).$$

$\mathbf{B}$ is also not conservative, but since $\operatorname{div}(\mathbf{a} \times \mathbf{r}) = 0$ for every constant vector $\mathbf{a}$, $\operatorname{div} \mathbf{B} = 0$. So on any simply connected region

$$\mathbf{B} = \operatorname{curl} \mathbf{A}$$

for a *vector potential* $\mathbf{A}$.

Problem 1.1. With which velocity must a rocket take off from the surface of the earth to escape the earth's gravitational field if no further acceleration takes place after take-off? (earth radius = 6700 km).

Systems of particles

Newton's law generalizes immediately to a system of N particles: $\mathbf{F}_i = m_i \ddot{\mathbf{r}}_i$, $i = 1, \ldots, N$, where the force $\mathbf{F}_i$ upon the ith particle may depend on the positions and velocities of all the particles. With $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$, $\mathbf{F} = (\mathbf{F}_1, \dots, \mathbf{F}_N)$ and $M = diag(m_1, \dots, m_N)$ this can be written as

$$\mathbf{F} = M\ddot{\mathbf{r}}.$$

More generally, we call a *Newtonian system* the equation $\mathbf{F} = M\ddot{\mathbf{r}}$ with $\mathbf{r} \in \mathbb{R}^n$ ($n = 3N$ for an N-particle system) and M a positive definite matrix.

1.2 Constraints

Consider a Newtonian system $\mathbf{F} = M\ddot{\mathbf{r}}$ on $\mathbb{R}^n$. Suppose the motion is subject to $n-k$ *holonomic constraints* $f_{k+1}(\mathbf{r}) = \cdots = f_n(\mathbf{r}) = 0$. We assume that the ∇f_i are linearly independent at every point satisfying the constraints, so that

$$Q = \{f_{k+1}(\mathbf{r}) = \cdots = f_n(\mathbf{r}) = 0\}$$

is a k-dimensional submanifold of $\mathbb{R}^n$. Newton's equations continue to hold in the form

$$M\ddot{\mathbf{r}} = \mathbf{F} + \mathbf{F}_{constr},$$

where $\mathbf{F}_{constr}$ is the constraint force that keeps the particles on Q. We assume that no work is done by the constraint force, i.e. no energy is lost or gained by keeping the particles on Q. This is expressed by

d'Alembert's principle: The constraint force is perpendicular to the submanifold Q.

Problem 1.2. (cf. [1], Section 21). Derive d'Alembert's principle by introducing a potential with a sharp minimum along Q and taking the limit as the minimum becomes sharper and sharper.

D'Alembert's principle together with the constraints and Newton's equations determine the motion as follows: Write

$$\mathbf{F}_{constr} = \sum_{i=k+1}^{n} \lambda_i \nabla f_i$$

with the (yet unknown) *Lagrange multipliers* $\lambda_i = \lambda_i(t, \mathbf{r}, \dot{\mathbf{r}})$. Taking two time derivatives of $f_j(\mathbf{r}) = 0$ yields $0 = \langle \nabla f_j(\mathbf{r}), \dot{\mathbf{r}} \rangle$ and

$$\begin{aligned} 0 &= \langle D^2 f_j \cdot \dot{\mathbf{r}}, \dot{\mathbf{r}} \rangle + \langle \nabla f_j, \ddot{\mathbf{r}} \rangle \\ &= \langle D^2 f_j \cdot \dot{\mathbf{r}}, \dot{\mathbf{r}} \rangle + \langle M^{-1}\mathbf{F}, \nabla f_j \rangle + \sum_{i=k+1}^{n} \lambda_i \langle M^{-1} \nabla f_i, \nabla f_j \rangle. \end{aligned}$$

Since M is positive definite and the ∇f_i linearly independent, the matrix $\langle M^{-1}\nabla f_i, \nabla f_j \rangle$ is positive definite. So the equations can be solved uniquely for $\lambda_{k+1}, \dots, \lambda_n$, and plugging the λ_i into Newton's equations determines the motion of the system.

Example 1.1. (centripetal force). Suppose that the motion is constrained to the sphere $Q = \{\|\mathbf{r}\| = R\}$, and there are no external forces. Taking two time derivatives yields $0 = \langle \mathbf{r}, \dot{\mathbf{r}} \rangle$ and $0 = v^2 + \langle \mathbf{r}, \ddot{\mathbf{r}} \rangle$, where $v = \|\dot{\mathbf{r}}\|$. Since $\ddot{\mathbf{r}} \perp Q$, this determines the *centripetal force*

$$\mathbf{F}_{constr} = m\ddot{\mathbf{r}} = -\frac{mv^2}{R}\frac{\mathbf{r}}{R}.$$

Problem 1.3. A particle slides down (without friction) a slope and at the bottom enters a circular looping of radius R. From which height h must the particle start (with initial velocity zero) so that it stays on track in the looping?

Problem 1.4. Show that for a Newtonian system constrained to a submanifold $Q \subset \mathbb{R}^n$ with mass matrix $M = m\mathbb{1}$ the constraint force is given by

$$\mathbf{F}_{constr} = m\, II(\dot{\mathbf{r}}, \dot{\mathbf{r}}) - \mathbf{F}^{\perp},$$

where II is the second fundamental form of $Q \subset \mathbb{R}^n$, and $\mathbf{F}^{\perp}$ is the component of the external force perpendicular to Q.

Chapter 2

Lagrangian mechanics

2.1 Hamilton's variational principle and Lagrange's equations

D'Alembert's principle and the method of Lagrange multipliers provide a way to treat systems with constraints, but a relatively clumsy one because it involves $2n - k$ unknowns $(\mathbf{r}, \lambda_{k+1}, \dots, \lambda_n)$. To find a simpler method, let $\mathbf{r} = \mathbf{r}(q_1, \dots, q_k)$ be a parametrization of Q (which always exists at least locally). Can we reformulate Newton's equations in terms of the k unknowns $(q_1, \dots, q_k)$?

Problem 2.1. Convince yourself that the desired equations are *not* $\ddot{q} = \mathbf{F}^{\|}$, where $\mathbf{F}^{\|}$ is the component of $\mathbf{F}$ tangent to Q expressed in terms of q.

Euler-Lagrange equations

Consider the following variational problem: Given a smooth *Lagrange function*

$$L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

find extremals of

$$\mathcal{L}(\mathbf{r}) = \int_a^b L(t, \mathbf{r}, \dot{\mathbf{r}}) dt$$

among paths $\mathbf{r} : [a, b] \to \mathbb{R}^n$ with fixed endpoints $\mathbf{r}(a) = \mathbf{r}_a$, $\mathbf{r}(b) = \mathbf{r}_b$.
We compute the variation of $\mathcal{L}$ in the direction $\delta\mathbf{r} : [a, b] \to \mathbb{R}^n$, $\delta\mathbf{r}(a) = \delta\mathbf{r}(b) =$

0:

$$
\begin{aligned}
d\mathcal{L}(\mathbf{r}) \cdot \delta\mathbf{r} &= \frac{d}{ds}\Big|_{s=0} \int_a^b L(t, \mathbf{r} + s\delta\mathbf{r}, \dot{\mathbf{r}} + s\dot{\delta\mathbf{r}})dt \\
&= \int_a^b \Big(\frac{\partial L}{\partial \mathbf{r}} \cdot \delta\mathbf{r} + \frac{\partial L}{\partial \dot{\mathbf{r}}} \cdot \dot{\delta\mathbf{r}}\Big)dt \\
&= \int_a^b \Big(\frac{\partial L}{\partial \mathbf{r}} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{r}}}\Big) \cdot \dot{\delta\mathbf{r}}\, dt. \qquad (2.1)
\end{aligned}
$$

So the variation vanishes in all directions $\delta\mathbf{r}$ if and only if $\mathbf{r}$ satisfies the *Euler-Lagrange equations*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{r}}} - \frac{\partial L}{\partial \mathbf{r}} = 0.$$

Hamilton's variational principle

Can we find a function L such that Newton's equations are the Euler-Lagrange equations for L? That is, we look for L such that $\frac{\partial L}{\partial \dot{\mathbf{r}}} = M\dot{\mathbf{r}}$ and $\frac{\partial L}{\partial \mathbf{r}} = -\nabla V$. Obviously the function $L = \frac{1}{2}\langle \dot{\mathbf{r}}, M\dot{\mathbf{r}}\rangle - V(t, \mathbf{r})$ does the job. So we have shown:

Newton's equations are equivalent to the Euler-Lagrange equations for the Lagrange function

$$L(t, \mathbf{r}, \dot{\mathbf{r}}) = \frac{1}{2}\langle \dot{\mathbf{r}}, M\dot{\mathbf{r}}\rangle - V(t, \mathbf{r}) = T - V.$$

Now consider a system constrained to $Q = \{f_{k+1} = \cdots = f_n = 0\}$. $\mathcal{L}$ is extremal for variations among paths $\mathbf{r} : [a, b] \to Q \subset \mathbb{R}^n$ iff the expression (2.1) vanishes for all variations $\delta\mathbf{r}$ tangent to Q, which is the case if and only if $\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{r}}} - \frac{\partial L}{\partial \mathbf{r}}$ is everywhere orthogonal to Q. But for $L = T - V$, $\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{r}}} - \frac{\partial L}{\partial \mathbf{r}}$ is the constraint force, so the orthogonality condition is just d'Alembert's principle! So we have shown:

Proposition 2.1. *(Hamilton's variational principle). The path* $\mathbf{r} : [a, b] \to Q$ *satisfies Newton's equations with holonomic constraints* $Q = \{f_{k+1} = \cdots = f_n = 0\} \subset \mathbb{R}^n$ *if and only if it is extremal for* $\int_a^b L(t, \mathbf{r}, \dot{\mathbf{r}})dt$ *among variations in* Q *with fixed endpoints, where* $L = T - V$ *is the Lagrangian.*

Corollary 2.2. *(Lagrange's equations). In a (local) parametrization* $\mathbf{r} = \mathbf{r}(q)$ *of* Q*, Newton's equations are equivalent to* Lagrange's equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0,$$

where the Lagrangian $L = T - V$ *is expressed in terms of* $(t, q, \dot{q})$*.*

Proof. Variations in Q are unconstrained in the parameters q, so a path is extremal if and only if it satisfies the Euler-Lagrange equations in the variables $q = (q_1, \ldots, q_k)$. □

Note that Lagrange's equations hold in whichever local parametrization we choose! Corollary 2.2 thus provides a very simple recipee to derive Newton's equations for constrained systems: Express the Lagrangian in terms of a convenient parametrization and work out Lagrange's equations. The following example illustrates this procedure:

Example 2.3. (spherical pendulum).

Consider a point of mass m at the end of a rigid rod of length l which is fixed at the other end, so the motion of the mass point is constrained to the sphere $Q = \{\|\mathbf{r}\| = l\}$. The pendulum is subject to a constant gravitational field of strength g in the negative z-direction.

The Lagrangian for $\mathbf{r} = (x, y, z)$ is given by

$$L = \frac{1}{2}m\|\dot{\mathbf{r}}\|^2 - mgz.$$

In spherical coordinates

$$\left\{\begin{aligned} x &= l\sin\phi\cos\theta \\ y &= l\sin\phi\sin\theta \\ z &= l\cos\phi \end{aligned}\right\}$$

the Lagrangian becomes

$$L = \frac{1}{2}ml^2(\dot{\phi}^2 + \sin^2\phi\dot{\theta}^2) - mgl\cos\phi,$$

and Lagrange's equations for (ϕ, θ) are

$$\frac{d}{dt}(ml^2\dot{\phi}) - ml^2\sin\phi\cos\phi\dot{\theta}^2 - mgl\sin\phi = 0$$
$$\frac{d}{dt}(ml^2\sin^2\phi\dot{\theta}) = 0.$$

The second equation implies $p_\theta := ml^2\sin^2\phi\dot{\theta} = \text{const}$. Inserting this into the first equation yields

$$\ddot{\phi} - \frac{p_\theta^2}{m^2l^4}\frac{\cos\phi}{\sin^3\phi} - \frac{g}{l}\sin\phi = 0.$$

Problem 2.2. (planar pendulum).

The planar pendulum is a spherical pendulum which can only move in a plane, say the plane $\theta = 0$. Then $p_\theta = 0$ and Lagrange's equation in the angle ϕ becomes

$$\ddot{\phi} - \frac{g}{l}\sin\phi = 0.$$

Draw the curves of constant energy $E = \frac{1}{2}\dot{\phi}^2 + \frac{g}{l}\cos\phi = \text{const}$ on the $(\phi, \dot{\phi})$ cylinder, and discuss the qualitative behaviour of the system (stable/unstable equilibria, contractible/noncontractible periodic orbits, homoclinic orbits).

Let us now discuss the motion of the spherical pendulum with $p_\theta \neq 0$ (the case $p_\theta = 0$ is the planar pendulum). It will be more convenient to use cylindrical coordinates

$$\left\{ \begin{array}{l} x = \rho\cos\theta \\ y = \rho\sin\theta \\ z = z \end{array} \right\}$$

with $\rho^2 + z^2 = l^2$. Set the mass to $m = 1$ since it cancels out anyway, and express the Lagrange function $L = T - V$ in terms of z and θ using $\dot\rho^2 = \frac{z^2}{\rho^2}\dot z^2$:

$$\begin{aligned} V &= gz, \\ T &= \frac{1}{2}(\dot\rho^2 + \rho^2\dot\theta^2 + \dot z^2) \\ &= \frac{1}{2}\Big(\frac{l^2\dot z^2}{l^2 - z^2} + (l^2 - z^2)\dot\theta^2\Big) \\ &= \frac{1}{2}\Big(\frac{l^2 - z^2}{l^2}p_z^2 + \frac{p_\theta^2}{l^2 - z^2}\Big), \end{aligned}$$

where

$$\begin{aligned} p_\theta &= \frac{\partial L}{\partial\dot\theta} = (l^2 - z^2)\dot\theta = \text{const}, \\ p_z &= \frac{\partial L}{\partial\dot z} = \frac{l^2\dot z}{l^2 - z^2}. \end{aligned}$$

The motion of the z-coordinate is described by the curves in the (z, p_z)-plane of constant total energy

$$E = \frac{1}{2}\Big(\frac{l^2 - z^2}{l^2}p_z^2 + \frac{p_\theta^2}{l^2 - z^2}\Big) + gz.$$

A point (z, p_z) is critical for E iff $p_z = 0$ and $V'_{eff}(z) = 0$, where

$$V_{eff}(z) = \frac{1}{2}\frac{p_\theta^2}{l^2 - z^2} + gz$$

is the *effective potential.* From

$$V'_{eff}(z) = \frac{p_\theta^2}{(l^2 - z^2)^2} + g$$

we see that $V'_{eff}(-l) = -\infty$, $V'_{eff}(0) = g > 0$ and $V''_{eff} > 0$. So for any p_θ there exists a unique critical point z_{crit} of V_{eff} in the interval $(-l, 0)$. It corresponds to a motion of the pendulum on the horizontal circle $z \equiv z_{crit}$.

Since $E \to \infty$ as $z \to \pm l$ or $p_z \to \pm\infty$, all other level curves of E are compact and therefore circles. The correspond to motions of the pendulum rotating around the z-axis with the z-coordinate oscillating between two values z_{min}

and z_{max}. Depending on the ratio between the frequencies of the rotation and the oscillation, the orbits are either closed, or dense in an annulus on the sphere $\{\|\mathbf{r}\| = l\}$.

The quantities p_θ and p_z are examples of *conjugate momenta* $p_i := \frac{\partial L}{\partial q_i}$ to a varible q_i. They have the obvious but important property that if L does not depend explicitly on q_i, then p_i is a constant of the motion.

2.2 Lagrangian systems on manifolds

The discussion in the previous section suggests a more general setting for Lagrangian systems. Let Q^n be a manifold of dimension n and

$$L : \mathbb{R} \times TQ \to \mathbb{R}$$

a smooth Lagrangian. Assign to every path $q : [a, b] \to Q$ its *action*

$$\mathcal{L}(q) := \int_a^b L\big(t, q(t), \dot{q}(t)\big) dt.$$

By definition, the path q is a solution of the Lagrangian system defined by L iff it is an extremal of $\mathcal{L}$ among variations with fixed endpoints. Equivalently, q satisfies Lagrange's equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0$$

in any local coordinates $(q_1, \dots, q_n)$.

Geodesics

A *natural Lagrangian* is one of the form

$$L(t, q, \dot{q}) = \frac{1}{2}\|\dot{q}\|^2 - V(t, q),$$

where $\|\ \|$ is the norm corresponding to a Riemannian metric $\langle\ ,\ \rangle$ on Q.

An important special case is *free motion* with $L = \frac{1}{2}\|\dot{q}\|^2$. Solutions of this Lagrangian systems are called *geodesics* of the Riemannian metric.

Problem 2.3. Show that this definition of geodesics coincides with the more customary one: geodesics are extremals of the length functional $\int_a^b \|\dot{q}\| dt$. However, the two concepts differ in the parametrizations of geodesics. How?

Problem 2.4. Let $g_{ij} = \langle \frac{\partial}{\partial q_i}, \frac{\partial}{\partial q_j} \rangle$ be the metric in local coordinates $(q_1, \dots, q_n)$ and g_{ij}^{-1} the inverse matrix. Show that Lagrange's equations of a natural Lagrangian system are

$$\ddot{q}_k + \sum_{i,j} \Gamma_{ij}^k \dot{q}_i \dot{q}_j + \sum_i g_{ik}^{-1} \frac{\partial V}{\partial q_i} = 0,$$

where Γ_{ij}^k are the Christoffel symbols

$$\Gamma_{ij}^k = \frac{1}{2}\sum_l g_{kl}^{-1}\Big(\frac{\partial g_{il}}{\partial q_j} + \frac{\partial g_{jl}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_l}\Big).$$

Problem 2.5. Show that geodesics on a submanifold $Q^k \subset \mathbb{R}^n$ with the induced metric are, up to parametrization, the curves on Q whose second derivative in $\mathbb{R}^n$ is everywhere perpendicular to Q. Conclude that the geodesics on a round sphere are the great circles.

Problem 2.6. Consider the ellipsoid $Q := \{\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} = 1\} \subset \mathbb{R}^3$, $0 < a < b < c$. Show that the intersections of Q with the 3 coordinate planes are geodesics. Describe the geodesics in a neighbourhood of these 3 geodesics. Are there other closed geodesics?

Conservation of energy

Conservation of energy is a consequence of the change-of-variable formula for integrals. Consider a change $t = t(\tau)$ of the time variable with $t(a) = a$ and $t(b) = b$. Denote derivatives with respect to t, τ by $\dot q = \frac{dq}{dt}$, $q' = \frac{dq}{d\tau}$. By the change-of-variables formula,

$$\int_a^b L\,dt = \int_a^b L\frac{dt}{d\tau}d\tau = \int_a^b \tilde L\,d\tau,$$

where

$$\tilde L(t, q, t', q') = L(t, q, \frac{q'}{t'})t'$$

is a Lagrange function on the *extended phase space* $T(\mathbb{R} \times Q)$. So the variational principles for L and $\tilde L$ agree, and the Euler-Lagrange equations for L are equivalent to those for $\tilde L$. In particular, the equation for $\tilde L$ with respect to the variable t reads

$$\begin{aligned}\frac{d}{d\tau}\frac{\partial \tilde L}{\partial t'} &= \frac{d}{d\tau}\Big(L + \langle\frac{\partial L}{\partial \dot q}, -\frac{1}{(t')^2}q'\rangle t'\Big)\\ &= \frac{d}{d\tau}\Big(L - \langle\frac{\partial L}{\partial \dot q}, \dot q\rangle\Big)\\ &= \frac{\partial \tilde L}{\partial t} = \frac{\partial L}{\partial t}t',\end{aligned}$$

or

$$\frac{d}{dt}\Big(L - \langle\frac{\partial L}{\partial \dot q}, \dot q\rangle\Big) = \frac{\partial L}{\partial t}.$$

So if L is time-independent, the *energy*

$$E := \sum_i \frac{\partial L}{\partial q_i}\dot q_i - L(q, \dot q)$$

is a constant of the motion. (This can also be directly verified from Lagrange's equations for L).

For an intrinsic expression of the energy, denote by $d^{fibre}L$ the *fibre derivative*

$$d^{fibre}_{(q,\dot q)}L \cdot v := \frac{d}{ds}\Big|_{s=0} L(q, \dot q + sv), \quad v \in T_qQ.$$

Then

$$E = d^{fibre}_{(q,\dot q)}L \cdot \dot q - L(q, \dot q).$$

For geodesic motion with $L = \frac{1}{2}\|\dot q\|^2$ the energy is $E = \frac{1}{2}\|\dot q\|^2$, so geodesics are parametrized with constant velocity.

2.3 Minimizing solutions

A solution of a Lagrangian system is called *(locally) minimizing* if it is a (local) minimum (not just an extremum) of $\mathcal{L}$ among curves with the same endpoints.

Proposition 2.4. *Suppose that L is strictly convex in the fibres in the following sense: For all (t, q) there exists an $\varepsilon > 0$ such that $\sum_{i,j} \frac{\partial^2 L}{\partial \dot q_i \partial \dot q_j} v_i v_j \geq \varepsilon \|v\|^2$ for all $v \in T_qQ$. Then given a solution $q : [a, b] \to Q$ of the Lagrangian system, $q_{[a',b']}$ is locally minimizing for any sufficiently small subinterval $[a', b'] \subset [a, b]$.*

Problem 2.7. Prove *Wirtinger's inequality*

$$\int_a^b \Big|\frac{df}{dt}\Big|^2 dt \geq \frac{\pi^2}{(b-a)^2} \int_a^b |f|^2 dt$$

for every $f \in C^1([a, b])$ with $f(a) = f(b) = 0$. *Hint:* Fourier series.

Proof. For $[a, b]$ sufficiently small we may assume that $q([a, b])$ is contained in a chart in $\mathbb{R}^n$. Let $\xi : [a, b] \to \mathbb{R}^n$ be a variation with $\xi(a) = \xi(b) = 0$. The second variation of $\mathcal{L}$ in direction ξ equals

$$\begin{aligned}
\frac{d^2}{ds^2}\Big|_{s=0} \mathcal{L}(q + s\xi) &= \int_a^b \sum_{i,j} \Big(\frac{\partial^2 L}{\partial \dot q_i \partial \dot q_j} \dot\xi_i \dot\xi_j + 2\frac{\partial^2 L}{\partial \dot q_i \partial q_j} \dot\xi_i \xi_j + \frac{\partial^2 L}{\partial q_i \partial q_j} \xi_i \xi_j \Big) dt \\
&\geq \int_a^b (\varepsilon \|\dot\xi\|^2 - c\|\dot\xi\|\,\|\xi\| - c\|\xi\|^2) dt \geq \int_a^b (\frac{\varepsilon}{2}\|\dot\xi\|^2 - C\|\xi\|^2) dt \\
&\geq (\frac{\varepsilon\pi^2}{2(b-a)^2} - C)\|\xi\|^2 dt > 0
\end{aligned}$$

for $(b - a)$ sufficiently small and $\xi \neq 0$. Here c, C are constants depending on $(q, \dot q)$, and we have used Wirtinger's inequality in the last line. $\square$

Corollary 2.5. *For every geodesic $q : [a, b] \to Q$ and sufficiently small subinterval $[a', b'] \subset [a, b]$, $q|_{[a',b']}$ is a local minimum of the length functional.*

Proof. Extremals of $\mathcal{L} = \frac{1}{2}\|\dot{q}\|^2$ are geodesics parametrized with $\|\dot{q}\| = \text{const}$. For an extremal q,

$$\begin{aligned} length(q) &= (b-a)\|\dot{q}\|, \\ \mathcal{L}(q) &= \frac{1}{2}(b-a)\|\dot{q}\|^2 = \frac{1}{2(b-a)} length(q)^2. \end{aligned}$$

So $\mathcal{L}$ is a monotone function of length on extremals. In particular, q is length minimizing iff it minimizes $\mathcal{L}$. □

Remark. In fact, geodesics are *absolute* minima of the length functional on small intervals (see, e.g.,[6]).

Problem 2.8. Let $L(q,\dot{q}) := \frac{1}{2}(|\dot{q}_1|^2 - |\dot{q}_2|^2)$ on $\mathbb{R}^2$. What are the extremals of $\mathcal{L}$? Are they unique for given endpoints? Are they locally/globally minimizing?

Existence of minimizers

Theorem 2.6. *Let $L = \frac{1}{2}\|\dot{q}\|^2 - V(t,q)$ be a natural Lagrangian on a compact manifold Q. Then, given $a < b$ and $q_a \neq q_b$ in Q, there exists a smooth minimizer $q : [a,b] \to Q$ of $\mathcal{L}$ with $q(a) = q_a$ and $q(b) = q_b$.*

Sketch of proof. I will use without proof some results from linear functional analysis. By Nash's Embedding Theorem, Q can be isometrically embedded into some $\mathbb{R}^N$. Let $\mathcal{H}$ be the completion of the linear space

$$\{u \in C^1([a,b], \mathbb{R}^N) \mid u(a) = u(b) = 0\}$$

with respect to the norm

$$\|u\|^2_{\mathcal{H}} := \int_a^b \|\dot{u}\|^2 dt.$$

Since on this space the norm $\| \ \|_{\mathcal{H}}$ is equivalent to the Sobolev norm $\int_a^b (\|u\|^2 + \|\dot{u}\|^2) dt$, $(\mathcal{H}, \| \ \|_{\mathcal{H}})$ is a Hilbert space.

Let $\inf \mathcal{L}$ be the infimum of $\mathcal{L}$ over all C^1-curves $q : [a,b] \to Q$ with $q(a) = q_a$ and $q(b) = q_b$. Let $q_n : [a,b] \to Q$ be a minimizing sequence with $\mathcal{L}(q_n) \to \inf \mathcal{L}$. Write $q_n = q_0 + u_n$ with $u_n \in \mathcal{H}$. Since $\int V(t,q_n)$ is bounded uniformly in n, it followew that $\|u_n\|_{\mathcal{H}}$ is bounded. By the Banach-Alaoglu Theorem there exists a subsequence, still denoted by (u_n), which converges weakly to some $u_\infty \in \mathcal{H}$. By the Arzela-Ascoli Theorem, we may further assume that $u_n \to u_\infty$ uniformly. Let $q_\infty := q_0 + u_\infty$.

By the uniform convergence, the image of q_∞ is contained in Q, and $q_\infty(a) = q_a$, $q_\infty(b) = q_b$. Moreover, $V(t, q_\infty) = \lim V(t, q_n)$.

By the weak convergence, $\|u_\infty\|_{\mathcal{H}} \leq \liminf \|u_n\|_{\mathcal{H}}$. This implies $\mathcal{L}(q_\infty) \leq \liminf \mathcal{L}(q_n) = \inf \mathcal{L}$. If the inequality were strict, then a smooth approximation of q_∞ would yield a contradiction to the definition of $\inf L$. So $\mathcal{L}(q_\infty) = \inf \mathcal{L}$. It follows that q_∞ satisfies Lagrange's equations in the weak sense and therefore is smooth. Thus q_∞ is a minimizing solution.

By the way, the fact that $\|u_n\|_{\mathcal{H}} \to \|u_\infty\|_{\mathcal{H}}$ implies that the subsequence u_n actually converges to u_∞ strongly in $\mathcal{H}$. $\square$

Corollary 2.7. *Any two points in a compact Riemannian manifold are connected by a length-minimizing geodesic.*

Problem 2.9. Show that Theorem 2.6 is not true in general for noncompact Q. Find some hypotheses for noncompact Q under which you can prove the theorem. *Remark:* For $V = 0$ such hypotheses are provided by the Hopf-Rinow Theorem (see [6]).

Problem 2.10. Modify the proof of Theorem 2.6 to show: If the natural Lagrange function L on the compact manifold Q is 1-periodic in t, then in every nontrivial free homotopy class of loops on Q there exists a minimizing solution.

What happens to this proof applied to the trivial free homotopy class of loops? Find an example where there exist no contractible solutions. *Remark:* There always exists a 1-periodic solution, even if Q is simply connected (see [5]; for $V = 0$ this is the Lusternik-Fet Theorem). However, in the simply connected case, this solution will in general not be minimizing, so different arguments are needed.

2.4 Noether's Theorem

Noether's Theorem formalizes the long-observed principle that symmetries of a Lagrangian system lead to conservation laws.

Theorem 2.8 (Noether's Theorem)**.** *Suppose that a 1-parameter family of diffeomorphisms $\phi_s : Q \to Q$, $s \in (-\varepsilon, \varepsilon)$, leaves the Lagrangian $L : \mathbb{R} \times TQ \to \mathbb{R}$ invariant: $L(t, \phi_{s*}\cdot) = L(t, \cdot)$. Then the quantity*

$$\mu(q, \dot q) := \langle \frac{\partial L}{\partial \dot q}, \frac{d}{ds}\big|_{s=0} \phi_s(q) \rangle$$

is a constant of the motion.

Notation. Any map $\phi : Q \to Q$ lifts to a map $\phi_* : TQ \to TQ$, $(q, \dot q) \mapsto \big(\phi(q), T_q\phi \cdot \dot q\big)$. Note that the quantity μ is intrinsically defined: $\frac{\partial L}{\partial \dot q}$ is the fibre derivative of L, and $\frac{d}{ds}\big|_{s=0}\phi_s(q)$ is a vector field on Q.

Proof. From $L(t, \phi_{s*}\cdot) = L(t, \cdot)$ we obtain in local coordinates:

$$\begin{aligned} 0 &= \frac{d}{ds}\big|_{s=0} L\big(t, \phi_{s*}(q, \dot q)\big) \\ &= \langle \frac{\partial L}{\partial q}, \frac{d}{ds}\big|_{s=0} \phi_s(q) \rangle + \langle \frac{\partial L}{\partial \dot q}, \frac{d}{ds}\big|_{s=0} T_q\phi_s \cdot \dot q \rangle. \end{aligned}$$

This identity and Lagrange's equations yield

$$\begin{aligned}\frac{d\mu}{dt} &= \langle \frac{d}{dt}\frac{\partial L}{\partial \dot q}, \frac{d}{ds}\big|_{s=0}\phi_s(q)\rangle + \langle \frac{\partial L}{\partial \dot q}, \frac{d}{ds}\big|_{s=0}T_q\phi_s \cdot \dot q\rangle \\ &= \langle \frac{d}{dt}\frac{\partial L}{\partial \dot q}, \frac{d}{ds}\big|_{s=0}\phi_s(q)\rangle - \langle \frac{\partial L}{\partial q}, \frac{d}{ds}\big|_{s=0}\phi_s(q)\rangle \\ &= 0.\end{aligned}$$

□

Example 2.9. (linear momentum).

Suppose that $L = \frac{1}{2}\langle M\dot{\mathbf{r}}, \dot{\mathbf{r}}\rangle - V(t, \mathbf{r})$ on $T\mathbb{R}^n$, and V is invariant under translations in the direction $\mathbf{e} \in \mathbb{R}^n$. Then the conserved quantity provided by Noether's Theorem is

$$\mu = \langle M\dot{\mathbf{r}}, \mathbf{e}\rangle,$$

the component in direction $\mathbf{e}$ of the *linear momentum* $M\dot{\mathbf{r}}$.

For example, this applies to a system of N particles in $\mathbb{R}^3$ exerting forces upon each other that depend only on their distances:

$$L = \frac{1}{2}\sum m_i\|\dot{\mathbf{r}}_i\|^2 - V(t, \mathbf{r}_1, \ldots, \mathbf{r}_N),$$

where V depends only on t and the $\|\mathbf{r}_i - \mathbf{r}_j\|$. Then V is invariant under translations in all directions $(\mathbf{e}, \ldots, \mathbf{e})$, $\mathbf{e} \in \mathbb{R}^3$, so the total linear momentum

$$\mu = \sum m_i\dot{\mathbf{r}}_i$$

is conserved. It follows that the *center of mass*

$$\mathbf{R} := \frac{1}{\sum m_i}\sum m_i\mathbf{r}_i$$

satisfies $\ddot{\mathbf{R}} = 0$, thus $\mathbf{R}(t) = \mathbf{R}_0 + \mathbf{V}_0 t$. So the motion decomposes into the linear motion of the center of mass and the relative motion which can be described in *center-of-mass coordinates* $\mathbf{r}_i' := \mathbf{r}_i - \mathbf{R}$ in which the center of mass is at rest.

Example 2.10. (angular momentum).

Suppose that $L = \frac{1}{2}m\|\dot{\mathbf{r}}\|^2 - V(t, \mathbf{r})$ on $T\mathbb{R}^3$, and V is invariant under rotations around the axis $\mathbf{e} \in \mathbb{R}^3$. Then the conserved quantity in Noether's Theorem is

$$\mu = \langle m\dot{\mathbf{r}}, \mathbf{e} \times \mathbf{r}\rangle = \langle m\mathbf{r} \times \dot{\mathbf{r}}, \mathbf{e}\rangle,$$

the component in direction $\mathbf{e}$ of the *angular momentum* $m\mathbf{r} \times \dot{\mathbf{r}}$.

Problem 2.11. Prove: If the Lagrangian $\tilde L$ on the extended phase space $\mathbb{R} \times Q$ is invariant under a 1-parameter family of diffeomorphisms $\tilde\phi_s : \mathbb{R} \times Q \to \mathbb{R} \times Q$, and $\tilde\mu$ is the corresponding conserved quantity, then

$$\mu(q, \dot q) = \tilde\mu(t, q, 1, \dot q)$$

is a conserved quantity of the Lagrangian system for L on Q.

Show that conservation of energy is a special case of this.

Chapter 3

Hamiltonian mechanics

Hamilton's variational principle implies that Lagrange's equations are invariant under arbitrary coordinate transformations $q \mapsto Q$: If the Lagrange function is expressed in the old and new coordinates as $L(t, q, \dot{q}) = \tilde{L}(t, Q, \dot{Q})$, then

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \Longleftrightarrow \frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{Q}} - \frac{\partial L}{\partial Q} = 0.$$

Given the usefulness of coordinate changes in mechanics, one may ask whether the equations of mechanics are invariant under more general transformations that involve coordinates as well as velocities. That this is indeed the case becomes apparent in the Hamiltonian formulation of mechanics.

3.1 The Legendre transform

Definition. Let V be a finite dimensional vector space. The *Legendre transform* of a convex function $f : V \to \mathbb{R}$ is $\mathcal{L}f : V^* \to \mathbb{R}$,

$$\mathcal{L}f(p) := \max_{x \in V} \Big[\langle p, x \rangle - f(x) \Big].$$

Since f is convex, the function $p \mapsto \langle p, x \rangle - f(x)$ attains its unique maximum at the point x where $p = df(x)$, thus

$$\mathcal{L}f(p) = \langle p, x \rangle - f(x) \Big|_{p = df(x)}.$$

Proposition 3.1. *The Legendre transform $\mathcal{L}f$ of a convex function f is again convex, and $\mathcal{L}(\mathcal{L}f) = f$.*

Proof. In one dimension this can be seen geometrically from Figure ???. In general, set $g := \mathcal{L}f$. Convexity of g follows from

$$\begin{aligned} g\Big((1-t)p_0 + tp_1\Big) &= \max_x \Big[(1-t)\Big(\langle p_0, x\rangle - f(x)\Big) + t\Big(\langle p_1, x\rangle - f(x)\Big)\Big] \\ &\le (1-t)\max_x \Big[\langle p_0, x\rangle - f(x)\Big] + t\max_x\Big[\langle p_1, x\rangle - f(x)\Big] \\ &= (1-t)g(p_0) + tg(p_1). \end{aligned}$$

For the involutivity note that

$$\begin{aligned} \langle p, x\rangle - g(p) &= \langle p, x\rangle - \max_z\Big[\langle p, z\rangle - f(z)\Big] \\ &\le \langle p, x\rangle - \Big[\langle p, x\rangle - f(x)\Big] = f(x) \end{aligned}$$

for all p, and

$$\begin{aligned} \langle p, x\rangle - g(p)\Big|_{p=df(x)} &= \langle p, x\rangle - \Big[\langle p, z\rangle - f(z)\Big]\Big|_{p=df(x)=df(z)} \\ &= \langle p, x\rangle - \Big[\langle p, z\rangle - f(z)\Big]\Big|_{z=x} = f(x). \end{aligned}$$

This shows that

$$\mathcal{L}g(x) = \max_p\Big[\langle p, x\rangle - g(p)\Big] = f(x),$$

and the maximum is attained at $p = df(x)$. □

3.2 Hamilton's equations

Now we apply the Legendre transform in $\dot q$ to the Lagrange function (with fixed t, q) to obtain a function of the momenta $p_i = \frac{\partial L}{\partial \dot q_i}$, the *Hamiltonian*

$$H(t, q, p) := \langle p, \dot q\rangle - L(t, q, \dot q)\Big|_{p=\frac{\partial L}{\partial \dot q}}.$$

This is defined whenever $L : \mathbb{R} \times TQ \to \mathbb{R}$ is convex on each fibre T_qQ, and yields a function $H : \mathbb{R} \times T^*Q \to \mathbb{R}$ which is again fibrewise convex.

To find Lagrange's equations on $\mathbb{R}^n$ in terms of H, express the differential dH in two different ways:

$$\begin{aligned} dH &= \sum_i \Big(\frac{\partial H}{\partial q_i}dq_i + \frac{\partial H}{\partial p_i}dp_i\Big) + \frac{\partial H}{\partial t}dt \\ &= \sum_i \Big(p_i d\dot q_i + \dot q_i dp_i - \frac{\partial L}{\partial q_i}dq_i - \frac{\partial L}{\partial \dot q_i}d\dot q_i\Big) - \frac{\partial L}{\partial t}dt\Big|_{p_i=\frac{\partial L}{\partial \dot q_i}} \\ &= \sum_i \Big(\dot q_i dp_i - \frac{\partial L}{\partial q_i}dq_i\Big) - \frac{\partial L}{\partial t}dt. \end{aligned}$$

From this we read off

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \frac{\partial L}{\partial q_i} = -\frac{\partial H}{\partial q_i}, \quad \frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t},$$

which combined with Lagrange's equations $\dot{p}_i = \frac{\partial L}{\partial q_i}$ yields *Hamilton's equations*

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

A more conceptual way to see the equivalence of Hamilton's and Lagrange's equations is via variational principles.

Proposition 3.2. *(extended Hamilton's variational principle).*
Let $L : \mathbb{R}\times TQ \to \mathbb{R}$ be fibrewise convex and $H : \mathbb{R}\times T^*Q \to \mathbb{R}$ the corresponding Hamiltonian. Then a path $q : [a,b] \to Q$ is extremal for $\mathcal{L}(q) = \int_a^b L(t,q,\dot{q})dt$ with fixed $q(a), q(b)$ if and only if $(q, p = \frac{\partial L}{\partial \dot{q}}) : [a,b] \to T^*Q$ is extremal for the action

$$\mathcal{A}(q,p) = \int_a^b \Big(\langle p, \dot{q}\rangle - H(t,q,p)\Big)dt$$

with fixed $q(a), q(b)$.

Proof. If (q,p) is extremal for $\mathcal{A}$, then in particular on each fibre T_q^*Q, p is extremal for the function $p \mapsto \langle p, \dot{q}\rangle - H(t,q,p)$. By Proposition 3.1 this means that $p = \frac{\partial L}{\partial \dot{q}}$ and $\langle p, \dot{q}\rangle - H(t,q,p) = L(t,q,p)$. So (q,p) is extremal for $\mathcal{A}$ if and only if $p = \frac{\partial L}{\partial \dot{q}}$ and q is extremal for $\mathcal{L}$. □

Note that the expression $\langle p, \dot{q}\rangle$ is intrinsically defined for a path $x = (q,p) : [a,b] \to T^*Q$. We can make this more apparent by writing $q = \pi \circ x$ with the canonical projection $\pi : T^*Q \to Q$. Then $\langle p, \dot{q}\rangle = \langle p, T_{(q,p)}\pi \cdot \dot{x}\rangle = \lambda(\dot{x})$, where λ_{st} is the *canonical 1-form* on T^*Q,

$$(\lambda_{\mathrm{st}})_{(q,p)}(v) = \langle p, T_{(q,p)}\pi \cdot v\rangle, \quad v \in T_{(q,p)}T^*Q.$$

If q_i are local coordinates on Q and p_i dual coordinates on T_q^*Q, i.e. $p = \sum p_i dq_i \in T_q^*Q$, then

$$\lambda_{\mathrm{st}} = \sum p_i dq_i.$$

This can be seen by writing a tangent vector to T^*Q as $v = \sum(v_i \frac{\partial}{\partial q_i} + w_i \frac{\partial}{\partial p_i}$, hence $\langle p, T\pi \cdot v\rangle = \sum p_i v_i$.

With the canonical 1-form the action functional can be written in the inherently intrinsic form

$$\mathcal{A}(x) = \int_a^b \Big(x^*\lambda_{\mathrm{st}} - H(t,x)dt\Big).$$

Physicists and Arnold write this as $\int p\,dq - H\,dt$.

Now in local coordinates (q_i, p_i) as above, the variation of $\mathcal{A}$ with $\delta q(a) = \delta q(b) = 0$ equals

$$\begin{aligned} d\mathcal{A}(q,p)(\delta q, \delta p) &= \int_a^b \Big(\langle p, \dot{\delta q} \rangle + \langle \dot q, \delta p \rangle - \langle \frac{\partial H}{\partial q}, \delta q \rangle - \langle \frac{\partial H}{\partial p}, \delta p \rangle \Big) dt \\ &= \int_a^b \Big(\langle -\dot p - \frac{\partial H}{\partial q}, \delta q \rangle + \langle \dot q - \frac{\partial H}{\partial p}, \delta p \rangle \Big) dt. \end{aligned}$$

So we have shown:

Proposition 3.3. *Let $H : \mathbb{R} \times T^*Q \to \mathbb{R}$ be a (not necessarily fibrewise convex) Hamiltonian. Then for a smooth path $x = (q,p) : [a,b] \to T^*Q$ the following are equivalent:*

(i) x is extremal for $\mathcal{A}(x)$ among variations with fixed $q(a), q(b)$;

(ii) x is extremal for $\mathcal{A}(x)$ among variations with fixed $x(a), x(b)$;

(iii) x satisfies Hamilton's equations in any local coordinates q_i on Q and dual coordinates p_i on T_q^*Q.

The two last propositions together yield the equivalence of Lagrange's and Hamilton's equations for fibrewise convex Hamiltonians.

3.3 Canonical transformations

In the extended Hamilton's variational principle it becomes apparent which transformations $\psi : (q,p) \mapsto (Q,P)$ on $\mathbb{R}^{2n}$ leave Hamilton's equations invariant. For this, note that the first term in the action functional $\mathcal{A}$ is the integral of the 1-form $\sum_i p_i dq_i$ over the path (q,p). If we add to this 1-form an exact 1-form dS the value of the integral changes only by the values of S at the end points, so the extremals with fixed end points remain the same. This shows that Hamilton's equations are invariant under all transformations which satisfy

$$\sum P_i dQ_i - \sum p_i dq_i = dS.$$

Since exact and closed forms coincide on $\mathbb{R}^{2n}$, this is equivalent to

$$\sum dP_i \wedge dQ_i = \sum dp_i \wedge dq_i,$$

i.e. the transformation preserves the *standard symplectic form* on $\mathbb{R}^{2n}$,

$$\omega_{\mathrm{st}} := \sum dq_i \wedge dp_i.$$

Such a transformation is called *canonical* by physicists, and *symplectic* by mathematicians. We have just shown:

Proposition 3.4. *Symplectic transformations of $\mathbb{R}^{2n}$ preserve Hamilton's equations.*

The following problem shows that the converse is also true, so symplectic transformations are the largest class of transformations under which the equations of mechanics are invariant.

Problem 3.1. Prove that a transformation of $\mathbb{R}^{2n}$ which preserves Hamilton's equations for every Hamiltonian is symplectic.

Examples of symplectic transformations:

(i) $\psi(q,p) = \Big(\phi(q), [D\phi(q)^{-1}]^T p\Big)$ (pure coordinate transformation);

(ii) $\psi(q,p) = (q+q_0, p+p_0)$ (translation);

(iii) $\psi(q,p) = (p, -q)$;

(iv) any area preserving diffeomorphism of $\mathbb{R}^2$.

Note that T^*Q carries the *canonical symplectic form* $\omega_{\mathrm{st}} = -d\lambda_{\mathrm{st}}$. By the last proposition, Hamilton's equations hold in any local coordinates (q_i, p_i) on T^*Q in which the canonical symplectic form becomes $\sum dq_i \wedge dp_i$. Coordinates q_i on Q and dual coordinates p_i on T_q^*Q are examples of such coordinates on T^*Q, but by far not the only ones.

Having preceeded thus far, it is natural to ask: Are there more general spaces than cotangent bundles on which Hamilton's equations can be defined? Call a *symplectic atlas* on a manifold an atlas all of whose transition maps are symplectic, and a *symplectic manifold* a manifold with a symplectic atlas. It follows from Proposition 3.4 that Hamilton's equations are defined intrinsically on a symplectic manifold.

It turns out that symplectic manifolds form indeed a large and interesting class of manifolds, and they provide the natural setting for Hamilton's equations. The geometry of symplectic manifolds and the study of Hamiltonian systems on symplectic manifolds is the subject of Part II.

3.4 Electromagnetic fields

We conclude this chapter by showing how electromagnetic forces fit into the Lagrangian and Hamiltonian formalisms. The electrostatic force is already built in because it is conservative, but the magnetostatic force is not. In order to include also dynamic electromagnetic forces, let us start from *Maxwell's equations*

$$\begin{aligned} \operatorname{div} \mathbf{D} &= \rho; \quad \operatorname{curl} \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0; \\ \operatorname{div} \mathbf{B} &= 0; \quad \operatorname{curl} \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{j}. \end{aligned}$$

Here $\mathbf{E}$ is the electric field, $\mathbf{D}$ the dielectric displacement, $\mathbf{B}$ the magnetic induction, $\mathbf{H}$ the magnetic field, ρ the charge density and $\mathbf{j}$ the current density. In vacuum the fields are related to each other and the velocity of light c by

$$\mathbf{D} = \varepsilon_0 \mathbf{E}; \quad \mathbf{B} = \mu_0 \mathbf{H}; \quad \frac{1}{\varepsilon_0 \mu_0} = c^2.$$

The force on a particle of charge e moving with velocity $\mathbf{v}$ equals

$$\mathbf{F} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

The third Maxwell equation implies $\mathbf{B} = \operatorname{curl} \mathbf{A}$ for a *vector potential* $\mathbf{A}(t, \mathbf{r})$. Plugging this into the second equation yields $\operatorname{curl}(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}) = 0$. So $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \phi$ for a *scalar potential* $\phi(t, \mathbf{r})$, and the force on a particle becomes

$$\mathbf{F} = e(\mathbf{v} \times \operatorname{curl} \mathbf{A} - \frac{\partial \mathbf{A}}{\partial t} - \nabla \phi).$$

Can we find a *generalized potential* $U(t, \mathbf{r}, \dot{\mathbf{r}})$ such that Lagrange's equations for $L = \frac{1}{2}\|\dot{\mathbf{r}}\|^2 - U$ are equivalent to $\mathbf{F} = m\ddot{\mathbf{r}}$ with the above force $\mathbf{F}$? This will be the case if we can find U such that

$$\mathbf{F} = \frac{d}{dt}\frac{\partial U}{\partial \dot{\mathbf{r}}} - \frac{\partial U}{\partial \mathbf{r}}.$$

Since the force $\mathbf{F}$ contains no $\ddot{\mathbf{r}}$, U must be linear in $\dot{\mathbf{r}}$,

$$U(t, \mathbf{r}, \dot{\mathbf{r}}) = \mathbf{G}(t, \mathbf{r}) \cdot \dot{\mathbf{r}} + V(t, \mathbf{r}).$$

For such U,

$$\frac{d}{dt}\frac{\partial U}{\partial \dot{\mathbf{r}}} - \frac{\partial U}{\partial \mathbf{r}} = \frac{\partial \mathbf{G}}{\partial t} + \Big[\frac{\partial \mathbf{G}}{\partial \mathbf{r}} - (\frac{\partial \mathbf{G}}{\partial \mathbf{r}})^T\Big]\dot{\mathbf{r}} - \nabla V.$$

Now the third component of $\big[\frac{\partial \mathbf{G}}{\partial \mathbf{r}} - (\frac{\partial \mathbf{G}}{\partial \mathbf{r}})^T\big]\mathbf{v}$ equals

$$\begin{aligned}\sum_{i=1}^{3}\Big(\frac{\partial G_3}{\partial x_i} - \frac{\partial G_i}{\partial x_3}\Big)v_i &= \Big(\frac{\partial G_3}{\partial x_1} - \frac{\partial G_1}{\partial x_3}\Big)v_1 + \Big(\frac{\partial G_3}{\partial x_2} - \frac{\partial G_2}{\partial x_3}\Big)v_2 \\ &= -(\operatorname{curl} \mathbf{G})_2 v_1 + (\operatorname{curl} \mathbf{G})_1 v_2 = (\operatorname{curl} \mathbf{G} \times \mathbf{v})_3,\end{aligned}$$

and similarly for the other components. Thus $\big[\frac{\partial \mathbf{G}}{\partial \mathbf{r}} - (\frac{\partial \mathbf{G}}{\partial \mathbf{r}})^T\big]\mathbf{v} = \operatorname{curl} \mathbf{G} \times \mathbf{v}$ and

$$\frac{d}{dt}\frac{\partial U}{\partial \dot{\mathbf{r}}} - \frac{\partial U}{\partial \mathbf{r}} = \frac{\partial \mathbf{G}}{\partial t} - \mathbf{v} \times \operatorname{curl} \mathbf{G} - \nabla V.$$

We see that this matches the force $\mathbf{F}$ above if $\mathbf{G} = -e\mathbf{A}$ and $V = e\phi$! So the electromagnetic force is derived from the Lagrange function

$$L = \frac{1}{2}m\|\dot{\mathbf{r}}\|^2 + e\mathbf{A}(t, \mathbf{r}) \cdot \dot{\mathbf{r}} - e\phi(t, \mathbf{r}).$$

The natural generalization of the vector potential $\mathbf{A}$ to a manifold Q is as a time-dependent 1-form A on Q. Incoroprating m and e into the Riemannian metric and the potentials,

$$L = \frac{1}{2}\|\dot q\|^2 + \langle A(t,q), \dot q\rangle - V(t,q)$$

is the general form of a natural Lagrange function including electromagnetic forces.

The inclusion of electromagnetism in the Hamiltonian formalism is straightforward from the Lagrange function. The momentum is

$$p = \frac{\partial L}{\partial q} = \dot q + A(t,q),$$

so $\dot q = p - A$ and

$$\begin{aligned} H &= \langle p, \dot q\rangle - L \\ &= \langle p, p - A\rangle - \frac{1}{2}\|p - A\|^2 - \langle A, p - A\rangle + V \\ &= \frac{1}{2}\|p - A\|^2 + V. \end{aligned}$$

We see that

$$H = \frac{1}{2}\|p - A(t,q)\|^2 + V(t,q)$$

is the general form of a natural Hamiltonian including electromagnetic forces. Again the magnetic potential A is a time-dependent 1-form on Q.

Now assume for the moment that A is time-independent. The diffeomorphism

$$\psi : T^*Q \to T^*Q, \quad (q,p) \mapsto \big(q, p + A(q)\big)$$

pulls back the Hamiltonian $H = \frac{1}{2}\|p - A\|^2 + V$ to $\psi^* H = \frac{1}{2}\|p\|^2 + V$, so it transforms away the magnetic term. But it also changes the symplectic structure of phase space: the 2-form $\psi^*\omega_{\mathrm{st}}$ is no longer the canonical symplectic form on T^*Q. Indeed, the canonical 1-form transforms to

$$\begin{aligned} (\psi^*\lambda_{\mathrm{st}})_{(q,p)}(v) &= (\lambda_{\mathrm{st}})_{(q,p+A)}(T\psi \cdot v) = \langle p + A, T\pi \cdot T\psi \cdot v\rangle \\ &= (\lambda_{\mathrm{st}})_{(q,p)}(v) + (\pi^* A)_{(q,p)}(v). \end{aligned}$$

So

$$\begin{aligned} \psi^*\lambda_{\mathrm{st}} &= \lambda_{\mathrm{st}} + \pi^* A, \quad \text{and} \\ \psi^*\omega_{\mathrm{st}} &= \omega_{\mathrm{st}} - \pi^* dA. \end{aligned}$$

We will see in Part II that a general Hamiltonian system is determined entirely by a Hamiltonian and a symplectic form. THen we have shown:

A natural Hamiltonian system with magnetic term A, $H = \frac{1}{2}\|p - A\|^2 + V$, is equivalent to a system without magnetic term, $H = \frac{1}{2}\|p\|^2 + V$, but with the symplectic structure twisted by A, $\omega_{\mathrm{st}} - \pi^* dA$.

For a time-dependent magnetic potential A the same is true, only that now we obtain a time-dependent symplectic form.

Now remember that the magnetic induction $\mathbf{B} = \operatorname{curl}\mathbf{A}$. If we write $\mathbf{A}$ as a 1-form A and $\mathbf{B}$ as a 2-form B this reads $dA = B$. So the exact 2-form $B = dA$ on a manifold Q should play the role of a magnetic field. Being mathematicians, we immediately feel compelled to call any *closed* 2-form B on Q a magnetic field. This indeed makes sense mathematically: the magnetic field enters by twisting the symplectic form to $\omega_{\mathrm{st}} - \pi^* B$. Note that of B is not exact we cannot incorporate it any more into the Hamiltonian and leave the symplectic structure the canonical one. So we are forced to consider more general symplectic structures if we want to study nonexact magnetic fields.

Physically, nonexact magnetic fields correspond to magnetic monopoles, none of which have ever been observed in nature. But mathematically they have very interesting properties, some of which we will see in Part II.

Chapter 4

Geodesics of left-invariant metrics on Lie groups

4.1 The rigid body

A *rigid body* is an ensemble of point masses m_i at positions $\mathbf{R}_i \in \mathbb{R}^3$ whose relative distances $\|\mathbf{R}_i - \mathbf{R}_j\|$ are fixed. Let us assume that the mass points are not all on a line (the degenerate case of all masses on a line will be treated in a problem). Fix a reference position of the body. Then an arbitrary position is obtained from the reference position by a uniquely determined rigid motion, i.e. a translation and rotation applied to all the positions $\mathbf{R}_i$. We will only consider the case that one point of the body remains fixed. For force-free motion without constraints this can be achieved by transforming to center-of-mass coordinates, and the fixed point is the center of mass. In general, the existence of a fixed point is an additional constraint, and the fixed point need not be the center of mass (even for force-free motion). Choose coordinates such that the fixed point is the origin. Then positions of the rigid body are obtained from the reference position by pure rotations, i.e. by elements of $SO(3)$.

So a motion of the rigid body is described by a curve $g(t)$ in the group $SO(3)$. the position of mass m_i at time t being $g(t)\mathbf{R}_i$. It will be useful to express vectors in two different coordinate systems: the *space coordinate system* which is fixed in space, and the *body coordinate system* which rotates with the body. Vectors in space coordinates are written by small letters and in body coordiantes by capital letters. A vector $\mathbf{R}$ in body coordinates corresponds to the vector $\mathbf{r} = g(t)\mathbf{R}$ is space coordinates.

The velocity of the mass m_i in space coordinates is $\dot{\mathbf{r}}_i = \dot{g}\mathbf{R}_i = (\dot{g}g^{-1})\mathbf{r}_i$, where the matrix $\dot{g}g^{-1}$ is skew-symmetric. Recall that every skew-symmetric operator on $\mathbb{R}^3$ can be written as the cross product with a uniquely determined vector (this follows simply because the cross product operators form a 3-dimensional subspace of the 3-dimensional space of skew-symmetric operators). So there exists a unique vector $\omega = \omega(t)$ such that

$$\dot{\mathbf{r}}_i = \omega \times \mathbf{r}_i \quad \text{for all } i.$$

ω is called the *angular velocity in space coordinates*. The angular velocity in body coordinates is $\Omega = g^{-1}\omega$.

The total kinetic energy of the rigid body equals

$$\begin{aligned}
T &= \frac{1}{2}\sum m_i \|\omega \times \mathbf{r}_i\|^2 = \frac{1}{2}\sum m_i \|\Omega \times \mathbf{R}_i\|^2 \\
&= \frac{1}{2}\sum m_i \langle \mathbf{R}_i \times (\Omega \times \mathbf{R}_i), \Omega\rangle = \frac{1}{2}\sum m_i \langle \Omega \|\mathbf{R}_i\|^2 - \mathbf{R}_i\langle \Omega, \mathbf{R}_i\rangle, \Omega\rangle \\
&= \frac{1}{2}\sum m_i \Big(\|\mathbf{R}_i\|^2 \|\Omega\|^2 - \langle \mathbf{R}_i, \Omega\rangle^2 \Big) \\
&=: \frac{1}{2}\langle I\Omega, \Omega\rangle.
\end{aligned}$$

The positive definite operator $I : \mathbb{R}^3 \to \mathbb{R}^3$ is called the *inertia tensor* of the rigid body around the origin.

For the *free rigid body* (with a fixed point), i.e. in the absence of external forces, the Lagrange function equals the kinetic energy, and the equations of motions can be derived from the expression for T. This will be done in Section 4.3. Since this causes no additional difficulties, we will derive the equations of motion for arbitrary Lie groups – which also gives me a good reason to review in the next section some basic facts about Lie groups.

Problem 4.1. Show that a rigid body all of whose mass points lie on one line in a constant gravitational field is equivalent to a spherical pendulum. How is the inertia tensor related to the mass and length of the pendulum?

4.2 Lie groups and Lie algebras

Definition. A *Lie group* is a manifold G with a group operation $\cdot$ such that $G \times G \to G$, $(g, h) \mapsto g \cdot h$ is a smooth map.

Problem 4.2. Prove that *left multiplication* $L_g : h \mapsto gh$, *right multiplication* $R_g : h \mapsto hg$ and inversion $g \mapsto g^{-1}$ are diffeomorphisms of G.

The Lie algebra

A vector field X on G is called *left-invariant (right-invariant)* if $L_{g*}X = X$ ($R_{g*}X = X$) for all $g \in G$. A left-invariant vector field is determined uniquely by its value at the *identity* $e \in G$, and conversely, every $X \in T_eG$ defines a left-invariant vector field $\underline{\mathrm{X}} = \underline{\mathrm{X}}^G$ via $\underline{\mathrm{X}}_e = T_eL_g \cdot X$. This defines an isomorphism between T_eG and the space of left-invariant vector fields.

Since $L_{g*}[X,Y] = [L_{g*}X, L_{g*}Y]$, the space of left-invariant vector fields is closed under the Lie bracket. Thus the Lie bracket on vector fields induces a bracket on T_eG via

$$[X,Y] := [\underline{\mathrm{X}}, \underline{\mathrm{Y}}]_e.$$

Definition. A *Lie algebra* is a vector space V with a bilinear operation $[\ ,\ ] : V \to V$ satisfying

(i) (skew-symmetry) $[X,Y] = -[Y,X]$;

(ii) (Jacobi identity) $[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$.

$\mathfrak{g} := T_eG$ with the bracket $[X,Y] := [\underline{\mathrm{X}}, \underline{\mathrm{Y}}]_e$ is called the *Lie algebra of the Lie group G*.

The exponential map

Given $X \in \mathfrak{g}$, let $\gamma_X(t)$ be the integral curve of $\underline{\mathrm{X}}$ starting at e. The curves $t \mapsto \gamma_X(s+t)$ and $t \mapsto \gamma_X(s)\gamma_X(t)$ are both integral curves of $\underline{\mathrm{X}}$ starting at $\gamma_X(s)$, hence they are equal:

$$\gamma_X(s+t) = \gamma_X(s)\gamma_X(t).$$

This shows that $\gamma_X(t)$ is defined for all $t \in \mathbb{R}$ and defines a smooth *1-parameter subgroup* of G.

Definition. The *exponential map* of G is $\exp : \mathfrak{g} \to G$, $\exp(X) := \gamma_X(1)$.

It follows that $\exp(tX) = \gamma_X(t)$, and the flow $\phi_t^{\underline{\mathrm{X}}}$ of $\underline{\mathrm{X}}$ is given by

$$\phi_t^{\underline{\mathrm{X}}}(g) = g \exp tX.$$

In particular $\underline{\mathrm{X}}_g = \frac{d}{dt}\big|_{t=0} g \exp tX$, so the left-invariant vector field $\underline{\mathrm{X}}$ is the vector field induced by the action of G on itself via *right* multiplication.

In general, $\exp : \mathfrak{g} \to G$ is not a group homomorphism. However, restricted to commuting elements of $\mathfrak{g}$ it is:

Lemma 4.1. *If* $[X,Y] = 0$ *then* $\exp(X+Y) = \exp X \exp Y$.

Proof. Denote the flows of $\underline{\mathrm{X}}, \underline{\mathrm{Y}}$ by $\phi_t^{\underline{\mathrm{X}}}, \phi_t^{\underline{\mathrm{Y}}}$. Since $[\underline{\mathrm{X}}, \underline{\mathrm{Y}}]_e = [X,Y] = 0$ and $[\underline{\mathrm{X}}, \underline{\mathrm{Y}}]$ is left-invariant, $[\underline{\mathrm{X}}, \underline{\mathrm{Y}}] = 0$. This implies $\frac{d}{dt}(\phi_t^{\underline{\mathrm{Y}}})^* \underline{\mathrm{X}} = (\phi_t^{\underline{\mathrm{Y}}})^* [\underline{\mathrm{Y}}, \underline{\mathrm{X}}] = 0$, and therefore $(\phi_t^{\underline{\mathrm{Y}}})_* \underline{\mathrm{X}} = \underline{\mathrm{X}}$. By Lemma A.1, the flow $\phi_t^{\underline{\mathrm{Y}}} \circ \phi_t^{\underline{\mathrm{X}}}$ is generated by the vector field $\underline{\mathrm{Y}} + (\phi_t^{\underline{\mathrm{Y}}})_* \underline{\mathrm{X}} = \underline{\mathrm{X}} + \underline{\mathrm{Y}}$, thus

$$\exp tX \exp tY = \phi_t^{\underline{\mathrm{Y}}} \circ \phi_t^{\underline{\mathrm{X}}}(e) = \phi_t^{\underline{\mathrm{X}}+\underline{\mathrm{Y}}}(e) = \exp t(X+Y).$$

□

The adjoint action

Definition. Conjugation $L_g R_{g^{-1}} : G \to G$, $h \mapsto ghg^{-1}$ induces linear maps

$$\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}, \quad \mathrm{Ad}_g := T_e(L_g R_{g^{-1}});$$

$$\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}, \quad \mathrm{ad}_X Y := \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp tX} Y.$$

Denote by $\mathrm{Ad}^*_g, \mathrm{ad}^*_X : \mathfrak{g}^* \to \mathfrak{g}^*$ the adjoint maps.

Obviously $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \circ \mathrm{Ad}_h$ and $\mathrm{Ad}^*_{gh} = \mathrm{Ad}^*_h \circ \mathrm{Ad}^*_g$, so Ad (resp. Ad^*) defines a left (resp. right) representation of G on the vector space $\mathfrak{g}$. The following lemma provides a useful way of computing brackets:

Lemma 4.2. $\mathrm{ad}_X Y = [X, Y]$.

Proof. For any $g \in G$,

$$(R^*_g \underline{\mathrm{Y}})_e = T_g R_{g^{-1}} \circ T_e L_g \cdot Y = \mathrm{Ad}_g Y.$$

So if $\phi_t = R_{\exp tX}$ is the flow of $\underline{\mathrm{X}}$,

$$(\phi^*_t \underline{\mathrm{Y}})_e = \mathrm{Ad}_{\exp tX} Y.$$

It follows that

$$\mathrm{ad}_X Y = \frac{d}{dt}\Big|_{t=0} (\phi^*_t \underline{\mathrm{Y}})_e = [\underline{\mathrm{X}}, \underline{\mathrm{Y}}]_e = [X, Y].$$

□

Lie group homomorphisms

Lemma 4.3. *For a Lie group homomorphism $f : G \to H$,*

$$f \circ \exp = \exp \circ T_e f : \mathfrak{g} \to H.$$

Proof. For $X \in \mathfrak{g}$ the curve $t \mapsto f(\exp tX)$ satisfies

$$\begin{aligned}\frac{d}{dt} f(\exp tX) &= T_{\exp tX} f \circ T_e L_{\exp tX} \cdot X = T_e(f \circ L_{\exp tX}) \cdot X \\ &= T_e(L_{f(\exp tX)} \circ f) \cdot X = T_e L_{f(\exp tX)} \circ T_e f \cdot X \\ &= \underline{(T_e f \cdot X)}_{f(\exp tX)},\end{aligned}$$

so it is the integral curve of $\underline{(T_e f \cdot X)}$ starting at e: $f(\exp tX) = \exp(t T_e f \cdot X)$. □

Corollary 4.4. *For a Lie group homomorphism $f : G \to H$, $T_e f : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism.*

Proof. Differentiating $L_{f(g)}R_{f(g)^{-1}}\big(f(h)\big) = f(ghg^{-1})$ with respect to h yields

$$\mathrm{Ad}_{f(g)} \circ T_e f = T_e f \circ Ad_g : \mathfrak{g} \to \mathfrak{h},$$

and therefore

$$\begin{aligned} T_e f \cdot [X, Y] &= T_e f \cdot \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp tX} Y \\ &= \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{f(\exp tX)} \circ T_e f \cdot Y \\ &= \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{\exp tT_e f \cdot X)} \circ T_e f \cdot Y \\ &= [T_e f \cdot X, T_e f \cdot Y]. \end{aligned}$$

□

Corollary 4.5. *If G is connected and $f_1, f_2 : G \to H$ are Lie group homomorphisms such that $T_e f_1 = T_e f_2$, then $f_1 = f_2$.*

Proof. By the hypothesis and Lemma 4.3, $f_1 \circ \exp = f_2 \circ \exp : \mathfrak{g} \to H$. This implies the $f_1 = f_2$ on a neighbourhood U of e in G, and consequently $f_1 = f_2$ on the subgroup $K \subset G$ generated by U. Now K is clearly open, and it is closed because $G \setminus K = \cup_{g \in G \setminus K} gK$ is a union of open sets. Since G is connected, $K = G$. □

Applying Lemma 4.3 and Corollary 4.4 to the Lie group homomorphism $L_g R_{g^{-1}} : G \to G$, we obtain

$$\begin{aligned} \exp(Ad_g X) &= g(\exp X)g^{-1}, \quad \text{and} \\ \mathrm{Ad}_g[X, Y] &= [\mathrm{Ad}_g X, \mathrm{Ad}_g Y]. \end{aligned}$$

Problem 4.3. Show that for a matrix group $G \subset GL(n, \mathbb{R})$,

- $\exp(X) = e^X$ is the exponential of the matrix $X \in \mathfrak{g} \subset \mathbb{R}^{n \times n}$;
- $\mathrm{Ad}_g X = gXg^{-1}$ for $g \in G$, $X \in \mathfrak{g}$;
- $[X, Y] = XY - YX$ for $X, Y \in \mathfrak{g}$.

Problem 4.4. Show that the differential of the determinant $\det : GL(n, \mathbb{R}) \to \mathbb{R}^*$ at the identity is the trace $\mathrm{tr} : \mathbb{R}^{n \times n} \to \mathbb{R}$, and conclude that

$$\det(e^X) = e^{\mathrm{tr} X}, \quad X \in \mathbb{R}^{n \times n}.$$

Let us summarize the definitions and formulae of this section.

Summary: formulae for Lie groups

For $g, h \in G$, $X, Y \in \mathfrak{g}$ and $s, t \in \mathbb{R}$,

- $\underline{\mathrm{X}}^G_g := T_e L_g \cdot X = \frac{d}{dt}\big|_{t=0} g \exp tX$;
- $[X, Y] := [\underline{\mathrm{X}}^G, \underline{\mathrm{Y}}^G]_e$;
- $\exp(s+t)X = \exp sX \exp tX$;
- $[X, Y] = 0$ implies $\exp(X+Y) = \exp(X)\exp(Y)$;
- If f is a Lie group homomorphism then $f \circ \exp = \exp \circ T_e f$, and $T_e f$ is a Lie algebra homomorphism;
- $\mathrm{Ad}_g := T_e(L_g R_{g^{-1}}) : \mathfrak{g} \to \mathfrak{g}$;
- $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \circ \mathrm{Ad}_h$, $\mathrm{Ad}^*_{gh} = \mathrm{Ad}^*_h \circ \mathrm{Ad}^*_g$;
- $\exp(\mathrm{Ad}_g X) = g(\exp X)g^{-1}$;
- $\mathrm{Ad}_g[X, Y] = [\mathrm{Ad}_g X, \mathrm{Ad}_g Y]$;
- $\mathrm{ad}_X Y := \frac{d}{dt}\mathrm{Ad}_{\exp tX} Y = [X, Y]$.

Definition. A *torus* is a compact connected abelian Lie group. A *maximal torus* in a Lie group G is a subtorus $T \subset G$ which is not contained in any bigger subtorus.

Problem 4.5. Prove that any n-dimensional torus is isomorphic to $\mathbb{R}^n/\mathbb{Z}^n$.

Problem 4.6. (cf. [3]). Let G be a compact connected Lie group. Prove:

(i) G possesses a maximal torus T.

(ii) The map $G/T \times T \to G$, $(g, t) \mapsto gtg^{-1}$, is surjective.

(iii) Every element of G is contained in a maximal torus.

(iv) Any two maximal tori are conjugated.

(v) The inclusion $T \subset G$ induces a surjection on the fundamental group.

Problem 4.7. Prove that $\pi_2(G) = 0$ for every compact Lie group G.

4.3 Geodesics of left-invariant metrics

Euler's equation

Let G be a Lie group with Lie algebra $\mathfrak{g}$. A *left-invariant metric* on G is specified by an inner product on the Lie algebra which we write in the form

$\langle I\cdot,\cdot\rangle$. Here $\langle\ ,\ \rangle$ is the duality pairing between a space and its dual, and the *inertia tensor* $I:\mathfrak{g}\to\mathfrak{g}^*$ is linear, self-adjoint and positive definite (although positive definiteness is not needed here). Denote by L_g left multiplication and by $L_{g*}:T_hG\to T_{gh}G$, $L_g^*:T_{gh}^*G\to T_h^*G$ the (co)tangential maps, and similarly for right multiplication R_g. Then the metric at a point $g\in G$ is given by $\langle I_g\cdot,\cdot\rangle$, where

$$I_g := L_{g^{-1}}^* I L_{g^{-1}*} : T_gG \to T_g^*G.$$

For a curve $g(t)$ in G consider the following quantities:

- $\dot g\in T_gG$ the *angular velocity*;
- $\omega := R_{g^{-1}*}\dot g\in\mathfrak{g}$ the *angular velocity in space coordinates*;
- $\Omega := L_{g^{-1}*}\dot g = \mathrm{Ad}_{g^{-1}}\omega\in\mathfrak{g}$ the *angular velocity in body coordinates*;
- $I_g\dot g\in T_g^*G$ the *angular momentum*;
- $\mu := R_g^* I_g\dot g\in\mathfrak{g}^*$ the *angular momentum in space coordinates*;
- $M := L_g^* I_g\dot g = \mathrm{Ad}_g^*\mu = I\Omega\in\mathfrak{g}^*$ the *angular momentum in body coordinates.*

These notations are justified as follows: We think of the motion $g(t)$ as describing a rigid body whose configuration at time t is obtained from its initial configuration by left roation by $g(t)$. Then the angular velocity in space coordinates is the unique $\omega\in\mathfrak{g}$ such that $\dot g = \frac{d}{ds}\big|_{s=0}\exp(s\omega)g = R_{g*}\omega$, etc.

Now suppose that $g(t)$ is a geodesic, i.e. a solution of the Lagrangian system with Lagrange function

$$L(g,\dot g)=\frac12\langle I_g\dot g,\dot g\rangle=\frac12\langle I\Omega,\Omega\rangle=\frac12\langle M,\Omega\rangle=\frac12\langle \mu,\omega\rangle.$$

Note that for $G=SO(3)$ this is indeed the Lagrange function of the free rigid body fixed at a point. Since L is invariant under left multiplication, Noether's Theorem provides a conserved quantity for every $X\in\mathfrak{g}$ which takes the form

$$\langle\frac{\partial L}{\partial\dot g},\frac{d}{ds}\big|_{s=0}\exp(sX)g\rangle=\langle I_g\dot g,R_{g*}X\rangle=\langle\mu,X\rangle.$$

So the space angular momentum is conserved,

$$\frac{d\mu}{dt}=0.$$

For $M=\mathrm{Ad}_g^*\mu$ we obtain

$$\begin{aligned}\frac{dM}{dt}&=\frac{d}{ds}\big|_{s=0}\mathrm{Ad}^*_{g(t+s)}\mu=\frac{d}{ds}\big|_{s=0}\mathrm{Ad}^*_{g(t)^{-1}g(t+s)}\mathrm{Ad}^*_{g(t)}\mu\\&=\mathrm{ad}^*_\Omega\mathrm{Ad}^*_g\mu=\mathrm{ad}^*_\Omega M,\end{aligned}$$

so the body angular momentum satisfies *Euler's equation*

$$\frac{dM}{dt} = \mathrm{ad}^*_\Omega M, \quad \Omega = I^{-1}M.$$

This is a first order ODE for Ω. From its solution $\Omega(t)$ the geodesic $g(t)$ can be recovered by solving the ODE on the group $\dot g = L_{g*}\Omega$. So we have proved:

Proposition 4.6. *A curve $g(t)$ in G is a geodesic ofthe left-invariant metric if and only if its body angular velocity $\Omega = L_{g^{-1}*}\dot g$ satisfies Euler's equation*

$$\frac{dM}{dt} = \mathrm{ad}^*_\Omega M, \quad \Omega = I^{-1}M.$$

Conserved quantities

Since $L = T$ equals the total energy, conservation of energy yields

$$T = \frac{1}{2}\langle M, \Omega\rangle = \text{const}.$$

This can also be verified directly from Euler's equation:

$$\frac{dT}{dt} = \langle \dot M, \Omega\rangle = \langle \mathrm{ad}^*_\Omega M, \Omega\rangle = \langle M, [\Omega, \Omega]\rangle = 0.$$

Further conserved quantities come from *Casimir functions*, i.e. functions $C : \mathfrak{g}^* \to \mathbb{R}$ satisfying $\mathrm{ad}^*_{d_\xi C}\xi = 0$ for all $\xi \in \mathfrak{g}^*$ (see Section ???). By Euler's equation,

$$\frac{d}{dt}C(M) = \langle \dot M, d_M C\rangle = \langle M, [\Omega, d_M C]\rangle = -\langle \mathrm{ad}^*_{d_M C} M, \Omega\rangle = 0.$$

So every Casimir function yields a conserved quantity

$$C(M) = \text{const}.$$

Next suppose that $\mathfrak{g}$ admits an Ad-invariant inner product $\langle\ ,\ \rangle_{inv}$ (which is always the case, e.g., if G is compact). Ad-invariance implies

$$\langle [X, Y], Z\rangle_{inv} = \langle X, [Y, Z]\rangle_{inv}, \quad X, Y, Z \in \mathfrak{g}.$$

Now identify $\mathfrak{g}$ and $\mathfrak{g}^*$ via the invariant inner product and drop the subscript *inv*. Then $\langle \mathrm{ad}^*_X Y, Z\rangle = \langle Y, [X, Z]\rangle = \langle [Y, X], Z\rangle$, thus

$$\mathrm{ad}^*_X Y = -[X, Y].$$

So we can write Euler's equation on $\mathfrak{g}$ as

$$\frac{dM}{dt} = -[\Omega, M].$$

The function $C(X) = \langle X, X\rangle$ is a Casimir function:

$$\mathrm{ad}^*_{d_X C} X = -[d_X C, X] = -[2X, X] = 0.$$

So we have a conserved quantity

$$\langle M, M\rangle = \text{const}.$$

In the case $G = SO(3)$ the two conserved quantities determine the motion of the body angular momentum (and hence of the body angular velocity): Trajectories of M are the intersections of spheres $\{\|M\| = \text{const}\}$ with ellipsoids $\{\langle I^{-1}M, M\rangle = \text{const}\}$. Figure ??? shows the trajectories on the sphere in the case of 3 distinct eigenvalues of the inertia tensor. The 6 critical points correspond to stationary rotations around the 3 principal axes (in positive and negative directions). The rotations around the middle axis are connected by 4 heteroclinic orbits, which shows that these rotations are unstable. The rotations around the largest and smallest axes are stable for the motion of M. But this does not reveal whether they are stable for the motion of the rigid body. We will see in the next section that they are indeed stable. All other orbits are periodic for the motion of M. The next section shows that they are not necessarily periodic for the motion of the rigid body, but "almost periodic".

Poinsot's Theorem

The motion on the Lie group G can be visualized by describing how its adjoint action moves points in the Lie algebra. For $G = SO(3)$ the adjoint representation coincides with the standard representation on $\mathbb{R}^3$, so the motion on the Lie algebra corresponds to the motion of the rigid body in space. For $G = SO(n)$, $n > 3$, this is no longer the case.

Let $E_0 := \{X \in \mathfrak{g} \mid \langle IX, X\rangle = 1\}$ be the *inertia ellipsoid* at time 0 and

$$E_t := \mathrm{Ad}_{g(t)} E_0 = \{X \in \mathfrak{g} \mid \langle \mathrm{Ad}^*_{g^{-1}} I \mathrm{Ad}_{g^{-1}} X, X\rangle = 1\}$$

the inertia ellipsoid at time t in space coordinates. The tangent plane to E_t is perpendicular to μ at the points X where the normal vector is parallel to μ, i.e. $\mathrm{Ad}^*_{g^{-1}} I \mathrm{Ad}_{g^{-1}} X = c\mu$, or equivalently,

$$X = c\mathrm{Ad}_g I^{-1} \mathrm{Ad}^*_g \mu = c\mathrm{Ad}_g I^{-1} M = c\mathrm{Ad}_g \Omega = c\omega.$$

The constant c is determined from the condition $X \in E_t$,

$$1 = \langle c\mu, X\rangle = \langle c\mu, c\omega\rangle = 2c^2 T,$$

so $c = \pm 1/\sqrt{T}$, where $T = L$ is the kinetic energy. It follows that

$$\langle \mu, X\rangle = \langle \mu, c\omega\rangle = \pm\sqrt{2T}.$$

So the ellipsoid E_t is tangent to the *invariant plane*

$$\{X \in \mathfrak{g} \mid \langle \mu, X \rangle = \sqrt{2T}\}$$

at the point $\omega(t)/\sqrt{2T}$. Since $\omega(t)$ is the instantaneous axis of rotation, the point of tangency has zero velocity, and we have proved

Theorem 4.7 (Poinsot's Theorem)**.** *The inertia ellipsoid rolls without slipping on the invariant plane* $\{X \in \mathfrak{g} \mid \langle \mu, X \rangle = \sqrt{2T}\}$.

For $G = SO(3)$, Poinsot's Theorem gives us a complete description of the motion of the free rigid bodyfixed at a point: Suppose that the three eigenvalues of the inertia tensor are distinct. There are stationary rotations around the 3 principal axes of inertia. The rotations around the largest and smallest axes are stable, while the roation around the middle axis is unstable. There are heteroclinic orbits between rotations around the middle axis in opposite directions. All other orbits lie on invariant 2-tori on which the flow is linear. Orbits on an invariant torus are either periodic (for rational slope) or dense in the torus (for irrational slope). The trajectory of the point of tangency in the invariant plane is either a closed curve or dense in an annulus.

Stationary rotations

A *stationary rotation* is a geodesic on G which is also a 1-parameter subgroup.

Write a stationary rotation as $g(t) = \exp tX$. Its angular velocity in body coordinates is

$$\Omega = L_{g^{-1}*}\dot{g} = X = \text{const}.$$

So $g(t) = \exp t\Omega$ is a geodesic iff $\Omega = \text{const}$ satisfies Euler's equation

$$0 = \dot{M} = \mathrm{ad}^*_\Omega(I\Omega).$$

For $G = SO(3)$ identify $\mathfrak{g}$ and $\mathfrak{g}^*$ via the Euclidean metric on $\mathfrak{g}$. Then the last equation becomes $[\Omega, I\Omega] = 0$, so Ω is an eigenvector of I. If I has 3 distinct eigenvalues, then the 3 eigendirections of I correspond to the 6 stationary rotations discussed earlier in this section.

Another special case occurs if I defines a biinvariant metric on G, i.e. the inner product $\langle I\cdot, \cdot\rangle$ on $\mathfrak{g}$ is Ad-invariant. Using this product to identify $\mathfrak{g}$ and $\mathfrak{g}^*$, I becomes the identity operator. Thus all motions are stationary rotations in this case. Every geodesic is either closed, or dense in the translate of a subtorus of G. In particular, there are no heteroclinic orbits. For $G = SO(3)$ this case corresponds to the round metric on S^3, in which all geodesics are closed.

Part II

Symplectic Manifolds

Chapter 5

Linear symplectic geometry

5.1 Symplectic vector spaces

Definition. A *symplectic vector space* (V^{2n}, ω) is a vector space V with a nondegenerate skew-symmetric bilinear form ω. Here nondegenerate means that $v \mapsto \omega(v, \cdot)$ defines an isomorphism $V \mapsto V^*$.

A linear map $\Psi : (V_1, \omega_1) \to (V_2, \omega_2)$ between symplectic vector spaces is called *symplectic* if $\Psi^*\omega_2 \equiv \omega_2(\Psi\cdot, \Psi\cdot) = \omega_1$.

If U is a vector space, $U \oplus U^*$ carries a *canonical symplectic structure* ω_{st} defined by

$$\omega_{\mathrm{st}}\Big((u, u^*), (v, v^*)\Big) := v^*(u) - u^*(v).$$

In coordinates q_i on U and dual coordinates p_i on U^*, the standard form is given by

$$\omega_{\mathrm{st}} = \sum dq_i \wedge dp_i.$$

The *ω-orthogonal complement* of a linear subspace $W \subset V$ is

$$W^{\perp_\omega} := \{v \in V \bigm| \omega(v, w) = 0 \text{ for all } w \in W\}.$$

Note that $\dim W + \dim W^{\perp_\omega} = 2n$, but $W \cap W^{\perp_\omega}$ may not be $\{0\}$. W is called

- *symplectic* if $W \cap W^{\perp_\omega} = \{0\}$;
- *isotropic* if $W \subset W^{\perp_\omega}$;
- *coisotropic* if $W^{\perp_\omega} \subset W$;
- *Lagrangian* if $W^{\perp_\omega} = W$.

Note that $\dim W$ is even for W symplectic, $\dim W \leq n$ for W isotropic, $\dim W \geq n$ for W coisotropic, and $\dim W = n$ for W Lagrangian.

The following obvious lemma will turn out to be very useful:

Lemma 5.1. *Let W be a linear subspace of a symplectic vector space (V^{2n}, ω). Then*

(i) $(W^{\perp_\omega})^{\perp_\omega} = W$;

(ii) $\big(W/(W \cap W^{\perp_\omega}), \omega\big)$ *is a symplectic vector space.*

Consider a subspace W of a symplectic vector space (V, ω) and let

$$N := W \cap W^{\perp_\omega}.$$

Choose subspaces $V_1 \subset W$, $V_2 \subset W^{\perp_\omega}$ and an isotropic subspace $V_3 \subset (V_1 \oplus V_2)^{\perp_\omega}$ such that

$$W = V_1 \oplus N, \quad W^{\perp_\omega} = N \oplus V_2, \quad (V_1 \oplus V_2)^{\perp_\omega} = N \oplus V_3.$$

The decomposition

$$V = V_1 \oplus N \oplus V_2 \oplus V_3$$

provides a symplectic isomorphism

$$(V, \omega) \to (W/N, \omega) \oplus (W^{\perp_\omega}/N, \omega) \oplus (N \oplus N^*, \omega_{\mathrm{st}}),$$
$$v_1 + n + v_2 + v_3 \mapsto \Big(v_1, v_2, (n, -i_{v_3}\omega)\Big).$$

The following lemma provides a linear normal form for a symplectic vector space with a linear subspace:

Lemma 5.2. *(linear normal form).*

Every symplectic vector space (V^{2n}, ω) possesses a symplectic basis $e_1, f_1, \ldots, e_n, f_n$ *satisfying $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$, $\omega(e_i, e_j) = \delta_{ij}$. Moreover, given a subspace $W \subset V$, the basis can be chosen such that*

- $W = \mathit{span}\{e_1, \ldots, e_{k+l}, f_1, \ldots, f_k\}$;
- $W^{\perp_\omega} = \mathit{span}\{e_{k+1}, \ldots, e_n, f_{k+l+1}, \ldots, f_n\}$;
- $W \cap W^{\perp_\omega} = \mathit{span}\{e_{k+1}, \ldots, e_{k+l}\}$.

Proof. Let $N := W \cap W^{\perp_\omega}$, $l := \dim N$ and $2k + l := \dim W$. Pick $e_1 \in W \setminus N$. There exists $f_1 \in W \setminus N$ such that $\omega(e_1, f_1) = 1$. Now intersect V, W and $W^{\perp_\omega}$ with $\mathit{span}\{e_1, f_1\}\perp_\omega$ and continue. Thus we find linearly independent $e_1, f_1, \ldots, e_k, f_k \in W \setminus N$ satisfying $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$, $\omega(e_i, e_j) = \delta_{ij}$.

In a similar way, construct $e_{k+l+1}, f_{k+l+1}, \ldots, e_n, f_n \in W\perp_\omega \setminus N$. Set $V_1 := \mathit{span}\{e_1, f_1, \ldots, e_k, f_k\}$ and $V_2 := \mathit{span}\{e_{k+l+1}, f_{k+l+1}, \ldots, e_n, f_n\}$. Note that $(V_1 \oplus V_2)^{\perp_\omega}$ is a symplectic subspace containing N.

Next pick any nonzero $e_{k+1} \in N$. There exists $f_{k+1} \in (V_1 \oplus V_2)^{\perp_\omega}$ such that $\omega(e_{k+1}, f_{k+1}) = 1$. Continuing like this, we find a linearly independent $e_{k+1}, \dots, e_{k+l} \in N$ and $f_{k+1}, \dots, f_{k+l} \in (V_1 \oplus V_2)^{\perp_\omega}$ satisfying $\omega(e_i, e_j) = \omega(f_i, f_j) = 0$, $\omega(e_i, e_j) = \delta_{ij}$. ☐

In particular, we get the following normal forms:

- $W = \text{span}\{e_1, f_1, \dots, e_k, f_k\}$ if W is symplectic;
- $W = \text{span}\{e_1, \dots, e_k\}$ if W is isotropic;
- $W = \text{span}\{e_1, \dots, e_n, f_1, \dots, f_k\}$ if W is coisotropic;
- $W = \text{span}\{e_1, \dots, e_n\}$ if W is Lagrangian.

Lemma 5.2 reduces the study of symplectic vector spaces to the *standard symplectic space* $(\mathbb{R}^{2n}, \omega_{\text{st}} = \sum dq_i \wedge dp_i)$.

5.2 Complex structures

A *complex structure* on a vector space V is an endomorphism $J : V \to V$ with $J^2 = -\mathbb{1}$. A complex structure J is called *compatible* with a symplectic form ω if

$$g_J := \omega(\cdot, J\cdot)$$

is an inner product (i.e. symmetric and positive definite).

Lemma 5.3. *Let (V, ω) be a symplectic vector space. There exists a natural continuous map from the space of all inner products to the space of all compatible complex structures which maps each induced inner product g_J to J. Thus the space of compatible complex structures is nonempty and contractible.*

Proof. An inner product g defines an isomorphism $A : V \to V$ via $\omega(\cdot, \cdot) = g(A\cdot, \cdot)$. Skew-symmetry of ω implies $A^T = -A$. Recall that each positive definite operator P possesses a unique positive definite square root $\sqrt{P}$, and $\sqrt{P}$ commutes with every operator with which P commutes. So we can define

$$J_g := (AA^T)^{-\frac{1}{2}} A.$$

It follows that $J_g^2 = -\mathbb{1}$, and

$$\omega(\cdot, J\cdot) = g(\sqrt{AA^T}\cdot, \cdot)$$

is an inner product. Continuity of the mapping $g \mapsto J_g$ follows from continuity of the square root. If $g = g_J$ for some J then $A = J = J_g$.

Contractibility follows from convexity of the space of inner products and the following general fact: If $f : X \to Y$, $g : Y \to X$ are continuous maps between topological spaces satisfying $f \circ g = \mathbb{1}$, then a contraction h_t of X induces a contraction $f \circ h_t \circ g$ of Y. ☐

On $(\mathbb{R}^{2n}, \omega_{\mathrm{st}} = dq \wedge dp)$ the *standard complex structure*

$$J_{\mathrm{st}} := \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$$

induces the Euclidean inner product $\sum(dq_i \otimes dq_i + dp_i \otimes dp_i)$.

A *Hermitian structure* on a vector space V^{2n} is a compatible pair (ω, J) of a symplectic and a complex structure. This induces a Hermitian product $(\cdot,\cdot)$ on the complex vector space (V, J) via

$$(v, w) := \langle v, w\rangle - i\omega(v, w),$$

where $\langle\cdot,\cdot\rangle = \omega(\cdot, J\cdot)$. (Check that $(Jv, w) = i(v, w)$ and $(w, v) = \overline{(v, w)}$). Conversely, a Hermitian metric $(\cdot,\cdot)$ on a complex vector space (V, J) induces a compatible symplectic form via

$$\omega := -\Im(\cdot,\cdot).$$

Note that the space of compatible symplectic structures on a complex vector space (V, J) is convex, in particular contractible. So we can summarize the above discussion as follows: *The forgetful maps $(\omega, J) \to \omega$, respectively $(\omega, J) \to J$, define fibrations with contractible fibres from the space of Hermitian structures onto the space of symplectic, respectively complex, structures.*

5.3 The symplectic group

Let $\omega = \omega_{\mathrm{st}}$ and $J = J_{\mathrm{st}}$ be the standard symplectic and complex structure on $\mathbb{R}^{2n}$ and $\langle\ ,\ \rangle$ the Euclidean inner product.

Definition. The *symplectic group* $Sp(2n)$ is the set of all linear symplectic maps of $(\mathbb{R}^{2n}, \omega_{\mathrm{st}})$, thus

$$Sp(2n) = \{\Psi \in GL(2n, \mathbb{R}) \mid \Psi^*\omega = \omega\} = \{\Psi \in GL(2n, \mathbb{R}) \mid \Psi^T J \Psi = J\}.$$

Remark. $Sp(2n)$ is *not* the "symplectic group" $Sp(n)$ considered in Lie group theory. E.g., the latter is compact, while "our" symplectic group is not.

Lemma 5.4. *(i) $Sp(2n)$ is a subgroup of $GL(2n, \mathbb{R})$ which is closed under transposition.*

(ii) Each symplectic matrix has determinant 1.

(iii) If λ is an eigenvalue of a symplectic matrix then λ^{-1}, $\bar\lambda$ and $\bar\lambda^{-1}$ are also eigenvalues.

(iv) If two eigenvalues λ, μ of a symplectic matrix satisfy $\lambda\mu \neq 1$ then the eigenspaces are ω-orthogonal.

Proof. (i) The subgroup property is obvious. The transposition property follows by inverting $\Psi^T J \Psi = J$.

(ii) This holds because a symplectic map preserves the standard volume form $\frac{1}{n!}\omega^n$.

(iii) Apply $\Psi^T J$ to $\Psi v = \lambda v$ to obtain $Jv = \lambda \Psi^T J v$, so λ^{-1} is an eigenvalue of Ψ^T hence of Ψ. The rest is clear.

(iv) If $\Psi v = \lambda v$, $\Psi w = \mu w$ then $\omega(v, w) = \omega(\Psi v, \Psi w) = \lambda\mu\, \omega(v, w)$. $\square$

Identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ by identifying multiplication by J in $\mathbb{R}^{2n}$ with multiplication by i in $\mathbb{C}^n$. Vectors $(x, y) \in \mathbb{R}^{2n}$ correspond to $x + iy \in \mathbb{C}^n$. A complex linear operator $X + iY$ corresponds to

$$(X + iY)(x + iy) = (Xx - Yy) + i(Xy + Yx) = \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

and the operator $X + iY$ is unitary iff

$$\mathbb{1} = (X + iY)(X^T - iY^T) = (XX^T + YY^T) + i(YX^T - XY^T).$$

This allows us to consider $GL(n, \mathbb{C})$ and $U(n)$ as subgroups of $GL(2n, \mathbb{R})$.

Lemma 5.5.

$$Sp(2n) \cap O(2n) = O(2n) \cap GL(n, \mathbb{C}) = GL(n, \mathbb{C}) \cap Sp(2n) = U(n).$$

Proof. It is easy to see that any two of the three conditions

$$\begin{aligned} \Psi \in Sp(2n) &\iff \Psi^T J \Psi = J, \\ \Psi \in O(2n) &\iff \Psi \Psi^T = \mathbb{1}, \\ \Psi \in GL(n, \mathbb{C}) &\iff \Psi J = J \Psi \end{aligned}$$

imply the third one. If $\Psi \in O(2n) \cap GL(n, \mathbb{C})$ it is of the form $\Psi = \left(\begin{smallmatrix} X & -Y \\ Y & X \end{smallmatrix}\right)$ with $XX^T + YY^T = \mathbb{1}$ and $YX^T - XY^T = 0$, so $\Psi \in U(n)$. $\square$

Proposition 5.6. *Every symplectic matrix has a unique decomposition*

$$\Psi = PQ,$$

where P is symplectic and positive definite, and $Q \in U(n)$. $(P^s Q)_{s \in [0,1]}$ defines a deformation retraction of $Sp(2n)$ onto $U(n)$.

The proof is based on the following lemma:

Lemma 5.7. *If P is symplectic and positive definite then P^s is symplectic for every real number $s > 0$.*

Proof. Decompose $\mathbb{R}^{2n}$ into eigenspaces V_λ of P to eigenvalues λ. Then V_λ is the eigenspace of P^s to the eigenvalue λ^s. For $v \in V_\lambda$ and $w \in V_\mu$,

$$\omega(P^s v, P^s w) = (\lambda\mu)^s \omega(v, w).$$

Now either $\lambda\mu = 1$; or $\lambda\mu \neq 1$, in which case $\omega(v, w) = 0$ by Lemma 5.4. In either case $\omega(P^s v, P^s w) = \omega(v, w)$. This proves that P^s is symplectic because the eigenvectors generate $\mathbb{R}^{2n}$. □

Proof of Proposition 5.6. Any invertible matrix Ψ can be uniquely decomposed as $\Psi = PQ$ where $P = \sqrt{\Psi\Psi^T}$ is positive definite and Q orthogonal. If Ψ is symplectic then P is symplectic by Lemma 5.7. Thus Q is symplectic and therefore unitary by Lemma 5.5. The continuous retraction $P^s Q$ stays in the symplectic group again by Lemma 5.7. □

Problem 5.1. Prove that the determinant $\det : U(n) \to S^1$ induces an isomorphism between $\pi_1\big(U(n)\big)$ and $\pi_1(S^1) \cong \mathbb{Z}$. In particular, $\pi_1\big(Sp(2n)\big) \cong \mathbb{Z}$.

Problem 5.2. (cf. [21]). Find an explicit diffeomorphism from $Sp(2)$ to an open solid torus. Describe the regions of real / complex eigenvalues.

5.4 Symplectic vector bundles

Definition. Let $E^{2n} \to M$ be a real vector bundle of rank $2n$ over a manifold.

A *symplectic structure* on E is a smooth section ω in the bundle $\Lambda^2 E^* \to M$ such that each $\omega_x \in \Lambda^2 E_x^*$ is a linear symplectic form.

A *complex structure* on E is a smooth section J in the bundle $Hom(E, E) \to M$ such that each $J_x \in Hom(E_x, E_x)$ satisfies $J_x^2 = -\mathbb{1}$. A complex structure J is called *compatible* with a symplectic structure ω if $\omega(\cdot, J\cdot)$ defines an inner product on E.

A *Hermitian structure* on E is a compatible pair (ω, J) of a symplectic and a complex structure, or equivalently, a complex structure with a Hermitian product.

The following proposition is an easy consequence of Lemma 5.3, where spaces of sections are equipped with any reasonable topology, e.g. the C^∞_{loc} topology:

Proposition 5.8. *(i) The space of compatible complex structures on a symplectic vector bundle (E, ω) is nonempty and contractible.*

(ii) The space of compatible symplectic structures on a complex vector bundle (E, J) is nonempty and convex.

(iii) A Hermitian vector bundle (E, ω, J) possesses local unitary trivializations*, i.e. trivializations $\Psi : E|_U \mapsto U \times \mathbb{C}^n$ satisfying $\Psi_x \circ J_x = J_{\mathrm{st}} \circ \Psi_x$ and $\Psi_x^* \omega_{\mathrm{st}} = \omega_x$ for each $x \in U$.*

This proposition implies that the homotopy theories of symplectic, complex and Hermitian vector bundles are the same. In particular, obstructions to trivialization of a symplectic vector bundle (E, ω) are measured by the *Chern classes* $c_k(E, \omega) = c_k(E, J)$ for any compatible complex structure J.

Problem 5.3. Let $E \to \Sigma$ be a Hermitian vector bundle over a compact Riemann surface Σ with boundary $\partial\Sigma$.

(i) Show: If $\partial\Sigma \neq \emptyset$ then E possesses a (global) unitary trivialization.

(ii) If $\partial\Sigma = \emptyset$ let $\Sigma = \Sigma_1 \cup \Sigma_2$ be a decomposition into two surfaces along their common boundary circle $C = \partial\Sigma_1 = -\partial\Sigma_2$, oriented as the boundary of Σ_1. Let $\Psi_i : E|_{\Sigma_i} \to \Sigma_i \times \mathbb{C}^n$ be unitary trivializations. Prove that $c_1(E)$ equals the degree of the map

$$\det(\Psi_1 \circ \Psi_2^{-1}|_C) : C \mapsto S^1.$$

Combined with the following problem, this problem provides a hands-on approach to the first Chern class in arbitrary manifolds.

Problem 5.4. (i) Prove that every singular homology class $\sigma \in H_k(M; \mathbb{Z})$ on a manifold M can be represented by a smooth map $f : N \to M$ of a smooth closed k-manifold N to M.

(ii) Prove that every 2-homology class $\sigma \in H_2(M; \mathbb{Z})$ on a manifold of dimension ≥ 4 can be represented by an *embedded* closed surface. Hint: This is easy except for $\dim M = 4$.

(iii) Show that the genus of an embedded surface representing a 2-homology class in a closed 4-manifold can always be chosen arbitrarily high. Find examples where the genus cannot be chosen arbitrarily low.

Remark. The *minimal genus* of an embedded closed surface representing a homology class in a closed 4-manifold is a fascinating and highly nontrivial problem. The following conjecture was only recently proved in full generality by Ozsváth and Szabó [17], building on work of Kronheimer, Mrowka and Taubes and using Seiberg-Witten theory:

Thom Conjecture: *Every symplectic 2-dimensional closed submanifold of a closed symplectic 4-manifold minimizes the genus of an embedded surface in its homology class.*

We end this section with a normal form for subbundles of symplectic vector bundles:

Proposition 5.9. *Let (E^{2n}, ω) be a rank $2n$ symplectic vector bundle and $W^{2k+l} \subset E$ a rank $2k+l$ subbundle such that $N := W \cap W^{\perp_\omega}$ has constant rank l. Then*

$$(E, \omega) \cong (W/N, \omega) \oplus (W^{\perp_\omega}/N, \omega) \oplus (N \oplus N^*, \omega_{\mathrm{st}}).$$

Proof. Pick a compatible almost complex structure J on (E, ω). Then

$$V_1 := W \cap JW, \quad V_2 := W^{\perp_\omega} \cap JW^{\perp_\omega}, \quad V_3 := JN$$

are smooth subbundles of E. As in the discussion preceding Lemma 5.2, the decomposition

$$E = V_1 \oplus N \oplus V_2 \oplus V_3$$

yields the desired isomorphism. $\square$

Chapter 6

Symplectic manifolds

6.1 Definition

Definition. A *symplectic atlas* on a manifold is an atlas all of whose transition maps are symplectic. A *symplectic manifold* is a manifold M^{2n} with a symplectic atlas.

Since Hamilton's equations are invariant under symplectic transformations, they are intrinsically defined on a symplectic manifold: Every Hamiltonian $H : M \to \mathbb{R}$ induces a Hamiltonian vector field X_H on M which has the form

$$\sum_i \Big(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} \Big)$$

in any symplectic coordinates (q_i, p_i).

A symplectic manifold carries a canonical 2-form ω obtained by pulling back the standard form $\omega_{st} = \sum dq_i \wedge dp_i$ on $\mathbb{R}^{2n}$ under any symplectic coordinate map. The form ω is closed because ω_{st} is exact. Moreover, ω is *nondegenerate* in the sense that $v \mapsto \omega_x(v, \cdot)$ defines an isomorphism $T_xM \to T_x^*M$.

Definition. A *symplectic manifold* (M, ω) is a manifold M^{2n} with a closed nondegenerate 2-form ω.

We have seen that the first definition implies the second one. The converse is a consequence of the following theorem (take as symplectic atlas all coordinate maps ϕ as in the theorem):

Darboux' Theorem: Let (M, ω) be a manifold with a closed nondegenerate 2-form. Then every $x \in M$ possesses a coordinate neighbourhood U and a coordinate map $\phi : U \to V \subset \mathbb{R}^{2n}$ such that $\phi^*\omega_{st} = \omega$.

Darboux' Theorem will be proved in the next chapter. But let us first see some examples of symplectic manifolds.

Example 6.1. (cotangent bundles).

Let $T^*Q \xrightarrow{\pi} Q$ be the cotangent bundle of a manifold Q^n. Recall from Part I that the expression $\sum p_i dq_i$ is independent of coordinates q_i on Q and dual coordinates p_i on T_q^*Q and thus defines a *canonical 1-form* $\lambda_{\rm st}$ on T^*Q. Intrinsically,

$$(\lambda_{\rm st})_{(q,p)} \cdot v = \langle p, T_{(q,p)}\pi \cdot v\rangle \quad \text{for } v \in T_{(q,p)}T^*Q,$$

where $\langle \ , \ \rangle$ is the pairing between T_q^*Q and T_qQ. The 2-form $\omega_{\rm st} := -d\lambda_{\rm st}$ is clearly closed, and the coordinate expression $\omega_{\rm st} = \sum dq_i \wedge dp_i$ shows that it is also nondegenerate. So $\omega_{\rm st}$ defines a *canonical symplectic form* on T^*Q. The standard form on $\mathbb{R}^{2n}$ is a particular case of this construction.

Example 6.2. (surfaces). A symplectic form on a surface is just an area form. So a surface admits a symplectic structure if and only if it is orientable.

Example 6.3. (complex projective space).

The *complex projective space* $\mathbb{C}P^n = (\mathbb{C}^{n+1} \setminus 0)/(\mathbb{C} \setminus 0)$ carries the following Hermitian metric: The tangent space at $[z] = [z_0 : \cdots : z_n] \in \mathbb{C}P^n$ can be identified with

$$T_{[z]}\mathbb{C}P^n = \{v \in \mathbb{C}^{n+1} \mid (v, z) = 0\},$$

where $(v, w) = \sum v_j \bar{w}_j$ is the standard Hermitian product on $\mathbb{C}^{n+1}$. Define the Hermitian metric on $\mathbb{C}P^n$ as the standard product of the projection onto $T_{[z]}\mathbb{C}P^n$:

$$\begin{aligned}(v, w)_{[z]} &:= \Big(v - \frac{(v,z)}{\|z\|^2} z, w - \frac{(w,z)}{\|z\|^2} z\Big) \\ &= \sum_j v_j \bar{w}_j - \sum_{j,k} \frac{\bar{z}_j z_k v_j \bar{w}_k}{\|z\|^2},\end{aligned}$$

or equivalently,

$$(\ ,\ )_{[z]} = \sum_j \frac{dz_j \otimes d\bar{z}_j}{\|z\|^2} - \sum_{j,k} \frac{\bar{z}_j z_k dz_j \otimes d\bar{z}_k}{\|z\|^4}.$$

Define a nondegenerate 2-form on $\mathbb{C}P^n$, the *Fubini-Study form*, by

$$\omega_{FS} := -\Im(\ ,\ ).$$

From

$$\begin{aligned}(\bar{z}_j z_k dz_j \wedge d\bar{z}_k + \bar{z}_k z_j dz_k \wedge d\bar{z}_j)(v, w) &= \bar{z}_j z_k (v_j \bar{w}_k - w_j \bar{v}_k) + \bar{z}_k z_j (v_k \bar{w}_j - w_k \bar{v}_j) \\ &= 2i\Im(\bar{z}_j z_k v_j \bar{w}_k + \bar{z}_k z_j v_k \bar{w}_j)\end{aligned}$$

we see that

$$(\omega_{FS})_{[z]} = \frac{i}{2}\Big(\sum_j \frac{dz_j \wedge d\bar{z}_j}{\|z\|^2} - \sum_{j,k} \frac{\bar{z}_j z_k dz_j \wedge d\bar{z}_k}{\|z\|^4}\Big).$$

Problem 6.1. Prove that the form ω_{FS} on $\mathbb{C}P^n$ is closed. Show that ω_{FS} is the unique symplectic form on $\mathbb{C}P^n$ which is $U(n+1)$-invariant and whose integral over a line $\mathbb{C}P^1 \subset \mathbb{C}P^n$ equals

$$\int_{\mathbb{C}P^1} \omega_{FS} = \pi.$$

Definition. A map $f : (M_1, \omega_1) \to (M_2, \omega_2)$ between symplectic manifolds is called *symplectic* if $f^*\omega_2 = \omega_1$. A symplectic diffeomorphism is called *symplectomorphism* by mathematicians, and *canonical transformation* by physicists.

Problem 6.2. Prove that a symplectic map is necessarily an immersion.

6.2 Almost complex structures

Definition. An *almost complex structure* on a manifold is a complex structure J on the vector bundle $TM \to M$. It is called *compatible* with a symplectic form ω if $\omega(\cdot, J\cdot)$ is a Riemannian metric.

In view of Proposition 5.8 there exist compatible almost complex structures on any symplectic manifold, and they form a contractible space.

An almost complex structure J on M is called *integrable* if there exist local *complex coordinates* $\phi : M \supset U \to V \in \mathbb{C}^n$ such that $T_x\phi \circ J_x = J_{\mathrm{st}} \circ T_x\phi$ for all $x \in U$.

Problem 6.3. Show that for an almost complex structure J the expression

$$N_J(X, Y) := [JX, JY] - J[JX, Y] - J[X, JY] - [X, Y]$$

for vector fields X, Y defines a skew-symmetric 2-tensor, and $N_J = 0$ if J is integrable.

N_J is called the *Nijenhuis tensor*. The *Newlander-Nirenberg Theorem* states that J is integrable if and only if $N_J = 0$.

Definition. A symplectic manifold (M, ω, J) with an integrable compatible almost complex structure is called *Kähler manifold*. Equivalently, a Kähler manifold is a complex manifold (M, J) with a Hermitian metric $(\ ,\ )$ such that the *Kähler form* $\omega = -\Im(\ ,\ )$ is a closed 2-form.

The basic example of a Kähler manifold is the complex projective space. Complex submanifolds of Kähler manifolds are again Kähler manifolds. In particular, all smooth projective varieties are Kähler manifolds.

Example 6.4. A particular case of a Kähler manifold is a *Stein manifold* (M, J, ϕ). This is a complex manifold (M, J) with a function $\phi : M \to \mathbb{R}$ such that

- ϕ is *exhausting*, i.e. proper and bounded from below;

- ϕ is *plurisubharmonic*, i.e. $-d(d\phi \circ J)(\cdot, J\cdot)$ is a Riemannian metric.

Then $\omega_\phi := -d(d\phi \circ J)$ is a Kähler symplectic form compatible with J. Examples of Stein manifolds are $(\mathbb{C}^n, J_{\rm st}, \phi = \frac{1}{4}\|z\|^2)$ and $(T^*Q, J, \phi = \frac{1}{2}\|p\|^2)$, where $\|p\|$ is the norm dual to a Riemannian metric on Q, and J is a suitable complex structure on T^*Q. In both examples the induced symplectic form is the standard symplectic form.

Problem 6.4. Show that a Riemannian metric g on a manifold Q induces a natural almost complex structure J_g on T^*Q compatible with $\omega_{\rm st}$, which interchanges the horizontal and vertical subspaces defined by the Levi-Civita connection. In general, this almost complex structure is not integrable. Find the conditions on the metric g such that J_g is integrable.

While cotangent bundles and Kähler manifolds provide obvious examples of symplectic manifolds, it is not obvious how to go beyond them. The first example of a compact symplectic manifold that is not Kähler was presented by Thurston in 1976 ([20]):

Example 6.5. (Thurston's manifold). Consider $\Gamma = \mathbb{Z}^4$ with the noncommutative group operation

$$a * b := a + L_a b, \quad a, b \in \mathbb{Z}^4, \quad L_q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & a_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Associativity follows from

$$L_a L_b = L_{a+b} = L_{a*b},$$

and the inverse is given by

$$a^{-1} = -L_{-a} a = (-a_1, -a_2, -a_3 + a_2 a_4, -a_4)^T.$$

Γ acts on $\mathbb{R}^4$ via

$$\rho_a(x) = a + L_a x,$$

and this action preserves the standard symplectic form $\omega_{\rm st} = dx_1 \wedge dx_2 + dx_3 \wedge dx_4$. Since

$$\|\rho_a(x) - x\|^2 = a_1^2 + a_2^2 + (a_3 + a_2 x_4)^2 + a_4^2 \geq 1$$

for all $0 \neq a \in \mathbb{Z}^4$ and $x \in \mathbb{R}^4$, the action is properly discontinuous. So the quotient

$$(M := \mathbb{R}^4/\Gamma, \omega_{\rm st})$$

is a compact symplectic 4-manifold with fundamental group $\pi_1(M) = \Gamma$. Let $[\ ,\ ]$ denote the commutator in Γ and e_i the standard basis of $\mathbb{Z}^4$. From $[e_2, e_4] = e_3$ it easily follows that the commutator subgroup of Γ is

$$[\Gamma, \Gamma] = \mathbb{Z} e_3.$$

So the first homology group of M is

$$H_1(M;\mathbb{Z}) = \Gamma/[\Gamma,\Gamma] \cong \mathbb{Z}^3.$$

But the odd Betti number of Kähler manifolds are even, so (M,ω) cannot be Kähler.

6.3 Normal forms

Moser's trick

The basis of all the normal form theorems is *Cartan's formula*

$$L_X\alpha = i_X d\alpha + d\, i_X\alpha$$

for a vector field X and a k-form α, which is proved in the appendix.

Now suppose we are given k-forms α_0, α_1 on a manifold M, and we are looking for a diffeomorphism $\phi : M \to M$ such that $\phi^*\alpha_1 = \alpha_0$. Moser's trick is to construct ϕ as the time-1 map of a time-dependent vector field X_t. For this, let α_t be a smooth family of k-forms connecting α_0 and α_1, and look for a vector field X_t whose flow ϕ_t satisfies

$$\phi_t^*\alpha_t \equiv \alpha_0. \tag{6.1}$$

The time-1 map $\phi = \phi_1$ the solves our problem. Now equation 6.1 follows once its linearized version

$$0 = \frac{d}{dt}\phi_t^*\alpha_t = \phi_t^*(\dot\alpha_t + L_{X_t}\alpha_t)$$

holds for every t. Inserting Cartan's formula, this reduces the problem to the algebraic problem of finding a vector field X_t that satisfies

$$\dot\alpha_t + d\, i_{X_t}\alpha_t + i_{X_t}d\alpha_t = 0. \tag{6.2}$$

Here is a first application of this method:

Theorem 6.6. *(Moser [16]). Let Ω_0, Ω_1 be volume forms on a closed manifold M. Then there exists a diffeomorphism ϕ isotopic to the identity such that $\phi^*\Omega_1 = \Omega_0$ if and only if*

$$\int_M \Omega_0 = \int_M \Omega_1.$$

Proof. The "only if" part is obvious. For the "if" part, let $\Omega_t := (1-t)\Omega_0 + \Omega_1$ be a family of volume forms connecting Ω_0 and Ω_1. The form $\dot\Omega_t = \Omega_1 - \Omega_0$ satisfies $\int_M \dot\Omega_t = 0$. So its cohomology class vanishes, and therefore $\dot\Omega_t = d\beta$ for some β. We want to solve equation 6.2,

$$0 = \dot\Omega_t + d\, i_{X_t}\Omega_t + i_{X_t}d\Omega_t = d(\beta + i_{X_t}\Omega_t).$$

Since Ω_t is nondegenerate, there exists a unique vector field X_t satisfying $\beta + i_{X_t}\Omega_t = 0$, and the time-1 map of the flow of X_t is the desired diffeomorphism ϕ. □

Combined with the classification of surfaces, Moser's theorem classifies closed symplectic 2-manifolds:

Corollary 6.7. *Two closed 2-dimensional symplectic manifolds are symplectomorphic if and only if they have the same genus and total area.*

Problem 6.5. Show that the corollary has no analog in higher dimension: Find two symplectic forms on the same smooth manifold of dimension ≥ 4 which have the same total volume but are not symplectomorphic.

A second application of Moser's trick shows that one does not obtain new symplectic forms by deformations which remain in the same cohomology class:

Theorem 6.8 (Moser's Stability Theorem)**.** *Let ω_t be a smooth family of cohomologous symplectic forms on a closed manifold. Then there exists a smooth family of diffeomorphisms ϕ_t such that $\phi_t^*\omega_t = \omega_0$.*

Proof. Since the ω_t are cohomologous, $\dot{\omega}_t$ is trivial in cohomology, so for every t there exists a 1-form β_t such that $d\beta_t = \dot{\omega}_t$. A priori β_t need not depend smoothly on t. But Hodge theory provides a unique particular choice of β_t which implies smoothness in t: Pick a Riemannian metric on the manifold M and let $d^* : \Omega^2(M) \to \Omega^1(M)$ be the L^2-adjoint of d. By Hodge theory, $\mathrm{im}(d^*) = \ker(d)^\perp$, so d is an isomorphism from $\mathrm{im}(d^*)$ to the exact 2-forms. The inverse of this isomorphism provides the particular choice for β_t.

Now we can solve equation 6.2,

$$0 = \dot{\omega}_t + d\, i_{X_t}\omega_t + i_{X_t}d\omega_t = d(\beta_t + i_{X_t}\omega_t)$$

by solving $\beta_t + i_{X_t}\omega_t = 0$, which has a unique solution X_t by nondegeneracy og ω_t. □

Problem 6.6. (cf. [15], Theorem 3.17). Prove that β_t can be chosen smoothly in t by local arguments in coordinate charts, without using Hodge theory.

Weinstein's neighbourhood theorems

Definition. An immersed submanifold W of a symplectic manifold (M, ω) is called *symplectic (isotropic, coisotropic, Lagrangian)* if $T_xW \subset T_xM$ is symplectic (isotropic, coisotropic, Lagrangian) for every $x \in W$.

All these types of submanifolds of a symplectic manifold have standard neighbourhoods. The proofs are based on the following

Lemma 6.9. *Let W be a compact submanifold of a manifold M, and let ω_0, ω_1 be symplectic forms on M which agree at all points of W. Then there exist tubular neighbourhoods U_0, U_1 of W and a diffeomorphism $\phi : U_0 \to U_1$ such that $\phi|_W = \mathbb{1}$ and $\phi^*\omega_1 = \omega_0$.*

Proof. Set $\omega_t := (1-t)\omega_0 + \omega_1$. Since $\omega_t \equiv \omega_0$ along W, ω_t are symplectic forms on some tubular neighbourhood U of W. By the relative de Rham Theorem, since $\dot{\omega}_t = \omega_1 - \omega_0$ is closed and vanishes along W, there exists a of 1-form β on U such that $\beta = 0$ along W and $d\beta = \dot{\omega}_t$ on U. As in the proof of Theorem 6.8, we solve equation 6.2 by setting $\beta + i_{X_t}\omega_t = 0$.

To apply Moser's trick, a little care is needed because U is noncompact, so the flow of X_t may not exist until time 1. However, since $\beta = 0$ along W, X_t vanishes along W. Thus there exists a tubular neighbourhood U_0 of W such that the flow $\phi_t(x)$ of X_t exists for all $x \in U_0$ and $t \in [0, 1]$, and $\phi_t(U_0) \subset U$ for all $t \in [0, 1]$. Now $\phi_1 : U_0 \to U_1 := \phi_1(U_0)$ is the desired diffeomorphism with $\phi_1^*\omega_1 = \omega_0$. □

Now we are ready for the main result of this section.

Theorem 6.10 (Normal Form Theorem)**.** *Let ω_0, ω_1 be symplectic forms on a manifold M and $W \subset M$ a compact submanifold such that $\omega_0|_W = \omega_1|_W$. Suppose that $N := \ker(\omega_0|_W) = \ker(\omega_1|_W)$ has constant rank, and the bundles $(TW^{\perp_{\omega_0}}/N, \omega_0$, $(TW^{\perp_{\omega_1}}/N, \omega_1$ over W are isomorphic as symplectic vector bundles. Then there exist tubular neighbourhoods U_0, U_1 of W and a diffeomorphism $\phi : U_0 \to U_1$ such that $\phi|_W = \mathbb{1}$ and $\phi^*\omega_1 = \omega_0$.*

Proof. By Proposition 5.9,

$$TM|_W, \omega_0) \cong (TW/N, \omega_0) \oplus (TW^{\perp_{\omega_0}}/N, \omega_0) \oplus (N \oplus N^*, \omega_{\mathrm{st}}),$$

and similarly for ω_1. By the hypotheses, the right-hand sides are isomorphic for ω_0 and ω_1. More precisely, there exists an isomorphism

$$\Psi : (TM|_W, \omega_0) \to (TM|_W, \omega_1)$$

with $\Psi|_{TW} = \mathbb{1}$. Extend Ψ to a diffeomorphism $\psi : U_0 \to U_1$ of tubular neighbourhoods such that $\psi|_W = \mathbb{1}$ and $\psi^*\omega_1 = \omega_0$ along W, and apply Lemma 6.9. □

All the normal forms are easy corollaries of this theorem.

Corollary 6.11 (Darboux' Theorem)**.** *Let (M, ω) be a manifold with a closed nondegenerate 2-form. Then every $x \in M$ possesses a coordinate neighbourhood U and a coordinate map $\phi : U \to V \subset \mathbb{R}^{2n}$ such that $\phi^*\omega_{st} = \omega$.*

Proof. This just follows from the fact that all symplectic vector spaces, hence all symplectic vector bundles over a point, are isomorphic. □

Corollary 6.12 (Symplectic Neighbourhood Theorem)**.** *Let ω_0, ω_1 be symplectic forms on a manifold M and $W \subset M$ a compact submanifold such that $\omega_0|_W = \omega_1|_W$ is symplectic, and the symplectic normal bundles $(TW^{\perp_{\omega_0}}, \omega_0)$, $(TW^{\perp_{\omega_1}}, \omega_1)$ over W are isomorphic (as symplectic vector bundles). Then there exist tubular neighbourhoods U_0, U_1 of W and a diffeomorphism $\phi : U_0 \to U_1$ such that $\phi|_W = \mathbb{1}$ and $\phi^*\omega_1 = \omega_0$.*

Corollary 6.13 (Isotropic Neighbourhood Theorem)**.** *Let ω_0, ω_1 be symplectic forms on a manifold M and $W \subset M$ a compact submanifold such that $\omega_0|_W = \omega_1|_W = 0$, and the symplectic normal bundles $(TW^{\perp_{\omega_0}}/TW, \omega_0)$, $(TW^{\perp_{\omega_1}}/TW, \omega_1)$ are isomorphic (as symplectic vector bundles). Then there exist tubular neighbourhoods U_0, U_1 of W and a diffeomorphism $\phi : U_0 \to U_1$ such that $\phi|_W = \mathbb{1}$ and $\phi^*\omega_1 = \omega_0$.*

Corollary 6.14 (Coisotropic Neighbourhood Theorem)**.** *Let ω_0, ω_1 be symplectic forms on a manifold M and $W \subset M$ a compact submanifold such that $\omega_0|_W = \omega_1|_W$ and W is coisotropic for ω_0 and ω_1. Then there exist tubular neighbourhoods U_0, U_1 of W and a diffeomorphism $\phi : U_0 \to U_1$ such that $\phi|_W = \mathbb{1}$ and $\phi^*\omega_1 = \omega_0$.*

Corollary 6.15 (Lagrangian Neighbourhood Theorem [23])**.** *Let $W \subset (M, \omega)$ be a compact Lagrangian submanifold of a symplectic manifold. Then there exist tubular neighbourhoods U of the zero section in T^*W and V of W in M and a diffeomorphism $\phi : U \to V$ such that $\psi|_W$ is the inclusion and $\phi^*\omega = \omega_{\mathrm{st}}$.*

Proof. Since W is Lagrangian, the map $v \mapsto i_v\omega$ defines an isomorphism from the normal bundle $TM/TW|_W$ to T^*W. Extend the inclusion $W \subset M$ to a diffeomorphism $\psi : U \to V$ of tubular neighbourhoods of the zero section in T^*W and of W in M. Now apply the Coisotropic Neighbourhood Theorem to the zero section in T^*W and the symplectic forms ω_{st} and $\psi^*\omega$. □

Chapter 7

Constructing symplectic manifolds

7.1 Blowing up in the complex category

The local model

The following discussion follows [15]. Consider the subset $L \subset \mathbb{C}^n \times \mathbb{C}P^{n-1}$ given by the incidence relation,

$$\begin{aligned}\tilde{\mathbb{C}}^n := L &:= \{(z,l) \in \mathbb{C}^n \times \mathbb{C}P^{n-1} \mid z \in l\} \\ &= \{\big((z_1 \dots, z_n), [w_1 : \dots : w_n]\big) \mid z_j w_k = z_k w_j \text{ for } 1 \leq j,k \leq n\}\end{aligned}$$

with the two obvious projections

$$\begin{aligned}\beta : L \to \mathbb{C}^n, \quad (z,l) \mapsto z; \\ \pi : L \to \mathbb{C}P^{n-1}, (z,l) \mapsto l.\end{aligned}$$

So $\pi : L \to \mathbb{C}P^{n-1}$ is just the tautological line bundle. Its zero section

$$E := L_0 := \beta^{-1}(0) \cong \mathbb{C}P^{n-1}$$

is called the *exceptional divisor*. Note that

$$\beta : \tilde{C}^n \setminus E \to \mathbb{C}^n \setminus 0$$

is a diffeomorphism. So $\tilde{\mathbb{C}}^n$ is obtained from $\mathbb{C}^n$ by replacing the origin by the $\mathbb{C}P^{n-1}$ of tangent lines at the origin. $\beta : \tilde{\mathbb{C}}^n \to \mathbb{C}^n$ is called the *blow-up of* $\mathbb{C}^n$ *at the origin.*

Lemma 7.1. *The normal bundle ν_E of E in $\tilde{\mathbb{C}}^n$ has first Chern class $c_1(\nu_E) = -c$, where $c \in H^2(\mathbb{C}P^{n-1};\mathbb{Z})$ is the canonical generator dual to a hyperplane. In particular, for $n = 2$, E has self-intersection number*

$$E \cdot E = \langle c_1(\nu_E), [E] \rangle = -1.$$

Proof. Since the normal bundle to E is just the line bundle $\pi : L \to \mathbb{C}P^{n-1}$, we have to show $c_1(L) = -c$. Since $c_1(L)$ is determined by its restriction to a line $\mathbb{C}P^1 \subset \mathbb{C}P^{n-1}$ and $L|_{\mathbb{C}P^1}$ is the tautological line bundle over $\mathbb{C}P^1$, it suffices to treat the case $n = 2$. Then $c_1(L) = kc$, where k is the self-intersection number of the zero section in L.

Now L does not possess holomorphic sections, But its dual, the *canonical line bundle* $L^* \to \mathbb{C}P^{n-1}$, has holomorphic sections s induced by homogeneous linear polynomials $\sum a_j z_j$ on $\mathbb{C}^n$:

$$s(l) : z \mapsto \sum a_j z_j, \quad l \in \mathbb{C}P^{n-1}, z \in l.$$

The section $s(l) : (z_1, z_2) \mapsto z_1$ of $L^* \to \mathbb{C}P^1$ has a unique transverse zero $l = [0 : 1]$. Since every intersection point of complex submanifolds counts with $+1$, this shows $c_1(L^*) = c$, hence $c_1(L) = -c$. □

Lemma 7.2. *Every biholomorphic map $\psi : \mathbb{C}^n \to \mathbb{C}^n$ with $\psi(0) = 0$ lifts uniquely to a biholomorphic map $\tilde{\psi} : \tilde{\mathbb{C}}^n \to \tilde{\mathbb{C}}^n$ with $\tilde{\psi}(E) = E$.*

Proof. Define

$$\tilde{(}z, l) := \begin{cases} \big(\psi(z), [\psi(z)]\big) & \text{if } z \neq 0, \\ (0, T_0\psi \cdot l) & \text{if } z = 0. \end{cases}$$

To see that this map is continuous, consider a sequence $z_k \to 0$ such that $[z_k] \to l$, i.e. $v_k := z_k/|z_k| \to v \neq 0$, $[v] = l$. Then

$$\lim \frac{\psi(z_k)}{|}z_k| = \lim \frac{\psi(|z_k|v_k)}{|z_k|} = \lim \frac{\psi(|z_k|v}{|z_k|} = T_0\psi \cdot v \neq 0,$$

and so $[\psi(z_n)] = [\psi(z_n)/|z_n|] \to [T_0\psi \cdot v] = T_0\psi \cdot l$. Similar arguments show that $\tilde{\psi}$ is smooth and holomorphic. □

This lemma allows us to define the *blow-up* $\pi : \tilde{M} \to M$ *of a complex manifold M at a point* p: Choose a holomorphic coordinate chart mapping p to the origin and blow up the origin in $\mathbb{C}^n$. The lemma ensures that the result in independent of the coordinate chart.

Lemma 7.3. *The blow-up $\tilde{M}$ is oriented diffeomorphic to the connected sum $M\#\overline{\mathbb{C}P^n}$, where $\overline{\mathbb{C}P^n}$ is $\mathbb{C}P^n$ with the opposite orientation.*

Proof. Recall the *connected sum* $M_1\#M_2$ of two oriented manifolds: Cut out little balls B_i from M_i and identify collar neighbourhoods of the boundaries of $M_i \setminus B_i$ via an orientation preserving diffeomorphism that maps the inner sphere of the collar of M_1 to the outer sphere of the collar of M_2. This construction is canonical if the diffeomorphism identifying the collars is required to be isotopic to the composition of an orientation reversing orthogonal map of the unit sphere in $\mathbb{R}^n$ with a reflection on that sphere.

Now the connected sum of M with either $\mathbb{C}P^n$ and $\overline{\mathbb{C}P^n}$ replaces a point by a $\mathbb{C}P^{n-1}$. The only question is which one gives the correct normal bundle of $\mathbb{C}P^{n-1}$. In the blow-up the normal bundle is the tautological bundle $L \to \mathbb{C}P^{n-1}$. The normal bundle of a hyperplane $\mathbb{C}P^{n-1} \subset \mathbb{C}P^n$ is the canonical bundle $L^* \to \mathbb{C}P^{n-1}$ (because $\mathbb{C}P^{n-1}$ is the zero set of a section in the canonical bundle $L^* \to \mathbb{C}P^n$). Hence the correct choice is $\overline{\mathbb{C}P^n}$. □

Example 7.4. $\tilde{\mathbb{C}}P^2 = \mathbb{C}P_2 \# \overline{\mathbb{C}P^2}$.

The blow-up $\tilde{\mathbb{C}}P^2$ of $\mathbb{C}P^2$ at one point p_0 fibres over the exceptional divisor with fibre $\mathbb{C}P^1$,

$$\mathbb{C}P^1 \to \tilde{\mathbb{C}}P^2 \xrightarrow{\pi} E \cong \mathbb{C}P^1.$$

The projection π maps each point $p \neq p_0$ of $\mathbb{C}P^2$ to the unique line in $\mathbb{C}P^2$ through p and p_0, and is the identity on E. More explicitly, the embedding $\tilde{\mathbb{C}}^n \subset \mathbb{C}^2 \times \mathbb{C}P^1$ extends to an embedding

$$\tilde{\mathbb{C}}P^2 = \{\big([z_0 : z_1 : z_2], [w_1 : w_2]\big) \mid z_z w_2 = z_2 w_1\} \subset \mathbb{C}P^2 \times \mathbb{C}P^1$$

with canonical projections $\beta : \tilde{\mathbb{C}}P^2 \to \mathbb{C}P^2$ and $\pi : \tilde{\mathbb{C}}P^2 \to E = \{p + 0\} \times \mathbb{C}P^1$.

Now S^2-bundles over S^2 are classified by

$$\pi_1\Big(Diff^+(S^2)\Big) = \pi_1\Big(SO(3)\Big) = \mathbb{Z}_2$$

according to *Smale's Theorem* [19] that the group $Diff^+(S^2)$ of orientation preserving diffeomorphisms of S^2 deformation retracts onto $SO(3)$. So there is only the trivial bundle $S^2 \times S^2$ and one nontrivial bundle. I claim that $\tilde{\mathbb{C}}P^2$ is the nontrivial bundle.

To see this, recall that the intersection form $Q : H_2(M;\mathbb{Z}) \times H_2(M;\mathbb{Z}) \to \mathbb{Z}$ of a closed 4-manifold is a diffeomorphism invariant up to $Q \mapsto AQA^T$, $A \in SL(2,\mathbb{Z})$. In particular, the *signature* (number of positive minus number of negative eigenvalues) and the *type* of Q are diffeomorphism invariants. Here Q is of even type if $\langle x, Qx\rangle$ is even for all $x \in \mathbb{Z}^2$ and odd otherwise. The intersection form of $S^2 \times S^2$ is

$$Q_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

in the basis $S^2 \times \mathrm{pt}$ and $\mathrm{pt} \times S^2$, which has signature zero and even type. The intersection form of $\tilde{\mathbb{C}}P^2$ has intersection form

$$Q_1 = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$$

in the basis E and the fibre F, or

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

in the basis E and $E+F$. This form has signature zero and odd type, so $\tilde{\mathbb{C}}P^2$ is not diffeomorphic to $S^2 \times S^2$.

The embedding $\tilde{\mathbb{C}}P^2 \subset \mathbb{C}P^2 \times \mathbb{C}P^1$ above shows that $\pi : \tilde{\mathbb{C}}P^2 \to \mathbb{C}P^1$ is the projectivization of the direct sum of the tautological line bundle $L \to \mathbb{C}P^1$ with the trivial line bundle $\mathbb{C}$,

$$\tilde{\mathbb{C}}P^2 = \P[L \oplus \mathbb{C}] \to \mathbb{C}P^1.$$

More generally, we can replace L by other line bundles over $\mathbb{C}P^1$. Since line bundles over surfaces are classified by their first Chern class, each line bundle over $\mathbb{C}P^1$ is isomorphic to $L^{\otimes d}$ for some $d \in \mathbb{Z}$. The zero section E of this line bundle has self-intersection $E \cdot E = \langle c_1(L^d), [E]\rangle = -d$, so $\P[L^{\otimes d} \oplus \mathbb{C}]$ has intersection form

$$Q_d = \begin{pmatrix} -d & 1 \\ 1 & 0 \end{pmatrix}$$

in the basis E, F. From

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -d & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix} = \begin{pmatrix} -d+2a & 1 \\ 1 & 0 \end{pmatrix}$$

we see that

$$Q_d \sim \begin{cases} Q_0 & \text{for } d \text{ even,} \\ Q_1 & \text{for } d \text{ odd.} \end{cases}$$

This shows the bundle isomorphisms

$$\P[L^{\otimes d} \oplus \mathbb{C} \cong \begin{cases} S^2 \times S^2 & \text{for } d \text{ even,} \\ \tilde{\mathbb{C}}P^2 & \text{for } d \text{ odd.} \end{cases}$$

Problem 7.1. Find explicit bundle isomorphisms from $\P[L^{\otimes d} \oplus \mathbb{C}]$ to $S^2 \times S^2$ respectively $\tilde{\mathbb{C}}P^2$. Show that $\P[L^{\otimes d} \oplus \mathbb{C}]$ and $\P[L^{\otimes d'} \oplus \mathbb{C}]$ are biholomorphically equivalent if and only if $d' = \pm d$.

Example 7.5. $X := \underline{\mathrm{X}}P^2 \# 9\overline{\mathbb{C}P^2}$. Let $\{P_0(z) = 0\}$, $\{P_1(z) = 0\}$ be nonsingular cubics in $\mathbb{C}P^2$ which intersect transversally in 9 points $p_1, \dots, p_9$. Consider the pencil (family parametrized by $\mathbb{C}P^1$) of cubics

$$C_{[a_0:a_1]} = \{a_0 P_0 + a_1 P_1 = 0\}$$

passing through $p_1, \dots p_9$. This pencil has the following two properties: Any two cubics in the pencil intersect transversally in $p_1, \dots, p_9$; and through every point of $\mathbb{C}P^2 \setminus \{p_1, \dots, p_9\}$ there passes exactly one cubic in the pencil. Hence we have a map $\mathbb{C}P^2 \setminus \{p_1, \dots, p_9\} \to \mathbb{C}P^1$ which assign to every z the point a such that $z \in C_a$. This map does not extend over $p_1, \dots, p_9$. But it does extend to the blow-up X of $\mathbb{C}P^2$ at the points $p_1, \dots, p_9$ by sending a tangent line l at p_j to the unique a such that $T_{p_j} C_a = l$. Thus we obtain a surjective holomorphic map

$$X \to \mathbb{C}P^1$$

with generic fibre a 2-torus T^2. Such a complex surface is called an *elliptic surface*. X has the following properties:

(i) A generic fibre has trivial normal bundle.

(ii) There is a finite number of singular fibres.

(iii) $X \setminus T^2$ is simply connected for a generic fibre T^2.

The first property is obvious. For the second property note that the second homotopy group ofthe total space of a torus bundle over S^2 injects into $\pi_2(S^2) = \mathbb{Z}$. Since X is simply connected, $H_2(X)$ would be at most one-dimensional if all fibres were nonsingular. But $H_2(X)$ is 10-dimensional, so there must be singular fibres.

To prove the third property, choose the cubics in such a way that the singular cubic

$$C = \{z_0 z_1^2 - z_2^3 = 0\}$$

occurs in the pencil. The holomorphic map

$$\mathbb{C}P^1 \to C, \quad [w_0, w_1] \mapsto [w_0^3 : w_2^3 : w_0 w_1^2]$$

is a homeomorphism, so C is homeomorphic to S^2, sitting in $\mathbb{C}P^2$ with a cusp at $[1:0:0]$. Now C appears as a singular fibre in the blow-up X. Since C is simply connected, any loop in a regular fibre of X can be contracted through C. On the other hand, a meridian of a regular fibre T^2 (a small loop encircling the fibre in X) can be contracted in $X \setminus T^2$ through an exceptional divisor because the exceptional divisors intersect all fibres. Since $\pi_1(X \setminus T^2)$ is generated by loops in the fibres and meridians, this proves $\pi_1(X \setminus T^2) = 0$.

7.2 Blowing up in the symplectic category

If (M, J, ω) is a Kähler manifold and $\beta : \tilde{M} \to M$ the blow-up at a point, then $\beta^*\omega$ is not symplectic on $\tilde{M}$ because it vanishes on the exceptional divisor E. So as it stands, blow-up does not preserve the symplectic category. There is, however, an obvious way to fix this: The exceptional divisor $E \cong \mathbb{C}P^{n-1}$ carries the Fubini-Study symplectic form ω_{FS}; adding a small multiple of ω_{FS} to $\beta^*\omega$ should yield a symplectic form on the blow-up. To see that this works, and can be done in a canonical way, we consider again the local model.

Let $L \subset \mathbb{C}^n \times \mathbb{C}P^{n-1}$ be the tautological line bundle with the projections $\beta : L \to \mathbb{C}^n$ and $\pi : L \to \mathbb{C}P^{n-1}$, and let L_0 be its zero section. Identify $L \setminus L_0$ with $\mathbb{C}^n \setminus 0$ via β, so the bundle projection becomes $\pi : \mathbb{C}^n \setminus 0 \to \mathbb{C}P^{n-1}$.

For a function $\phi : \mathbb{C}^n \to \mathbb{C}$ define $d^c\phi := d\phi \circ J$, where J is the standard complex structure on $\mathbb{C}^n$. A simple computation using $dz_j \circ J = i dz_j$ and $d\bar{z}_j \circ J = -i d\bar{z}_j$ shows that

$$-dd^c\phi = 2i\partial\bar{\partial}\phi = 2i \sum_{j,k} \frac{\partial^2 \phi}{\partial z_j \partial \bar{z}_k} dz_j \wedge d\bar{z}_k.$$

Applying this formula we compute for $z \neq 0$,

$$-dd^c \frac{1}{4}|z|^2 = \frac{i}{2}\sum_j dz_j \wedge d\bar{z}_j = \omega_{\mathrm{st}};$$

$$-dd^c \frac{1}{4}\log|z|^2 = \frac{i}{2}\Big(\sum_j \frac{dz_j \wedge d\bar{z}_j}{\|z\|^2} - \sum_{j,k} \frac{\bar{z}_j z_k dz_j \wedge d\bar{z}_k}{\|z\|^4}\Big) = \pi^*\omega_{FS},$$

where $\pi : \mathbb{C}^n \setminus 0 \to \mathbb{C}P^{n-1}$ is the projection (see Section 6.1).

Define the 1-form

$$\theta := -d^c \frac{1}{4}\log|z|^2,$$

and the vector fields

$$X := \sum_j (z_j \frac{\partial}{\partial z_j} + \bar{z}_j \frac{\partial}{\partial \bar{z}_j}), \quad JX = i\sum_j (z_j \frac{\partial}{\partial z_j} - \bar{z}_j \frac{\partial}{\partial \bar{z}_j}).$$

X is the radial vector field and JX the vector field generating the standard S^1-action on $\mathbb{C}^n$. The 1-form θ is invariant under homotheties $z \mapsto \lambda z$, $\lambda \in \mathbb{C}^*$, and satisfies $\theta(X) = 0$, $\theta(JX) = \frac{1}{2}$. So up to normalization, θ is a connection 1-form on the tautological line bundle $\pi : \mathbb{C}^n \setminus 0 \to \mathbb{C}P^{n-1}$.

In terms of θ,

$$d\theta = \pi^*\omega_{FS}, \quad d(|z|^2\theta) = \omega_{\mathrm{st}}.$$

More generally, for a function $\rho : \mathbb{R}^+ \to \mathbb{R}$,

$$d\Big(\rho(|z|^2)\theta\Big) = d\rho \wedge \theta + \rho\pi^*\omega_{FS}$$

is a Kähler form on $\mathbb{C}^n \backslash 0$ provided that $\rho(0) > 0$ and $\rho' > 0$. This is true because $d\rho \wedge \theta(X, JX) > 0$ and $\pi^*\omega_{FS}$ is nondegenerate in the transverse directions. This form does not extend over the origin. But $d(|z|^2) \wedge \theta = \omega_{\mathrm{st}} - |z|^2\pi^*\omega_{FS}$, and $\pi^*\omega_{FS}$ extends to the blow-up L. Therefore $d\big(\rho(|z|^2)\theta\big)$ extends to a Kähler form on L.

Picking a function ρ with $\rho(0) > 0$, $\rho' > 0$ and $\rho(|z|^2) = |z|^2$ for $|z|^2 \geq \delta$, we can put a symplectic form onto the blow-up. But we can do this more precisely. Let $B^{2n}(r)$ denote the ball of radius r and

$$L(\delta) := \{(z, [w]) \in L \mid |z| \leq \delta\}.$$

Lemma 7.6. *For each $\lambda > 0$,*

$$\omega_\lambda := \beta^*\omega_{\mathrm{st}} + \lambda^2\pi^*\omega_{FS}$$

is a Kähler form on L. Moreover, for all $\delta > 0$ we have symplectomorphisms

$$\Big(L(\delta) \setminus L_0, \omega_\lambda\Big) \cong \Big(B^{2n}(\sqrt{\lambda^2 + \delta^2}) \setminus B^{2n}(\lambda), \omega_{\mathrm{st}}\Big).$$

Proof. The first statement follows from the discussion above applied to the function $\rho(|z|^2) = \lambda^2 + |z|^2$. For the second statement, consider the function $f : \mathbb{C}^n \setminus 0 \to \mathbb{C}^n \setminus 0$,

$$f(z) := \sqrt{\lambda^2 + |z|^2}\frac{z}{|z|}.$$

Since the connection form θ has no radial component, θ and $f^*\theta$ equal their restriction to spheres around the origin. But between such spheres f acts as a homothety under which θ is invariant, so $f^*\theta = \theta$. From here it follows that

$$f^*\omega_{\text{st}} = d\, f^*(|z|^2\theta) = d(\lambda^2 + |z|^2)\theta = \omega_{\text{st}} + \lambda^2\omega_{FS}.$$

□

Remark. On $L \setminus L_0$ the form ω_λ is given by

$$\omega_\lambda = \frac{i}{2}\partial\bar{\partial}(|z|^2 + \lambda^2 \log|z|^2).$$

Definition. Let (M, ω) be a symplectic manifold and $\psi : B^{2n}(\lambda) \hookrightarrow M$ a symplectic embedding of a standard ball. Extend ψ to a symplectic embedding $\tilde{\psi} : B^{2n}(\sqrt{\lambda^2 + \delta^2}) \hookrightarrow M$ for some $\delta > 0$. Define

$$\tilde{M} := M \setminus \tilde{\psi}\Big(B^{2n}(\sqrt{\lambda^2 + \delta^2})\Big) \cup L(\delta),$$

glued via the symplectomorphism of Lemma 7.6, and a symplectic form $\tilde{\omega}$ on $\tilde{M}$ which equals ω on $M \setminus \tilde{\psi}\Big(B^{2n}(\sqrt{\lambda^2 + \delta^2})\Big)$ and ω_λ on $L(\delta)$. $(\tilde{M}, \tilde{\omega})$ is called the *blow-up of* M *at the ball* $\psi\big(B^{2n}(\lambda)\big)$.

We will identify $\tilde{M}$ as a smooth manifold with the blow-up $\beta : \tilde{M} \to M$ of M at the point $\psi(0)$, picking some integrable complex structure near this point. This allows us to speak of the projection β (which is not symplectic!) and the exceptional divisor E, and to identify blow-ups at different balls with the same center as smooth manifolds.

Lemma 7.7. *(i) The blow-up does not depend (up to symplectomorphism) on δ and the extension $\tilde{\psi}$.*

(ii) $\mathrm{vol}(\tilde{M}, \tilde{\omega}) = \mathrm{vol}(M, \omega) - \mathrm{vol}\big(B^{2n}(\lambda)\big)$.

(iii) $[\tilde{\omega}] = [\beta^*\omega] - \pi\lambda^2 PD[E] \in H^2(\tilde{M}; \mathbb{Z})$.

(iv) If $\psi_0, \psi_1 : B^{2n}(\lambda) \hookrightarrow M$ are isotopic through symplectic embeddings, then the corresponding forms $\tilde{\omega}_0$ and $\tilde{\omega}_1$ on $\tilde{M}$ are isotopic.

Proof. (i) For any extensions of ψ the forms $\tilde{\omega}$ agree on $\tilde{M} \setminus E$ by Lemma 7.6, so they agree on $\tilde{M}$ by continuity.

(ii) is clear.

(iii) The form $\tilde{\omega} - \beta^*\omega$ vanishes on $M \setminus \tilde{\psi}\Big(B^{2n}(\sqrt{\lambda^2+\delta^2})\Big)$. On $L(\delta)$ the map $\beta : \tilde{M} \to M$ is given by a map $L(\delta) \to \mathbb{C}^n$ which equals the symplectomorphism of Lemma 7.6 near $\partial L(\delta)$ and the usual map $\beta : L \to \mathbb{C}^n$ near L_0. So $\tilde{\omega} - \beta^*\omega$ is a form with compact support on $L(\delta$ which equals $\lambda^2\pi^*\omega_{FS}$ near L_0. Since the second cohomology of L with compact support is one-dimensional, $[\tilde{\omega} - \beta^*\omega]$ is a multiple of $PD[E]$. The factor is determined by evaluating both classes on a nontrivial relative homology class of (L, L_0). A line $\mathbb{C}P^1 \subset L_0$ provides such a class. On it $PD[E]$ takes the value -1 and $\lambda^2\pi^*\omega_{FS}$ the value $\pi\lambda^2$. Hence $[\tilde{\omega} - \beta^*\omega] = -\pi\lambda^2 PD[E]$.

(iv) Let $\psi_t : B^{2n}(\lambda) \hookrightarrow M$ be symplectic embeddings. After composing with symplectomorphisms of M sending $\psi_t(0)$ to $\psi_0(0)$ we may assume $\psi_t(0) = \psi_0(0)$ for all t. Then as smooth manifolds the blow-up are given by the same manifold $\tilde{M}$ with a map $\beta : \tilde{M} \to M$ and exceptional divisor E. By (iii) the symplectic forms $\tilde{\omega}_t$ induced by the embeddings ψ_t are cohomologous, so by Moser's Stability Theorem they are isotopic. □

Problem 7.2. Use the compatibility between (ii) and (iii) in the preceding lemma to derive a formula for $\mathrm{vol}\big(B^{2n}(\lambda)\big)$, and check that it is correct by direct integration.

7.3 Fibre connected sum

Connected sum

Let us start out by trying to do connected sums in the symplectic category. Take two symplectic manifolds and cut out small Darboux balls of the same volume. The boundary spheres have tubular neighbourhoods symplectomorphic to the annulus

$$A^{2n}(\delta, \varepsilon) := B^{2n}(\delta) \setminus B^{2n}(\varepsilon)$$

for some $\varepsilon < \delta$. All we need for the connected sum is a symplectomorphism of this annulus that interchanges the boundary components. Here is the bad news:

Lemma 7.8. *There exists a symplectomorphism of $A^{2n}(\delta, \varepsilon)$ that interchanges the boundary components if and only if $n = 1$.*

Proof. Consider a compact hypersurface S in R^{2n} with the standard symplectic form $\omega = \omega_{\mathrm{st}}$. Then $\omega|_S = d\lambda$ for some 1-form λ on S.

Claim: For $n \geq 2$, $\int_S \lambda \wedge \omega^{n-1}$ is independent of the choice of λ.

Indeed, suppose $d\lambda = d\mu = \omega|_S$. Then

$$\int_S (\lambda - \mu) \wedge \omega^{n-1} = \int_S d\Big((\lambda - \mu) \wedge \lambda \wedge \omega^{2n-2}\Big) = 0.$$

Note that the statement is false for $n = 1$.

Let us call S *convex (concave)* if there exists a λ on S such that $d\lambda = \omega|_S$ and $\lambda \wedge \omega^{n-1}|_S$ is a positive (negative) volume form on S with the natural orientation. For $n \geq 2$, if S is convex, $\int_S \lambda \wedge \omega^{n-1} > 0$ for all λ with $d\lambda = \omega|_S$, so S is not concave. For $n = 1$ every hypersurface is both convex and concave.

Now suppose $\phi : A^{2n}(\delta, \varepsilon) \to A^{2n}(\delta, \varepsilon)$ is symplectic and interchanges the boundary components. The sphere $S := S^{2n-1}(\delta)$ is convex. Choose a λ on S with $d\lambda = \omega|_S$ and $\lambda \wedge \omega^{n-1}|_S$ positive. Then $\lambda' := \phi^*\lambda$ on $S' := \phi^{-1}(S) = S^{2n-1}(\varepsilon)$ satisfies $d\lambda' = \omega|_{S'}$, and $\lambda' \wedge \omega^{n-1}|_{S'}$ is negative if S' is oriented as a hypersurface in $\mathbb{R}^{2n}$. Thus S' is concave. But S' is also convex, which is impossible for $n \geq 2$.

For $n = 1$, the unique S^1-equivariant symplectomorphism of the open unit ball without zero that interchanges the boundary components is

$$\phi(z) = \left(\frac{1}{|z|} - 1\right) \bar{z}.$$

□

Fibre connected sum

Despite the preceding lemma, the connected sum construction can be used for higher dimensions by applying it to the fibres of the normal bundle of a codimension 2 submanifold. This *fibre connected sum* construction was first observed by Gromov [9] and further developed by Gompf [8].

Consider an oriented 2-plane bundle $\pi : E \to N$ over a symplectic manifold (N, ω_N). Pick a Hermitian metric with norm $r = |z|$ in the fibres, and a connection 1-form θ on $E \setminus N$, where N is identified with the zero section. Then

$$\eta := d(\frac{1}{2}r^2\theta)$$

is an exact 2-form on E which restricts to the standard area form on each fibre, and

$$\pi^*\omega_N + t\eta$$

is a symplectic form on E for $t > 0$ sufficiently small.

The diffeomorphism

$$\iota(z) = \left(\frac{1}{|z|} - 1\right) z$$

of the open unit ball without zero in $\mathbb{R}^2$ interchanges the boundary components and satisfies $\iota^*\omega_{\mathrm{st}} = -\omega_{\mathrm{st}}$. ($\iota$ is the conjugate of the map ϕ in the previous proof). It induces a diffeomorphism of the open unit disk bundle without the zero section, $\iota : E^1 \setminus N \to E^1 \setminus N$, satisfying

$$\iota^*(\pi^*\omega_N + t\eta) = \pi^*\omega_N - t\eta.$$

Now consider two disjoint symplectic embeddings

$$j_1, j_2 : (N^{2n-2} \hookrightarrow (M^{2n}, \omega_M)$$

whose normal bundles are opposite as oriented 2-plane bundles,

$$\nu_1 \cong -\nu_2.$$

Let $\psi : \nu_1 \to -\nu_2$ be a symplectic bundle isomorphism, so

$$\psi^*(\pi^*\omega_N + t\eta_2) = \pi^*\omega_N - t\eta_1.$$

By the Symplectic Neighbourhood Theorem, a neighbourhood V_1 of $j_1(N)$ is symplectomorphic to the unit disk bundle,

$$(V_1, \omega_M) \cong (\nu_1^1, \pi_1^*\omega_N + t\eta_1),$$

for t sufficiently small, and similarly for $j_2(N)$. Under these identifications,

$$\phi := \iota \circ \psi : (V_1 - j_1(N), \omega_M) \to (V_2 - j_2(N), \omega_M)$$

is a symplectomorphism that identifies the inner boundary to the outer one.

Definition. The manifold

$$\#_\psi M := M - j_1(N) - j_2(N) \Big/ V_1 \underset{\phi}{\sim} V_2$$

obtained by cutting out $j_1(N)$ and $(j_2(N)$ and gluing the collars via ϕ is called the *fibre connected self-sum* of M. If $M = M_1 \amalg M_2$ is a disjoint union and $j_i : N \hookrightarrow M_i$, then

$$M_1 \#_\psi M_2 = \#_\psi(M_1 \amalg M_2)$$

is called the *fibre connected sum* of M_1 with M_2.

The discussion above implies

Proposition 7.9. *The fibre connected (self-)sum carries a symplectic structure which agrees with ω_M outside a neighbourhood of $j_1(N) \cup j_2(N)$.*

Remark. Gompf [8] actually showed that the symplectic structure on $\#_\psi M$ is canonical. But we do not need this for the applications below.

Applications

The main result of this chapter is

Theorem 7.10 (Gompf [8])**.** *Every finitely presented group is the fundamental group of a closed symplectic 4-manifold.*

Proof. 1. Let

$$< g_1, \dots, g_k \mid r_1, \dots, r_l >$$

be a presentation of G with generators g_i and relations r_i. Let F be a closed surface of genus k and $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k$ embedded oriented loops on F whose intersection numbers satisfy

$$\alpha_i \cdot \beta_j = \delta_{ij}.$$

Note that

$$\pi_1(F) \Big/ < \beta_1, \dots, \beta_k >$$

is the free group generated by $\alpha_1, \dots, \alpha_k$. Let γ_i, $i = 1, \dots, l$, be immersed oriented loops on F representing the relations r_i (with g_j replaced by α_j). Set $\gamma_{l+i} := \beta_i$ for $i = 1, \dots, k$. Then

$$\pi_1(F) \Big/ < \gamma_1, \dots, \gamma_{k+l} > \cong G.$$

2. We wish to construct a closed 1-form ρ on F such that $\rho|_{\gamma_i}$ is a volume form for all i. This is not possible in general due to homological reasons, but as we will see, it becomes possible after attaching sufficiently many handles to F.

Pick two distinct points $x, y \in S^1$ and the corresponding loops

$$\alpha = S^1 \times x, \quad \beta = x \times S^1, \quad \gamma = y \times S^1$$

on the torus $T^2 = S^1 \times S^1$. Pick a disk $D \subset T^2$ which is disjoint from α and β and intersects γ in a single arc (see Figure).

Make $\gamma_1, \dots, \gamma_{k+l}$ intersect transversally in double points and consider the oriented graph

$$\Gamma := \cup \gamma_i.$$

For each edge e of Γ take the connected sum of F with T^2 at a disk on e such that e gets summed with γ. Add α and β to the collection of the $\gamma)i$. This yields a new surface, still called F, of higher genus with a collection of immersed loops $\gamma_1, \dots, \gamma)m$ such that still

$$\pi_1(F) \Big/ < \gamma_1, \dots, \gamma_m > \cong G,$$

and each edge of $\Gamma = \cup \gamma_i$ has a segment that equals α, β or γ in some T^2.

3. Now there exists a closed 1-form ρ_0 on T^2 satisfying $\rho_0 = 0$ on D and $\int \rho_0 > 0$ over α, β and γ. (Take the form δ dual to $\alpha + \beta$, $\delta = df$ on D, extend f, and set $\rho_) := \delta - df$). Let ρ^* be the closed 1-form on F whichequals ρ_0 on $T^2 - D$ and zero otherwise. It satisfies

$$\int_e \rho^* > 0$$

for each oriented edge e of Γ. So for $i = 1, \ldots, m$ we find volume forms θ_i on (the domain of the loop) γ_i such that

$$\int_e \theta_i = \int_e \rho^*$$

for each edge e of Γ in γ_i. This implies

$$\theta_i - \rho^*|_{\gamma_i} = df_i$$

for a function $f_i : \gamma_i \to \mathbb{R}$ that vanishes at each vertex of Γ on γ_i. Extend the function $f_1 + \cdots + f_m$ on Γ to a smooth function f on F (which is possible since the f_i vanish at the vertices). Now

$$\rho := \rho^* + df$$

is a the desired closed 1-form on F such that $\rho|_{\gamma_i} = \theta_i$ is a volume form on each γ_i, $i = 1, \ldots, m$.

4. Now consider $F \times T^2$ with a product symplectic form ω and projections π_1, π_2 onto the factors. For $i = 1, \ldots, m$,

$$T_i := \gamma_i \times \alpha$$

is an immersed Lagrangian torus in $F \times T^2$. We also have a closed 2-form

$$\eta := \pi_1^* \rho \wedge \pi_2^* \theta$$

on $F \times T^2$, where θ is the pullback of a volume form on α to T^2, and ρ is the 1-form on F constructed in Steps 3 and 4. By construction, $\eta|_{T_i}$ is symplectic for all i. Thus for $t > 0$ sufficiently small,

$$\omega' := \omega + t\eta$$

is symplectic on $F \times T^2$, on T_i, and on $z \times T^2$, where z is any point in $F - \cup\gamma_i$.

Next write the symplectic manifold as

$$F \times T^2 = (F \times \beta) \times \alpha.$$

Perturb the loops γ_i in the 3-manifold $F \times \beta$ to disjointly embedded loops γ_i'. The resulting tori

$$T_i' := \gamma_i' \times \alpha$$

are disjointly embedded in $F \times T^2$, and they remain symplectic for small perturbations. Moreover, they remain disjoint from $z \times T^2$.

Note that the normal bundle of T_i' is the pullback of the normal bundle of γ_i' in $F \times \beta$, so it is trivial.

5. Let X be the rational elliptic surface from Example 7.5. A generic fibre $N \subset X$ is a symplectically embedded 2-torus with trivial normal bundle. Use

the fibre connected sum to attach a copy of X to $F \times T^2$ along each T_i' and along $z \times T^2$ (identifying T_i' with N after rescaling the symplectic form on X). Denote the resulting symplectic manifold by M.

we have shown in Example 7.5 that $X - N$ is simply connected. Therefore $\pi_1(T_i')$ and $\pi_1(z \times T^2)$ get killed in $\pi_1(M)$. In particular, the relations $[\gamma_i'] \in \pi_1(T_i')$ die in $\pi_1(M)$, and

$$\pi_1(M) = \pi_1(F) \Big/ < \gamma_1, \ldots, \gamma_m > \cong G.$$

$\square$

Example 7.11. (Gompf). $\mathbb{C}P^2 \# \overline{\mathbb{C}P^2}$ contains two disjoint symplectic spheres: the exceptional divisor E, and the lift H of a line in $\mathbb{C}P^2$ not passing through the blow-up point. They have self-intersections

$$E \cdot E = -1, \quad H \cdot H = +1.$$

If there $\int_E \omega = \int_H \omega$, then we could form the fibre connected self-sum identifying E and H to obtain a symplectic manifold diffeomorphic to $S^1 \times S^3$. Since this is impossible, we have shown:

There exists no symplectic form on $\mathbb{C}P^2 \# \overline{\mathbb{C}P^2}$ for which E and H are symplectic of the same area.

Problem 7.3. Suppose (M^{2n}, ω) has a symplectic submanifold $E \cong \mathbb{C}P^{n-1}$ whose normal bundle is the tautological bundle over $\mathbb{C}P^{n-1}$. Show that the fibre connected sum $M \#_\psi \mathbb{C}P^n$ identifying E with a hyperplane in $\mathbb{C}P^n$ equals the blow-down of M along E.

Problem 7.4. (Gompf). Consider the symplectic manifolds with product structures,

$$M_1 = T^2 \times S^2 \supset T^2 \times z, \quad M_2 = T^2 \times T^2 \supset T^2 \times w.$$

Show that $M_1 \#_\psi M_2$ identifying $T^2 \times z$ and $T^2 \times w$, with a suitable twisted framing ψ, is the Thurston manifold of Example 6.5.

Chapter 8

Hamiltonian systems on symplectic manifolds

8.1 Basic properties

Let (M, ω) be a symplectic manifold and $H : \mathbb{R} \times M \to \mathbb{R}$ a *time-dependent Hamiltonian.*

Definition. The *Hamiltonian vector field* X_H is the unique vector field satisfying

$$dH_t = i_{X_H}\omega, \quad \text{where } H_t = H(t, \cdot).$$

The *Hamiltonian system*

$$\dot{x}(t) = X_H\big(t, x(t)\big), \quad x : \mathbb{R} \to M$$

defines (at least locally) the *Hamiltonian flow (or phase flow)* $\phi_t^H : M \to M$.

In Darboux coordinates (q_i, p_i), $\omega = \sum dq_i \wedge dp_i$, the Hamiltonian vector field is given by

$$X_H = \sum \Big(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} \Big),$$

and the Hamiltonian system $\dot{x} = X_H(t, x)$, $x = (q, p)$, is equivalent to Hamilton's equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

The Hamiltonian flow obeys two fundamental conservation laws:

Proposition 8.1. *(i) The Hamiltonian flow (of a time-dependent Hamiltonian) preserves the symplectic form ω.*

(ii) If H is time-independent the Hamiltonian flow preserves the Hamiltonian H (conservation of energy).

Proof. 1. By Cartan's formula and since ω is closed,

$$L_{X_H}\omega = d\,i_{X_H}\omega + i_{X_H}d\omega = d(dH_t) = 0.$$

Thus $\frac{d}{dt}(\phi_t^H)^*\omega = (\phi_t^H)^* L_{X_H}\omega = 0$, and therefore $(\phi_t^H)^*\omega = \omega$.
2. The total time derivative of H equals

$$\frac{dH}{dt} = \frac{d}{dt}H\big(t, x(t)\big) = \frac{\partial H}{\partial t} + dH_t \cdot \dot{x} = \frac{\partial H}{\partial t} + \omega(X_H, \dot{x}) = \frac{\partial H}{\partial t},$$

which vanishes if H is time-independent. □

Remark. The preceding prove shows that for a time-dependent Hamiltonian,

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}.$$

The conservation of ω has two important consequences. The first one is geometric: Hamiltonian flows provide a wealth of symplectomorphisms.

Definition. A diffeomorphism $\phi : M \to M$ is called a *compactly supported Hamiltonian diffeomorphism* if $\phi = \phi_1^H$ for some compactly supported time-dependent Hamiltonian H.

Problem 8.1. Prove that the compactly supported Hamiltonian diffeomorphisms form a group under composition. Prove that this group acts transitively on any connected symplectic manifold.

The second consequence is dynamical:

Proposition 8.2 (Liouville's Theorem)**.** *The Hamiltonian flow preserves the* canonical volume form *on M,*

$$\mathrm{vol} = \frac{\omega^n}{n!}.$$

So $\mathrm{vol}\big(\phi_t^H(A)\big) = \mathrm{vol}(A)$ for every Borel measurable set $A \subset M$. In physicists' language, "the phase flow is incompressible". This property plays a crucial role in statistical mechanics. Its importance for classical mechanics rest on the following result:

Proposition 8.3 (Poincaré Recurrence Theorem)**.** *Let (X, μ) be a measure space with $\mu(X) < \infty$, and $f : X \to X$ a measure preserving bijection. Then for every $A \subset X$ with $\mu(A) > 0$ there exist infinitely many $N \in \mathbb{N}$ such that $f^N(A) \cap A \neq \emptyset$.*

Proof. If all $f^N(A)$, $N \in \mathbb{N}$, were disjoint then $\mu(X) \geq \mu(A) + \mu\big(f(A)\big) + \cdots = \infty$. This proves existence of one N such that $f^N(A) \cap A \neq \emptyset$. The existence of infinitely many follows by repeating this argument. □

Since open sets in $\mathbb{R}^{2n}$ have positive volume, the Poincaré Recurrence Theorem implies:

Corollary 8.4. *If the Hamiltonian flow remains in a compact region then it is* recurrent*, i.e. for every open set U there exist infinitely many $T > 0$ such that $\phi_T^H(U) \cap U \neq \emptyset$.*

Problem 8.2. (cf. [18]). Let (Σ, ω) be a 2-dimensional compact connected symplectic manifold and $H : \Sigma \to \mathbb{R}$ a nonconstant function. Let $Per \subset (0, \infty)$ be the set of minimal periods of nonconstant periodic orbits of $\dot{x} = X_H(x)$.

(i) Prove that Per is nonempty and connected.

(ii) Find examples where Per is: (i) a point, (ii) an interval $[a, b]$, (iii) an interval (a, ∞).

(iii) Prove: If Per is bounded from above then $\Sigma \cong S^2$,

(iv) Prove that Per always contains the value $area(\Sigma)/(\max H - \min H)$.

Solution. First study the periods of periodic orbits in the vicinity of a critical point. This can be done in arbitrary dimension $2n$. So let $H : \mathbb{R}^{2n} \to \mathbb{R}$ be a Hamiltonian with a critical point at the origin. Consider a nonconstant T-periodic orbit x of $\dot{x} = X_H(x)$.

Claim 1. If 0 is a degenerate critical point then for every T_0 there exists a δ such that if x remains in the δ-ball around 0 then $T \geq T_0$.

Proof. Recall that a 2π-periodic map $\gamma : \mathbb{R}/2\pi\mathbb{Z} \to \mathbb{R}^m$ with $\int_0^{2\pi} \gamma(t)dt = 0$ satisfies Wirtinger's inequality

$$\int_0^{2\pi} |\gamma|^2 dt \leq \int_0^{2\pi} |\dot{\gamma}|^2 dt.$$

By rescaling this implies

$$\int_0^T |\gamma|^2 dt \leq \left(\frac{T}{2\pi}\right)^2 \int_0^T |\dot{\gamma}|^2 dt$$

for every T-periodic map $\gamma : \mathbb{R}/T\mathbb{Z} \to \mathbb{R}^m$ with $\int_0^T \gamma(t)dt = 0$. Apply this to the map $\gamma := \dot{x} : \mathbb{R}/T\mathbb{Z} \to \mathbb{R}^{2n}$. Since 0 is a critical point, for every ε there exists a δ such that $|DX_H(x)| < \varepsilon$ for $|x| < \delta$. Hence

$$\int_0^T |\dot{\gamma}|^2 dt = \int_0^T |DX_H(x)\gamma|^2 dt \leq \varepsilon^2 \int_0^T |\gamma|^2 dt \leq \left(\frac{\varepsilon T}{2\pi}\right)^2 \int_0^T |\dot{\gamma}|^2 dt.$$

Since $\dot{\gamma}$ does not vanish identically this implies $T \geq 2\pi/\varepsilon$, which proves Claim 1.

Claim 2. For every T_0 and ε there exists a $\delta < \varepsilon$ such that if x meets the δ-ball but does not remain in the ε-ball then $T \geq T_0$.

Proof. On the ε-ball we have $|X_H(x)| \le c|x|$ for some constant c. Consider times $t_0 < t_1$ with $|x(t_0)| = \varepsilon/2$, $|x(t_1)| = \varepsilon$ and $|x(t)| \le \varepsilon$ for $t \in [t_0, t_1]$. Then

$$\varepsilon/2 = |x(t_1)| - |x(t_0)| \le |x(t_1) - x(t_0)| \le \int_{t_0}^{t_1} |\dot{x}| dt \le \int_{t_0}^{t_1} c|x| dt \le c\varepsilon(t_1 - t_0),$$

thus $t_1 - t_0 \ge 1/2c$. Note that this estimate remains the same if we make ε smaller. Using the estimate with ε replaced by $\varepsilon/2, \varepsilon/4, \ldots$ we see that T must arbitrarily large, and Claim 2 is proved.

Claim 3. Suppose $n = 1$ and $0 \in \mathbb{R}^2$ is a nondegenerate local maximum or minimum. Then there exist δ and c such that if x remains in the δ-ball then $1/c \le T \le c$.

Proof. Since the periodic orbits are close to ellipses for δ small, there exists a constant c such that every periodic orbit in the δ-ball satisfies

$$\max_t |x(t)|/c \le l(x) \le c \min_t |x(t)|.$$

Here $l(x) = \int_)^T |\dot{x}| dt$ denotes the length of x. Moreover, we may assume $|x|/c \le |X_H(x)| \le c|x|$ for $|x| \le \delta$. Then

$$\begin{aligned} l(x) &\le \int_0^T c|x| dt \le cT \max_t |x(t)| \le c^2 T l(x), \\ l(x) &\ge \int_0^T \frac{|x|}{c} dt \ge \frac{T}{c} \min_t |x(t)| \ge \frac{T}{c^2} l(x). \end{aligned}$$

Hence $1/c^2 \le T \le c^2$, and Claim 3 is proved.

Now consider the Hamiltonian system on the surface (Σ, ω). Let $(x_s)_{s \in (0,1)}$ be an embedded family of nonconstant periodic orbits of periods T_s. Suppose the family is maximal in the sense that the closure of the family in Σ contains critcal points p_0 at $s = 0$ and p_1 at $s = 1$. Call such a family a *regular family*. Call the end $s = 0$ of *type 1* if the critical point p_0 is either degenerate or hyperbolic (nondegenerate of index 1). Then it follows from Claims 1 and 2 that $T_s \to \infty$ as $s \to 0$.

Call the end $s = 0$ of *type 2* if it is not of type 1. Then p_0 is an elliptic critical point (a nondegenerate maximum or minimum). It follows from Claim 3 that $T_s \to a$, $0 < a < \infty$, as $s \to 0$.

Let $I := \{T_s \mid s \in (0,1)\} \subset (0, \infty)$. Note that I is connected. The discussion above gives us all possibilities for I.

- If both ends are of type 1 then $I = [a, \infty)$ for some $0 < a < \infty$.
- If one end is of type 1 and the other one of type 2 then $I = [a, \infty)$ or $I = (a, \infty)$ for some $0 < a < \infty$.

- If both ends are of type 2 then $I = \{a\}$, $I = [a, b]$, $I = [a, b)$, $I = (a, b]$ or $I = (a, b)$ for some $0 < a < b < \infty$.

Now (a) follows by considering connected components of $H^{-1}((a, b))$ for intervals (a, b) consisting of regular values.

For (b) observe that all possible forms of Per can be realized by a Hamiltonian on S^2 with only 2 critical points, the maximum and minimum. In particular, $Per = \{2\pi\}$ for the height function on the unit sphere $S^2 \subset \mathbb{R}^3$ with the standard area form.

(c) If Per is bounded from above then for every regular family both ends must be of type 2. Hence all critical points of H are nondegenerate of index 0 or 2. Since Σ is connected, this implies that H has only 2 critical points, the maximum and minimum. This can only happen for $\Sigma \cong S^2$.

(d) Consider a regular family $(x_h)_{h\in(h_0,h_1)}$, parametrized such that $H(x_h) = h$, with periods T_h. Let $R := \{(t, h) \in \mathbb{R}^2 \mid 0 < t < T_h, h_0 < h < h_1\}$ and define $\phi : R \to \Sigma$, $\phi(t, h) := x_h(t)$. Then

$$1 \equiv \frac{\partial(H \circ \phi)}{\partial s} = dH(\frac{\partial \phi}{\partial s}) = \omega(X_H, \frac{\partial \phi}{\partial s}) = \omega(\frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial s}).$$

So the area of $\phi(R)$ is given by

$$\int_R \phi^* \omega = \int_{h_0}^{h_1} \int_0^{T_h} dt\, dh = \int_{h_0}^{h_1} T_h\, dh.$$

Choose for each interval of regular values of H such a regular family. This yields a measurable function $h \mapsto T_h$ on $(\min H, \max H)$. Let A be the sum of the areas of all the regular families. Then $h \mapsto T_h$ is integrable and

$$\int_{\min H}^{\max H} T_h\, ds = A \leq area(\Sigma).$$

Now distinguish 2 cases. If $Per = [a, \infty)$ or (a, ∞) the statement follows from the inequality

$$\min_h T_h(\max H - \min H) \leq \int_{\min H}^{\max H} T_h\, ds \leq area(\Sigma).$$

If Per is a finite interval then H has only 2 critical points, its maximum and minimum. Hence $h \mapsto T_h$ extends to a continuous map on $[\min H, \max H]$ and $A = area(\Sigma)$. Now by the intermediate value theorem, there exists a $h \in [\min H, \max H]$ with $T_h(\max H - \min H) = area(\Sigma)$. □

Problem 8.3. Let $S^1 = \mathbb{R}/\mathbb{Z}$ be the unit circle.

(i) Show that every orientation preserving diffeomorphism of S^1 is the time-one map of a time-dependent vector field.

(ii) Find an orientation preserving diffeomorphism of S^1 which is *not* the time-one map of any time-independent vector field.

Solution. (i) Lift $\bar{\phi} : S^1 \to S^1$ to $\phi : \mathbb{R} \to \mathbb{R}$. Then $\bar{\phi}$ is an orientation preserving diffeomorphism iff $\phi' > 0$ and $\phi(x+1) = \phi(x)+1$. Now $\phi_t(x) := (1-t)x+t\phi(x)$ satisfies $\phi_t' > 0$ and $\phi_t(x+1) = \phi_t(x) + 1$, hence it descends to an isotopy $\bar{\phi}_t : S^1 \to S^1$ connecting $\bar{\phi}$ to the identity. Its t-derivative is the desired time-dependent vector field.

(ii) Consider an orientation preserving diffeomorphism $\phi : S^1 \to S^1$ without fixed points. Suppose that $\phi = \phi_1$ is the time-one map of a time-independent vector field $f : S^1 \to \mathbb{R}$. Since ϕ has no fixed points, we must have $f \neq 0$, say $f > 0$ everywhere.

Claim: If ϕ^2 has a fixed point then ϕ^2 is the identity.

To see this, consider the flow ϕ_t of f, and let $\phi^2(x_0) = x_0$. By the group property of the flow (here we use that f is time-independent),

$$\phi^2 \circ \phi_t(x_0) = \phi_{2+t}(x_0) = \phi_t \circ \phi^2(x_0) = \phi_t(x_0),$$

so every $\phi_t(x_0)$ is a fixed point of ϕ^2. But as t increases, $\phi_t(x_0)$ runs around the whole circle, hence ϕ^2 is the identity.

In view of this claim, it suffices to find an orientation preserving diffeomorphism ϕ without fixed points such that ϕ^2 has a fixed point but is not the identity. Such ϕ's clearly exist. Here is an explicit example: Let $f(t,x)$ be a time-dependent vector field on S^1 with the following properties: $f(t,x) = 1/2$ for $0 \le t \le 1/4$ and $0 \le x \le 3/4$, and for $1/4 \le t \le 1$ and all x; $f(t,x) > 1/2$ for $0 < t < 1/4$ and $3/4 < x < 1$. Its time-one map ϕ maps $0 \to 1/2$, $1/4 \to 3/4$, $1/2 \to 1 \equiv 0$ and $3/4 \to p > 1/4$. Thus $\phi^2(0) = 0$ but $\phi^2(1/4) = p > 1/4$, so ϕ^2 is not the identity. □

Problem 8.4. Use the preceding problem to construct a Hamiltonian diffeomorphism of the 2-torus which is *not* the time-one map of a time-independent Hamiltonian.

8.2 Variational principles

Hamilton's principle

Consider an *exact symplectic manifold* $(M, \omega = d\lambda)$ with a time-dependent Hamiltonian H. Define the *Hamiltonian action* of a path $x : [a,b] \to M$ by

$$\mathcal{A}_H(x) := \int_a^b \Big(-x^*\lambda - H(t,x)dt\Big).$$

Let $L_0, L_1 \subset M$ be submanifolds with $\lambda|_{L_i} = 0$ for $i = 1, 2$.

Proposition 8.5 (Hamilton's principle)**.** *For a path $x : [a,b] \to M$ with $x(a) \in L_0$, $x(b) \in L_1$ the following are equivalent:*

(i) x is an extremal of $\mathcal{A}_H$ among variations with endpoints on L_0 and L_1.

(ii) x is an extremal of $\mathcal{A}_H$ among variations with fixed endpoints.

(iii) x is a solution of $\dot{x} = X_H(t,x)$.

Proof. Let ξ be a vector field along x with $\xi(a) \in TL_0$ and $\xi(b) \in TL_1$. Let $\phi : [0,\varepsilon] \times [a,b] \to M$ satisfy

$$\phi(0,t) = x(t), \quad \phi(s,a) \in L_0, \quad \phi(s,b) \in L_1, \quad \frac{\partial \phi}{\partial s}\Big|_{s=0} = \xi.$$

By Stokes' Theorem and since $\lambda|_{L_i} = 0$,

$$\int_{[0,\varepsilon]\times[a,b]} \phi^* d\lambda = \int_{\varepsilon} \times [a,b] \phi^* \lambda - \int_0 \times [a,b] \phi^* \lambda.$$

Dividing by ε and letting $\varepsilon \to 0$, the right-hand side tends to the derivative of $\int_x \lambda$ in the direction ξ, and the left-hand side converges to $\int_a^b d\lambda(\xi, \dot{x}) dt$. From this it follows that

$$\begin{aligned} d\mathcal{A}_H(x) \cdot \xi &= \int_a^b \Big(-d\lambda(\xi, \dot{x}) - dH(t,x) \cdot \xi \Big) dt \\ &= \int_a^b \omega\Big(\dot{x} - X_H(t,x), \xi \Big) dt, \end{aligned}$$

which implies the proposition. □

For a cotangent bundle $M = T^*Q$, $\lambda = -p\,dq$ and $L_i = T^*_{q_i}Q$ the fibres at points q_0, q_1, this proposition recovers Hamilton's extended variational principle of Proposition 3.2.

Remark. Versions of Hamilton's principle hold for periodic, homoclinic or heteroclinic orbits. This variational principle has turned out to be a powerful tool to prove existence of such special solutions for very general Hamiltonian systems. For example, for a large class of 1-periodic Hamiltonian systems on a cotangent bundle over a compact manifold (including all natural Hamiltonians), there exist infinitely many 1-periodic solutions [5]. On a compact symplectic manifold, the number of 1-periodic solutions is estimated below by topology of the manifold (sum of Betti numbers, cuplength). This is the famous *Arnold conjecture* that led to the invention of *Floer homology* (see [7]).

Maupertuis' principle

Consider now a time-independent Hamiltonian on a symplectic manifold (M, ω). By conservation of energy, each regular level set $S := H^{-1}(c)$ is invariant under the Hamiltonian flow. The flow on S can be described geometrically as follows: The hypersurface S carries the *characteristic line field*

$$\ker(\omega|_S) = (TS)^{\perp_\omega} \subset TS.$$

Integral curves of this line field are called *characteristics*. They form a 1-dimensional foliation of S called the *characteristic foliation*. Since $\omega(X_H, v) = dH(v) = 0$ for every $v \in TS$, the Hamiltonian vector field X_H spans the characteristic line field. So we see:

Proposition 8.6. *The orbits of the Hamiltonian flow on a regular energy surface $S = H^{-1}(c)$ are the characteristics of S. So if $\tilde{H}$ is another Hamiltonian with S as a regular energy surface, then the orbits of $\tilde{H}$ on S differ from those of H only in the parametrization.*

If $\omega|_S = d\lambda$ is exact we can give a variational characterization of characteristics on S. For an (unparametrized) path γ on S define its *action*

$$\mathcal{A}(\gamma) := -\int_\gamma \lambda.$$

Let $L_0, L_1 \subset S$ be submanifolds with $\lambda|_{L_i} = 0$. Then the variation of a path γ on S from L_0 to L_1 among such paths is given by

$$d\mathcal{A}(\gamma) \cdot \xi = -\int_\gamma i_\xi \omega,$$

and we conclude:

Proposition 8.7. *For an unparametrized path γ on S from L_0 to L_1 the following are equivalent:*

(i) γ is an extremal of $\mathcal{A}$ among variations with endpoints on L_0 and L_1.

(ii) γ is an extremal of $\mathcal{A}$ among variations with fixed endpoints.

(iii) γ is a characteristic on S.

Now consider a cotangent bundle $M = T^*Q$ with a fibrewise convex Hamiltonian H and its Legendre transform L. Fix a regular energy surface $H^{-1}(E)$. Let $L_i = T^*_{q_i}Q$ be the fibres at fixed points $q_0, q_1 \in Q$.

A path $\big(q(\tau), p(\tau)\big)$ in T^*Q is extremal for the Hamiltonian action among variations on each fibre if and only if $p = \frac{\partial L}{\partial \dot{q}}$, and solutions of the Hamiltonian systems are extremals of the Hamiltonian action among such paths $(q, p = \frac{\partial L}{\partial \dot{q}})$.

Conversely, let q be an unparametrized path in Q from q_0 to q_1. For each parametrization consider the path $(q, p = \frac{\partial L}{\partial \dot q})$ in T^*Q. Up to a time shift, there exists a unique parametrization of q such that the path $(q, p = \frac{\partial L}{\partial \dot q})$ lies on $H^{-1}(E)$. Define the *reduced action* of q to be

$$\mathcal{A}_E(q) := \int_q p\, dq = \int_a^b \langle \frac{\partial L}{\partial \dot q}, \dot q\rangle d\tau$$

with this parametrization. Then the preceding proposition implies:

Corollary 8.8 (Maupertuis' principle)**.** *A path from q_0 to q_1, parametrized such that $H(q, \frac{\partial L}{\partial \dot q}) \equiv E$, is a solution of the Hamiltonian system if and only if it is an extremal for the reduced action $\mathcal{A}_E$ among all paths from q_0 to q_1.*

Now specialize to a natural Hamiltonian $H = \frac{1}{2}\|p\|^2 + V(q)$ on T^*Q. Assume that V is bounded from above, so that $H^{-1}(E)$ is a regular level set for every $E > \sup V$. We can rewrite this level set as

$$\{H = E\} = \Big\{ \frac{\|p\|^2}{2\big(E - V(q)\big)} = 1 \Big\}.$$

So by Proposition 8.6, solutions of the Hamiltonian system for H of energy E are, up to parametrization, solutions of the Hamiltonian system for $\frac{\|p\|^2}{2\left(E-V(q)\right)}$ of energy 1. But $\frac{\|p\|^2}{\left(E-V(q)\right)}$ is just a metric on T_q^*Q, dual to the Riemannian metric $\big(E - V(q)\big)\|\dot q\|^2$ on T_qQ. So we have shown:

Corollary 8.9. *The solutions of the natural Hamiltonian system $H = \frac{1}{2}\|p\|^2 + V(q)$ of energy $E > \sup V$ are, up to parametrization, precisely the geodesics in the* Jacobi metric $\big(E - V(q)\big)\|\dot q\|^2$ *on Q.*

This corollary shows that time-independent natural Hamiltonian systems behave for high energy as geodesic flows. So all general properties of geodesic flows carry over to natural Hamiltonian systems. For example, as a consequence of the Lusternik-Fet Theorem, if Q is compact then for every energy $E > \sup V$ there exists a periodic orbit of the Hamiltonian system of energy E.

Problem 8.5. Use the Jacobi metric to prove the following special case of a theorem of Bolotin: Let $V : \mathbb{R}^n \to \mathbb{R}$ and E satisfy

$$\langle q, \nabla V(q)\rangle > 2\big(V(q) - E\big) \quad \text{for all } q \text{ with } V(q) \le E.$$

Then there exists a libration motion of energy E, i.e. a solution of $\ddot q = -\nabla V(q)$ oscillating between two points on the level surface $\{V = E\}$.

Chapter 9

Poisson manifolds

9.1 The Poisson bracket on a symplectic manifold

Let (M, ω) be a symplectic manifold. The time evolution of functions under Hamiltonian flows can be described very elegantly by the Poisson brackets:

Definition. The *Poisson bracket* of $F, G \in \Omega^0(M)$ is

$$\{F, G\} := \omega(X_F, X_G) = \underline{X}_G \cdot F = -X_F \cdot G.$$

In Darboux coordinates, $X_G = \sum(\frac{\partial G}{\partial p_i}\frac{\partial}{\partial q_i} - \frac{\partial G}{\partial q_i}\frac{\partial}{\partial p_i})$, hence

$$\{F, G\} = X_G \cdot F = \sum\Big(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial G}{\partial q_i}\frac{\partial F}{\partial p_i}\Big).$$

For example,

$$\begin{aligned}\{q_i, q_j\} = \{p_i, p_j\} &= 0,\\ \{q_i, p_j\} &= \delta_{ij}.\end{aligned}$$

Lemma 9.1. *The Poisson bracket is invariant under symplectomorphisms ϕ,*

$$\{\phi^* F, \phi^* G\} = \phi^*\{F, G\}.$$

Proof. Using $X_{\phi^* F} = \phi^* X_F$,

$$\begin{aligned}\{\phi^* F, \phi^* G\} &= \omega(X_{\phi^* F}, X_{\phi^* G}) = \omega(\phi^* X_F, \phi^* X_G)\\ &= \phi^*\big(\omega(X_F, X_G)\big) = \phi^*\{F, G\}.\end{aligned}$$

□

The importance of the Poisson bracket lies in the following observation:

Lemma 9.2. *The time evolution of a function under the Hamiltonian flow of a time-independent Hamiltonian H is given by*

$$\frac{d}{dt}(F \circ \phi_t^H) = \{F \circ \phi_t^H, H\}.$$

Proof. Using Lemma 9.1, this follows from the definitions,

$$\frac{d}{dt}(F \circ \phi_t^H) = (\phi_t^H)^*(X_H \cdot F) = (\phi_t^H)^*\{F, H\} = \{F \circ \phi_t^H, H \circ \phi_t^H\},$$

and $H \circ \phi_t^H = H$. □

For example, conservation of energy reads as $\{H, H\} = 0$, and constants of the motions are precisely the functions F that Poisson commute with H, $\{F, H\} = 0$.

The Poisson bracket is obviously *skew-symmetric*, $\{F, G\} = -\{G, F\}$, and it has the *derivation property*, $\{FG, H\} = F\{G, H\} + G\{F, H\}$. A less obvious property follows from the closedness of ω:

Lemma 9.3.

$$d\omega(X_F, X_G, X_H) = -\Big(\{\{F, G\}, H\} + \{\{G, H\}, F\} + \{\{H, F\}, G\}\Big).$$

Proof. We use the formula

$$d\omega(X_F, X_G, X_H) = X_F \cdot \omega(X_G, X_H) - \omega([X_F, X_G], X_H) + \text{cyclic}.$$

Evaluate the terms on the right-hand side:

$$X_F \cdot \omega(X_G, X_H) = \{\{G, H\}, F\},$$

$$\begin{aligned} -\omega([X_F, X_G], X_H) &= [X_F, X_G] \cdot H = X_F \cdot \{H, G\} - X_G \cdot \{H, F\} \\ &= \{\{H, G\}, F\} - \{\{H, F\}, G\} = -\{\{G, H\}, F\} - \{\{H, F\}, G\} \end{aligned}$$

Adding up all cyclic permutations yields the desired identity. □

So $d\omega = 0$ is equivalent to the *Jacobi identity*

$$\{\{F, G\}, H\} + \{\{G, H\}, F\} + \{\{H, F\}, G\} = 0.$$

9.2 Poisson manifolds

Definition. A *Poisson manifold* is a manifold M with an $\mathbb{R}$-bilinear operation $\{\ ,\ \} : \Omega^0(M) \times \Omega^0(M) \to \Omega^0(M)$ satisfying

(i) (skew-symmetry) $\{F, G\} = -\{G, F\}$;

(ii) (derivation property) $\{FG, H\} = F\{G, H\} + G\{F, H\}$;

(iii) (Jacobi identity) $\{\{F, G\}, H\} + \{\{G, H\}, F\} + \{\{H, F\}, G\} = 0$.

A map ϕ between Poisson manifolds is called *Poisson map* if $\{\phi^* F, \phi^* G\} = \phi^*\{F, G\}$ for all functions F, G.

We have seen in the previous section that every symplectic manifold carries a Poisson structure. But not every Poisson structure is symplectic:

Problem 9.1. Prove that the *rigid body bracket*

$$\{F, G\}(\xi) := \langle \xi, \nabla_\xi F \times \nabla_\xi G \rangle$$

for $\xi \in \mathbb{R}^3$, $F, G \in \Omega^0(\mathbb{R}^3)$ defines a Poisson structure on $\mathbb{R}^3$. Hint: The Jacobi identity follows from the Jacobi identity for the cross product.

In view of the derivation property, the value of $\{F, G\}$ at a point depends only on the differentials of F and G at this point. So a Poisson structure defines a skew-symmetric 2-tensor field $B \in \Omega_2(M)$, called the *Poisson tensor*, via

$$B_x(d_x F, d_x G) := \{F, G\}.$$

A diffeomorphism ψ between manifolds with Poisson tensors B_1, B_2 is a Poisson map if and only if $\psi_* B_1 = B_2$.

Problem 9.2. (cf. [14]). Prove that there exists a unique bilinear operation $[\ ,\ ] : \Omega_*(M) \times \Omega_*(M) \to \Omega_*(M)$, called the *Schouten bracket*, with the following properties:

(i) $\deg[A, B] = \deg A + \deg B - 1$;

(ii) $[A, B] = (-1)^{\deg A \deg B}[B, A]$;

(iii) $[A, B \wedge C] = [A, B] \wedge C + (-1)^{(\deg A + 1)\deg B} B \wedge [A, C]$;

(iv) $[\ ,\ ]$ is given on functions F, G and vector fields X, Y by: $[F, G] = 0$, $[X, F] = X \cdot F$, and $[X, Y]$ is the Lie bracket.

Prove that the Jacobi identity for a Poisson structure is equivalent to $[B, B] = 0$ for the Poisson tensor.

Hamiltonian flows

We associate to every function H its *Hamiltonian vector field* X_H defined by

$$X_H \cdot F := \{F, H\}, \quad F \in \Omega^0(M).$$

In terms of the Poisson tensor,

$$X_H = -i_{dH} B.$$

For time-dependent $H(t,x) = H_t(x)$ define its time-dependent Hamiltonian vector field by $X_H(t,x) = X_{H_t}(x)$. Denote the *Hamiltonian flow* of H by ϕ_t^H.

If the Poisson structure is symplectic the Hamiltonian vector field defined by the Poisson structure agrees with the one defined by the symplectic structure, which justifies using the same notation X_H for both of them.

Lemma 9.4. *Let $(M, \{\ \})$ be a Poisson manifold.*

(i) The Jacobi identity for $\{\ ,\ \}$ is equivalent to $L_{X_H} B = 0$ for every function H.

(ii) $X_{\psi^* H} = \psi^* X_H$ for every Poisson map ψ.

(iii) $F \mapsto X_F$ defines a Lie algebra antihomomorphism,

$$X_{\{F,G\}} = -[X_F, X_G].$$

Proof. (i) Using $L_{X_H} dF = d(X_H \cdot F) = d\{F, H\}$, we obtain

$$\begin{aligned}(L_{X_H} B)(dG, dG) &= X_H \cdot B(dF, dG) - B(L_{X_H} dF, dG) - B(dF, L_{X_H} dG) \\ &= X_H \cdot \{F, G\} - B(d\{F, H\}, dG) - B(dF, d\{G, H\}) \\ &= \{\{F, G\}, H\} - \{\{F, H\}, G\} - \{F, \{G, H\}\}.\end{aligned}$$

(ii) follows from $\psi^* B = B$:

$$\psi^* X_H = -\psi^*(i_{dH} B) = -i_{\psi^* dH} B = X_{\psi^* H}.$$

(iii) follows from the Jacobi identity:

$$\begin{aligned}X_{\{F,G\}} \cdot H &= -\{\{F, G\}, H\} = \{\{G, H\}, F\} + \{\{H, F\}, G\} \\ &= X_F\{G, H\} + X_G\{H, F\} = -X_F X_G H + X_G X_F H\end{aligned}$$

□

Proposition 9.5. *Let ϕ_t^H be the flow of a time-dependent Hamiltonian on a Poisson manifold.*

(i) For a time-dependent function F_t,

$$\frac{d}{dt}(F \circ \phi_t^H) = \frac{\partial F}{\partial t} \circ \phi_t^H + \{F, H\} \circ \phi_t^H.$$

(ii) If H is time-independent, then $H \circ \phi_t^H = H$ (conservation of energy).

(iii) The Hamiltonian flow is a Poisson map,

$$\{F \circ \phi_t^H, G \circ \phi_t^H\} = \{F, G\} \circ \phi_t^H.$$

Proof. (i) follows straight from the definitions:

$$\begin{aligned}\frac{d}{dt}(F \circ \phi_t^H) &= \frac{\partial F}{\partial t} \circ \phi_t^H + (X_H \cdot F) \circ \phi_t^H \\ &= \frac{\partial F}{\partial t} \circ \phi_t^H + \{F, H\} \circ \phi_t^H.\end{aligned}$$

(ii) Put $F = H$ in (i).

(iii) By Lemma 9.4,

$$\frac{d}{dt}(\phi_t^H)^* B = (\phi_t^H)^* L_{X_{H_t}} B = 0,$$

thus $(\phi_t^H)^* B = B$. □

Remark. (cf. [14]). If F and H are time-independent, $\frac{d}{dt}(F \circ \phi_t^H) = \{F \circ \phi_t^H, H\}$. This leads to a more explicit proof of (iii) for time-independent H:

Set $u(t) := \{F \circ \phi_t^H, G \circ \phi_t^H\} - \{F, G\} \circ \phi_t^H$. Property (i) and the Jacobi identity yield

$$\begin{aligned}\frac{du}{dt} &= \{\{F \circ \phi_t^H, H\}, G \circ \phi_t^H\} + \{F \circ \phi_t^H, \{G \circ \phi_t^H, H\}\} - \{\{F, G\} \circ \phi_t^H, H\} \\ &= \{\{F \circ \phi_t^H, G \circ \phi_t^H\}, H\} - \{\{F, G\} \circ \phi_t^H, H\} \\ &= \{u, H\} = X_H \cdot u.\end{aligned}$$

It follows that

$$\frac{d}{dt}(\phi_{-t}^H)^* u(t) = (\phi_{-t}^H)^* (\dot{u} - X_H \cdot u) = 0.$$

Thus $u(t) = u(0) \circ \phi_t^H = 0$ since $u(0) = 0$.

Definition. A diffeomorphism of a Poisson manifold is called *compactly supported Hamiltonian diffeomorphism* if it is the time-1 map of a time-dependent compactly supported Hamiltonian.

Lemma 9.6. *(cf. [7]). The compactly supported Hamiltonian diffeomorphisms on a Poisson manifold form a group under composition. More precisely, for time-dependent Hamiltonians H, K and a Poisson diffeomorphism ψ,*

(i) $\phi_t^H \circ \phi_t^K$ is the Hamiltonian flow of $H(t, x) + K\big(t, (\phi_t^H)^{-1}(x)\big)$;

(ii) $(\phi_t^H)^{-1}$ is the Hamiltonian flow of $-H\big(t, (\phi_t^H)^{-1}(x)\big)$;

(iii) $\psi \circ \phi_t^H \circ \psi^{-1}$ is the Hamiltonian flow of $H\big(t, \psi^{-1}(x)\big)$.

Proof. This follows from $X_{\psi^*H} = \psi^* X_H$ and Lemma A.1:

(iii) The flow $\psi \circ \phi_t^H \circ \psi^{-1}$ is generated by the time-dependent vector field $\psi_* X_{H_t} = X_{\psi_* H_t}$.

(i) The flow $\phi_t^H \circ \phi_t^K$ is generated by the time-dependent vector field $X_{H_t} + (\phi_t^H)_* X_{K_t} = X_{H_t} + X_{(\phi_t^H)_* K_t}$.

(ii) The flow $(\phi_t^H)^{-1}$ is generated by the time-dependent vector field $-(\phi_t^H)^* X_{H_t} = -X_{(\phi_t^H)^* H_t}$. □

9.3 Symplectic leaves

Let us now investigate a little more the structure of Poisson manifolds. In contrast to a symplectic structure a Poisson structure may be degenerate, i.e. the Poisson tensor may not induce an isomorphism $T^*M \to TM$. For example, the rigid body bracket on $\mathbb{R}^3$ is degenerate at the origin. At the extreme, the zero bracket is a perfectly nice Poisson structure. On the other hand, a nondegenerate Poisson structure is just a symplectic structure:

Lemma 9.7. *If a Poisson structure is nondegenerate in the sense that $\{F, G\} = 0$ for all G implies $F = \text{const}$ (equivalently, the Poisson tensor defines an isomorphism $T^*M \to TM$, $\lambda \mapsto i_\lambda B$), then it is induced by a symplectic structure.*

Proof. Define ω by $\omega(i_\lambda B, i_\mu B) := B(\lambda, \mu)$. In view of the nondegeneracy of B, ω is well-defined and nondegenerate. It has the following properties:

1. The Hamiltonian vector field X_H defined by ω agrees with the one defined by the Poisson structure, $-i_{dF}B$.

To see this, note that

$$\omega(-i_{dF}B, -i_\mu B) = B(dF, \mu) = dF(-i_\mu B)$$

for all μ, hence $\omega(-i_{dF}B, \cdot) = dF$.

2. $\{F, G\} = \omega(X_F, X_G)$.

This follows from the first claim:

$$\omega(X_F, X_G) = \omega(-i_{dF}B, -i_{dG}B) = B(dF, dG) = \{F, G\}.$$

Now by Lemma 9.3, the Jacobi identity for the Poisson structure implies $d\omega = 0$. □

Remark. The isomorphism $\lambda \mapsto -i_\lambda B$ is the inverse of $v \mapsto i_v\omega$.

A Poisson manifold $(M, \{\ ,\ \})$ carries a natural distribution $\mathcal{F}$ spanned by the Hamiltonian vector fields,

$$\mathcal{F}_x := \{X_F \mid F \in \Omega^0(M)\} = \{i_\lambda B \mid \lambda \in T_x^*M\}.$$

Theorem 9.8 (symplectic leaves)**.** *The leaves of the distribution $\mathcal{F}$ are injectively immersed submanifolds, and the restriction of the Poisson structure to each leaf is symplectic.*

Proof. The first statement follows from the General Frobenius Theorem A.5 once we find a collection of vector fields tangent to $\mathcal{F}$ whose flows leave $\mathcal{F}$ invariant, and such that every $\mathcal{F}_x$ is spanned by these vector fields. But Hamiltonian vector fields provide such a collection: Their flows preserve the Poisson structure and therefore also $\mathcal{F}$.

For the second statement, note that for a function F which is constant on a leaf $\mathcal{L}$, $X_F \cdot G = -X_G \cdot F = 0$ for every G, thus $X_F = 0$ along $\mathcal{L}$. This shows that the Hamiltonian vector field of a function at a leaf depends only on the restriction of the function to the leaf, so the Poisson structure on M induces Poisson structures on the leaves. Since the tangent space to a leaf is spanned by Hamiltonian vector fields, the Poisson structure on the leaf is nondegenerate, hence symplectic by Lemma 9.7. □

Corollary 9.9 (Darboux' Theorem for Poisson structures of constant rank)**.** *Around every point near which the symplectic leaves of a Poisson structure have constant dimension there exist coordinates q_i, p_i, z_k in which the Poisson structure satisfies*

$$\{q_i, q_j\} = \{p_i, p_j\} = 0, \quad \{q_i, q_j\} = \delta_{ij};$$
$$\{z_k, q_i\} = \{z_k, p_i\} = \{z_k, z_l\} = 0.$$

Proof. Let z_k be functions that are constant on the leaves and whose differentials are a basis of the orthogonal complement to the leaves (in the cotangent spaces). The Hamiltonian vector field of z_k vanishes, so $\{z_k, F\} = 0$ for every function F. Now apply a parametric version of the symplectic Darboux' Theorem to find symplectic coordinates q_i, p_i on the leaves. □

Remark. At a point x of nonconstant rank it is still true that the Poisson structure splits into a symplectic structure and a transverse Poisson structure which is singular at x ([24]). But the transverse structure can be arbitrarily complicated (see [1] for the singularity theory of 2-dimensional Poisson structures).

Casimir functions

A *Casimir function* for a Poisson structure is a function C that Poisson commutes with all other functions, $\{C, F\} = 0$ for all F. The importance of Casimir functions lies in the obvious fact that they are preserved under *any* Hamiltonian flow.

If the Poisson structure is nondegenerate the constants are the only Casimir functions. The coordinates z_k in Darboux' Theorem are local Casimir functions, so locally the symplectic leaves are the common level sets of Casimir functions.

However, there may not exist any nonconstant global Casimir functions for a degenerate structure:

Problem 9.3. Define a Poisson structure on the 3-torus whose symplectic leaves form a 2-dimensional linear foliation with irrational slope, and conclude that this structure admits no nonconstant Casimir functions.

Problem 9.4. Prove that the Casimir functions for the rigid body bracket on $\mathbb{R}^3$ are precisely the functions of the form $C(\xi) = f(\|\xi\|)$.

Problem 9.5. Let $(\Sigma, \{\ ,\ \})$ be a 2-dimensional Poisson manifold.

(i) Prove that the Jacobi identity is automatically satisfied.

(ii) Which values can the rank of the Poisson tensor take?

(iii) Investigate flows of time-independent Hamiltonians on $(\Sigma, \{\ ,\ \})$. Which orbit types occur?

(iv) Prove: If a compact $(\Sigma, \{\ ,\ \})$ admits a time-independent Hamiltonian flow with only 2 stationary points, then $\Sigma = S^2$ and the Poisson structure is symplectic.

(v) Find a time-independent Hamiltonian flow on a $(\Sigma, \{\ ,\ \})$ which cannot be a Hamiltonian flow with respect to any symplectic structure. Hint: Hamiltonian flows with respect to symplectic structures satisfy Poincaré recurrence.

Chapter 10

Coadjoint orbits

10.1 Actions of compact Lie groups

Definitions

Throughout this section, G is a compact Lie group and M a manifold. A *(left) action* of G on M is a group homomorphism $G \mapsto Diff(M)$ such that the map

$$G \times M \to M, \quad (g, x) \mapsto g \cdot x$$

is smooth. Similarly, a right action is group antihomomorphism $G \mapsto Diff(M)$ such that the map

$$G \times M \to M, \quad (g, x) \mapsto x \cdot g$$

is smooth. The map $g \mapsto g^{-1}$ interchanges left and right actions. Unless otherwise stated, all actions will be left actions.

Denote the orbit of x by

$$G \cdot x := \{g \cdot x \mid g \in G\}$$

and the stabilizer (or isotropy group) of x by

$$G_x := \{g \in G \mid g \cdot x = x\}.$$

Note that $G_{g \cdot x} = gG_xg^{-1}$, so all stabilizers along an orbit are conjugate. This conjugacy class (G_x) is called the *type* of the orbit.

The *infinitesimal (or linearized) action* is

$$\mathfrak{g} \mapsto \mathcal{X}(M), \quad X \mapsto \underline{\mathrm{X}}_x = \underline{\mathrm{X}}_x^M := \frac{d}{dt}\Big|_{t=0} \exp tX \cdot x.$$

Lemma 10.1. *The infinitesimal action has the following properties:*

(i) The map $\mathfrak{g} \mapsto \mathcal{X}(M)$, $X \mapsto \underline{X}$ *is linear.*

(ii) $\underline{\mathrm{Ad}_g X} = g_* \underline{X}$ *for* $g \in G$.

(iii) $\underline{[X, Y]} = -[\underline{X}, \underline{Y}]$.

Proof. For (i) just note that for each $x \in M$, $X \mapsto \underline{X}_x$ is the differential of the map $G \to M$, $g \mapsto g \cdot x$.
(ii) follows from $\exp(\mathrm{Ad}_g X) = g(\exp X)g^{-1}$:

$$\begin{aligned} \underline{\mathrm{Ad}_g X}_x &= \frac{d}{dt}\Big|_{t=0} \exp(t\mathrm{Ad}_g X) \cdot x = \frac{d}{dt}\Big|_{t=0} g(\exp tX)g^{-1} \cdot x \\ &= T_{g^{-1}x} g \cdot \underline{X}_{g^{-1}x} = (g_* \underline{X})_x . \end{aligned}$$

(iii) follows from (i), (ii) and $[X, Y] = \frac{d}{dt}\big|_{t=0} \mathrm{Ad}_{\exp tX} Y$:

$$\begin{aligned} \underline{[X, Y]} &= \frac{d}{dt}\Big|_{t=0} \underline{\mathrm{Ad}_{\exp tX} Y} \\ &= \frac{d}{dt}\Big|_{t=0} (\exp tX)_* \underline{Y} = -[\underline{X}, \underline{Y}] \end{aligned}$$

because $\exp tX$ is the flow of $\underline{X}$. □

An action is called

- *free* if $G_x = \{e\}$ for all x;
- *effective* if $\cap_{x \in M} G_x = \{e\}$.

A free right action of G on the manifold P gives rise to a G-principal bundle

$$P \to B := P/G.$$

Here a G-principal bundle $P \to B$ is a locally trivial fibration with fibre G and a free fibre-preserving right action of G on P. For a left action of G on another manifold M we get an *associated bundle*

$$P \times_G M := P \times M \Big/ (p, x) \sim (pg^{-1}, gx), g \in G.$$

$P \times_G M \to B$, $[p, x] \mapsto [p]$ is a fibre bundle with fibre M.
A subgroup $H < G$ acts on G by right multiplication and thus induces an H-principal bundle $G \mapsto G/H$. A left action of H on M yields the associated bundle $G \times_H M \to G/H$. In this case we have an additional action of G on $G \times_H M$ by left multiplication,

$$\tilde{g} \cdot [g, x] := [\tilde{g}g, x].$$

Below we will consider the case that H acts linearly on a vector space V, so that $G \times_H V \to G/H$ is a vector bundle with fibre V.

The Slice Theorem

Consider a left action $G \times M \to M$ and a point $x \in M$. To economize on notation, denote the linearized action of G_x on T_xM again by $(g, v) \mapsto g \cdot v$. This action leaves the tangent space to the orbit $T_x(G \cdot x)$ invariant: $T_x(G \cdot x)$ is generated by the $\underline{X}_x$, $X \in \mathfrak{g}$, and for $h \in G_x$,

$$h \cdot \underline{X}_x \equiv \frac{d}{dt}\Big|_{t=0} h \exp tX \cdot x = \underline{Ad_h X}_x.$$

Thus we get an induced linear action of G_x on

$$V_x := T_xM/T_x(G \cdot x).$$

Theorem 10.2 (Slice Theorem)**.** *Let a compact Lie group G act on a manifold M, and $x \in M$. Then there exists a G-equivariant diffeomorphism from a neighbourhood of the zero section in $G \times_{G_x} V_x$ onto a neighbourhood of the orbit $G \cdot x$ in M.*

Proof. Pick a G-invariant Riemannian metric on M and let $\exp_x : T_xM \to M$ be the corresponding exponential map (defined on a neighbourhood of zero). Since G maps geodesics to geodesics,

$$g \cdot \exp_x v = \exp_{g \cdot x} g \cdot v.$$

Identify V_x with the orthogonal complement $T_x(G \cdot x)^\perp$ of $T_x(G \cdot x)$ in T_xM, and consider the map

$$G \times V_x \to M, \quad (g, v) \mapsto g \cdot \exp_x v.$$

This map is G_x-invariant, $gh^{-1} \cdot \exp_x(h \cdot v) = g \cdot \exp_x v$ for $h \in G_x$, and thus induces a map

$$f : G \times_{G_x} V_x \to M, \quad [g, v] \mapsto g \cdot \exp_x v.$$

The map f is G-equivariant:

$$f([\tilde{g}g, x]) = \tilde{g}g \cdot \exp_x v = \tilde{g} \cdot f([g, x]).$$

I claim that Tf is bijective at points of zero section. In view of the G-equivariance, it suffices to compute the differential at the point $[e, 0]$:

$$T_{[e,0]}f \cdot [X, w] = \underline{X}_x + w, \quad X \in \mathfrak{g}, w \in V_x.$$

This is surjective and hence bijective by dimension reasons.

Since f embeds the zero section, it follows from the problem below that f embeds a neighbourhood of the zero section. $\square$

Problem 10.1. Let $f : N \to M$ be a smooth map between manifolds of the same dimension. Let $Z \subset N$ be a compact submanifold such that $f|_Z$ is injective and Tf is bijective at points of Z. Then f is an embedding on a neighbourhood of Z.

Proof. Bijectivity of Tf implies that there exists an open covering $U_1, \dots, U_k$ of Z such that f_{U_i} are embeddings. Let $\varepsilon > 0$ be a Lebesgue number for this covering, i.e. such that $d(x, y) < \varepsilon$ for $x, y \in \cup U_i$ implies that x and y lie in the same U_i for some i. Injectivity of $f|_Z$ implies the existence of a neighbourhood V of Z such that $d(x, y) \geq \varepsilon$ for $x, y \in V$ implies $f(x) \neq f(y)$. Thus f is an embedding on $\cup U_i \cap V$. □

Let us examine the stabilizers of the G-action in a tube $G \times_H V$. We have $g \cdot [e, v] = [e, v]$ iff there exists a $h \in H$ such that $gh^{-1} = e$ and $hv = v$. So $G_{[e,v]} = H_v$, and therefore

$$G_{[g,v]} = gH_v g^{-1}.$$

In particular, $G_{[g,v]}$ is conjugate to H iff $H_v = H$, i.e. v lies in the linear subspace $V_{(H)} < V$ of fixed points of H. So the union of orbits in $G \times_H V$ of orbit type (H) is

$$\{[g, v] \mid v \in W\} = G \times_H V_{(H)} = G/H \times V_{(H)},$$

which is a subbundle (in particular a submanifold) of $G \times_H V$. In view of the Slice Theorem, this proves

Proposition 10.3. *The union $M_{(H)}$ of orbits of a given orbit type (H) is a submanifold of M.*

Corollary 10.4. *The fixed point set $M_{(G)}$ of G is a submanifold of M.*

Proposition 10.5. *If M is compact there are only finitely many orbit types.*

Proof. The proof goes by induction on $n = \dim M$. The case $n = 0$ is clear. For the induction step, in view of the Slice Theorem and compactness of M, it suffices to show that every tube $G \times_H V$ contains only finitely many orbit types. Let SV be the unit sphere with respect to an H-invariant inner product on V. Since the action of H on V is linear, $[g, \lambda v]$ and $[g, v]$ have the same orbit type (H_v) for $\lambda \neq 0$. Thus the orbit types in $G \times_H V$ are that of the central orbit $G \times_H 0$ and the orbit types in $G \times_H SV$, which is a finite number by the induction hypothesis. □

Proposition 10.6. *If M/G is connected there is one orbit type (H) for which $M_{(H)}$ is open and dense in M. Moreover, $M_{(H)}/G$ is a connected manifold.*

Definition. Such orbits are called *principal orbits*

Proof. First let us reduce to the case that M and G are connected. If $M = \cup M_i$ is disconnected, let $G_i := \{g \in G \mid g \cdot M_i = M_i\}$. Let M_i^* be the union of G_i-principal orbits in M_i. For $i \neq j$ there exists a $g \in G$ such that $g \cdot M_i = M_j$, and therefore $g \cdot M_i^* = M_j^*$ since both are open and dense. This shows that M_i^* and M_j^* have the same G-principal orbit type (conjugated by g).

Next suppose that M is connected and $G = \cup g_j G_0$ disconnected, where G_0 is the identity component. Let M_0^* be the union of the principal orbits of the

G_0-action on M. Then $M^* := \cup g_j M_0^*$ is open and dense, and all orbits in M^* have the same type.

So it suffices to prove the first statement for M and G connected. The proof uses again induction on $n = \dim M$. The case $n = 0$ is clear. For the induction step, choosing a locally finite covering of M by tubes as in the Slice Theorem, it suffices to prove the statement for a tube $G \times_H V$. Let SV be the unit sphere with respect to an H-invariant inner product. If $G \times_H SV$ is connected the statement follows from linearity of the H-action on V and the induction hypothesis. Since G is connected, the only way that $G \times_H SV$ can be disconnected is if $\dim V = 1$ and $G \times_H SV \to G/H$ is the trivial two-fold covering. But in this case H acts trivially on V, so all orbits in $G \times_H V = G/H \times V$ are of the same type, and the induction step follows.

Connectivity of $M_{(H)}/G$ follows also by induction. To see that $M_{(H)}/G$ is a manifold, note that at a principal orbit $G \cdot x$ the representation of G_x on V_x is trivial, thus $(G \times_{G_x} V_x)/G \cong V_x$ is a manifold. □

We conclude this section with a result on the adjoint representation.

Proposition 10.7. *For the adjoint representation of a compact Lie group on its Lie algebra,*

(i) the stabilizer of any point contains a maximal torus;

(ii) the stabilizer of a point on a principal orbit is abelian.

Proof. (i) Any $X \in \mathfrak{g}$ is contained in the Lie algebra $\mathfrak{t}$ of a maximal torus T. But then the infinitesimal stabilizer of X, $\mathfrak{g}_X = \ker(\mathrm{ad}_X)$, contains $\mathfrak{t}$. So G_X contains T.

(ii) Let $\langle\ ,\ \rangle$ be an Ad-invariant inner product on $\mathfrak{g}$. Then the slice of an element $X \in \mathfrak{g}$ is

$$\begin{aligned} V_X &\equiv T_X(\mathrm{Ad}_G \cdot X)^\perp = [\mathfrak{g}, X]^\perp \\ &= \{Y \in \mathfrak{g} \mid 0 = \langle Y, [Z, X]\rangle = \langle Z, [X, Y]\rangle \text{ for all } Z \in \mathfrak{g}\} \\ &= \ker(\mathrm{ad}_X) \\ &= \mathfrak{g}_X. \end{aligned}$$

Now suppose X lies on a principal orbit. By the proof of Proposition 10.3, the stabilizer of $X + Y$ equals the stabilizer of X for any $Y \in V_X$ sufficiently close to zero. For the infinitesimal stabilizers this means that $\mathfrak{g}_{X+Y} = \mathfrak{g}_X$ for all $Y \in \mathfrak{g}_X$ close to zero. In particular, $0 = [X + Y, Z] = [Y, Z]$ for all $Y, Z \in \mathfrak{g}_X$ close to zero. This proves that $\mathfrak{g}_X$ is abelian. □

The following problem shows that the stabilizer of a principal orbit can be strictly bigger than a maximal torus:

Problem 10.2. Find all the stabilizers of the adjoint representations of $SO(3)$ and $SU(2)$ on $\mathbb{R}^3$.

10.2 The Lie-Poisson bracket

The interest in Poisson manifolds mainly arises from the fact that the dual $\mathfrak{g}^*$ of any Lie algebra $\mathfrak{g}$ carries a natural Poisson structure, the *Lie-Poisson bracket*

$$\{F, G\}(\xi) := \langle \xi, [d_\xi F, d_\xi G] \rangle, \quad F, G \in \Omega^0(\mathfrak{g}^*), \xi \in \mathfrak{g}^*.$$

Here $d_\xi F$ lies in $T^*_\xi \mathfrak{g}^*$ which is naturally identified with $\mathfrak{g}$.

Proposition 10.8. *The Lie-Poisson bracket defines a Poisson structure on $\mathfrak{g}^*$.*

Proof. Skew-symmetry and the derivation property are obvious. For the Jacobi identity, first compute the differential of $\{F, G\}$ at ξ in the direction $\eta \in T_\xi \mathfrak{g}^* \cong \mathfrak{g}^*$:

$$\begin{aligned} d\{F, G\} \cdot \eta &= \langle \eta, [dF, dG] \rangle + \langle \xi, [D^2 F \cdot \eta, dG] \rangle + \langle \xi, [dF, D^2 G \cdot \eta] \rangle \\ &= \langle \eta, [dF, dG] \rangle - \langle \mathrm{ad}^*_{dG} \xi, D^2 F \cdot \eta \rangle + \langle \mathrm{ad}^*_{dF} \xi, D^2 G \cdot \eta \rangle \\ &= \langle \eta, [dF, dG] \rangle - D^2 F(\mathrm{ad}^*_{dG} \xi, \eta) + D^2 G(\mathrm{ad}^*_{dF} \xi, \eta). \end{aligned}$$

So the differential of $\{F, G\}$ at ξ is

$$d\{F, G\} = [dF, dG] - D^2 F(\mathrm{ad}^*_{dG} \xi, \cdot) + D^2 G(\mathrm{ad}^*_{dF} \xi, \cdot),$$

and

$$\begin{aligned} \{\{F, G\}, H\} &= \langle \xi, [d\{F, G\}, dH] \rangle \\ &= \Big\langle \xi, \big[[dF, dG], dH\big] \Big\rangle - \Big\langle \mathrm{ad}^*_{dH} \xi, -D^2 F(\mathrm{ad}^*_{dG} \xi, \cdot) + D^2 G(\mathrm{ad}^*_{dF} \xi, \cdot) \Big\rangle \\ &= \Big\langle \xi, \big[[dF, dG], dH\big] \Big\rangle + D^2 F(\mathrm{ad}^*_{dG} \xi, \mathrm{ad}^*_{dH} \xi) - D^2 G(\mathrm{ad}^*_{dF} \xi, \mathrm{ad}^*_{dH} \xi). \end{aligned}$$

Adding up cyclic permutations, the terms involving second derivatives cancel because the $D^2 F$ is symmetric in the two arguments, and the terms involving triple Lie brackets cancel by the Jacobi identity on $\mathfrak{g}$. ☐

Example 10.9. (rigid body bracket) The rigid body bracket is the Lie-Poisson bracket on the dual of the Lie algebra $(\mathbb{R}^3, \times)$.

Let us compute the Hamiltonian vector field in the Lie-Poisson bracket:

$$X_H \cdot F(\xi) = \{F, H\}(\xi) = \langle \xi, [dF, dH] \rangle = -\langle \mathrm{ad}^*_{dH} \xi, dF \rangle$$

for all functions F, thus

Lemma 10.10. *The Hamiltonian vector field in the Lie-Poisson bracket is*

$$X_H(\xi) = -\mathrm{ad}^*_{d_\xi H} \xi.$$

But

$$\underline{X}_\xi^{\mathfrak{g}^*} = -\mathrm{ad}_X^*\xi$$

is just the fundamental vector field of $X \in \mathfrak{g}$ associated to the coadjoint (left) action

$$G \times \mathfrak{g}^* \to \mathfrak{g}^*, \quad (g, \xi) \mapsto \mathrm{Ad}_{g^{-1}}^*\xi.$$

So the Hamiltonian vector fields span precisely the tangent spaces to the coadjoint orbits! Combined with the Symplectic Leaf Theorem this implies

Proposition 10.11. *The coadjoint orbits in $\mathfrak{g}^*$ are the symplectic leaves of the Lie-Poisson bracket, with the symplectic form*

$$\omega_\xi(\underline{X}_\xi^{\mathfrak{g}^*}, \underline{Y}_\xi^{\mathfrak{g}^*}) := \langle \xi, [X, Y] \rangle, \quad X, Y \in \mathfrak{g}.$$

10.3 Examples of coadjoint orbits

Example 10.12. ($G = SO(3)$). Identify the Lie algebra $\mathfrak{so}(3)$ with $\mathbb{R}^3$ via

$$\mathbb{R}^3 \ni \begin{pmatrix} x \\ y \\ z \end{pmatrix} = X \mapsto \hat{X} := \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \in \mathfrak{so}(3).$$

To understand this identification, note that $\left(\begin{smallmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{smallmatrix}\right)$ is a counterclockwise rotation in the plane with infinitesimal rotation $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$.
Under this identification, the adjoint action on $\mathfrak{so}(3)$ corresponds to left multiplication on $\mathbb{R}^3$ by $SO(3)$ matrices, the Lie bracket corresponds to the cross product, and $-\frac{1}{2}\mathrm{tr}$ corresponds to the dot product $\langle \ , \ \rangle$ on $\mathbb{R}^3$:

$$\begin{aligned} Ad_A\hat{X} &= A \cdot X; \\ [\hat{X}, \hat{Y}] &= X \times Y; \\ -\frac{1}{2}\mathrm{tr}(\hat{X}\hat{Y}) &= \langle X, Y \rangle. \end{aligned}$$

If we identify $\mathfrak{so}(3)$ with its dual via this inner product, the coadjoint action becomes

$$\mathrm{Ad}_A^*\hat{\xi} = -A \cdot \xi, \quad \mathrm{ad}_{\hat{X}}^*\hat{\xi} = -X \times \xi.$$

The coadjoint orbits are the spheres of radius $r > 0$,

$$\mathcal{O}_r := \{\xi \mid \|\xi\| = r\},$$

and the singular orbit $\{0\}$.

Claim: The induced symplectic form on $\mathcal{O}_r$ is $\frac{1}{r}\sigma_r$, where σ_r is the standard surface element on $\mathcal{O}_r$.

Proof. Use the vector identity on $\mathbb{R}^3$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{B}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$$

to get $\xi \times (v \times \xi) = v\|\xi\|^2 - \xi(\xi \cdot v)$. This shows that given $v \in T_\xi \mathcal{O}_r$, the vector $X = \frac{v \times \xi}{\|\xi\|^2}$ satisfies $\xi \times X = v$, so X is a Lie algebra element whose fundamental vector field at ξ equals v. Now by definition the value of the symplectic form on $v, w \in T_\xi \mathcal{O}_r$ is

$$\begin{aligned}
\omega_\xi(v, w) &= \langle \xi, \frac{(v \times \xi) \times (w \times \xi)}{\|\xi\|^4} \rangle \\
&= \langle \frac{\xi}{\|\xi\|^4}, w\big((v \times \xi) \cdot \xi\big) - \xi\big((v \times \xi) \cdot w\big) \rangle \\
&= -\frac{(v \times \xi) \cdot w}{\|\xi\|^2} = \frac{(v \times w) \cdot \xi}{\|\xi\|^2}.
\end{aligned}$$

Thus for with $\xi = (x, y, z)$, ω can be written as

$$\omega_\xi = \frac{x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy}{\|\xi\|^2} = \frac{\sigma}{\|\xi\|}.$$

$\square$

Problem 10.3. Describe the coadjoint orbits on $\mathfrak{so}(4)$.

Example 10.13. ($G = U(n)$).

The Lie algebra $\mathfrak{u}(n)$ carries the Ad-invariant inner product $\mathrm{tr}(AB^*) = -\mathrm{tr}(AB)$. Identify $\mathfrak{u}(n)^*$ with the space $\mathcal{H}$ of Hermitian matrices, acting on $\mathfrak{u}(n)$ via

$$\langle \xi, X \rangle = \mathrm{tr}(i\xi X), \quad \xi \in \mathcal{H}, X \in \mathfrak{u}(n).$$

With these identifications, for $A \in U(n)$, $X \in \mathfrak{u}(n)$, $\xi \in \mathcal{H}$,

$$\mathrm{Ad}^*_A \xi = A^{-1}\xi A, \quad \mathrm{ad}_X \xi = \xi X - X\xi.$$

The coadjoint orbits are

$$\mathcal{H}_\lambda = \{\xi \in \mathcal{H} \mid \mathrm{spectrum}(\xi) = \lambda\},$$

parametrized by the eigenvalues $\lambda = (\lambda_1, \ldots, \lambda_n)$, $\lambda_1 \leq \cdots \leq \lambda_n$.

If $\lambda_1 < \cdots < \lambda_n$, then a matrix $\xi \in \mathcal{H}_\lambda$ is characterized by the eigenspaces L_j of the λ_j, or equivalently, by the *complete flag*

$$E_1 \subset E_2 \subset \cdots \subset E_n = \mathbb{C}^n, \quad E_i = \bigoplus_{j \leq i} L_j.$$

Thus the coadjoint orbit $\mathcal{H}_\lambda$ is the *complete flag manifold* on $\mathbb{C}^n$. The coadjoint action on $\mathcal{H}$ corresponds to simultaneous multiplication of all subspaces of a flag by a $U(n)$-matrix. To each vector λ there corresponds a symplectic structure on the complete flag manifold induced from the Lie-Poisson bracket on $\mathcal{H}$.

If $\lambda_1 < \lambda_2 = \cdots = \lambda_n$, then a matrix $\xi \in \mathcal{H}_\lambda$ is characterized by the eigenspace L_1 to the eigenvalue λ_1. So the coadjoint orbit $\mathcal{H}_\lambda$ is diffeomorphic to $\mathbb{C}P^{n-1}$. The coadjoint action on $\mathcal{H}_\lambda$ corresponds to the standard linear action of $U(n)$ on $\mathbb{C}P^{n-1}$. Different values of λ_1, λ_2 yield a 2-parameter family of symplectic structures on $\mathbb{C}P^{n-1}$.

For other types of λ the coadjoint orbits are noncomplete flag manifolds.

Problem 10.4. Express the symplectic forms on $\mathcal{H}_\lambda$, $\lambda_1 < \lambda_2 = \cdots = \lambda_n$, in terms of the standard form on $\mathbb{C}P^{n-1}$.

Problem 10.5. Describe the 6-manifold $\mathcal{H}_\lambda$, $\lambda_1 < \lambda_2 < \lambda_3$, and the induced symplectic forms on it.

Appendix A

Calculus on manifolds

A.1 Flows and vector fields

A time-dependent vector field $X(t,x) = X_t(x)$ on a manifold M defines a *flow*, i.e. a 1-parameter family of diffeomorphisms $\phi_t^X : M \to M$, via

$$\frac{d}{dt}\phi_t(x) = X\big(t, \phi_t(x)\big).$$

The flow may only exist for some finite time interval, but for simplicity let us assume it exists for all times. Every smooth 1-parameter family of diffeomorphisms is the flow of a time-dependent vector field. The flow is a *1-parameter group*, $\phi_{s+t}^X = \phi_s^X \circ \phi_t^X$, if and only if the vector field X is time-independent.

The following lemma describes how operations on flows correspond to operations on vector fields.

Lemma A.1. *Let X, Y be time-dependent vector fields with flows ϕ_t^X, ϕ_t^Y, ψ a diffeomorphism, and $h(t,x), f(x)$ smooth functions.*

(i) (composition). $\phi_t^X \circ \phi_t^Y$ is the flow of the vector field $X_t + (\phi_t^X)_* Y_t$.

(ii) (inverse). $(\phi_t^X)^{-1}$ is the flow of the vector field $-(\phi_t^X)^* X_t$.

(iii) (conjugation). $\psi \circ \phi_t^X \circ \psi^{-1}$ is the flow of the vector field $\psi_* X_t$.

(iv) (reparametrization). $\phi_{h(t,x)}^X$ is the flow of the vector field Y_t defined by

$$Y_t\Big(\phi_{h(t,x)}^X(x)\Big) = \frac{\partial h}{\partial t}(t,x) X_{h(t,x)}\Big(\phi_{h(t,x)}^X(x)\Big).$$

Conversely, if X is time-independent, the flow of $f(x)X$ is $\phi_{h(t,x)}^X$, where for every x the function $t \mapsto h(t,x)$ is the unique solution of the ordinary *differential equation*

$$\frac{\partial h}{\partial t}(t,x) = f\Big(\phi^X_{h(t,x)}(x)\Big);$$
$$h(0,x) = 0.$$

Proof. 1.

$$\begin{aligned}\frac{d}{dt}(\phi_t^X \circ \phi_t^Y) &= X_t(\phi_t^X \circ \phi_t^Y) + T\phi_t^X \cdot Y_t(\phi_t^Y) \\ &= X_t(\phi_t^X \circ \phi_t^Y) + (\phi^X_{t*}Y_t)(\phi_t^X \circ \phi_t^Y).\end{aligned}$$

2.

$$\begin{aligned}0 &= \frac{d}{dt}\Big((\phi_t^X)^{-1} \circ \phi_t^X\Big) = Y_t + T(\phi_t^X)^{-1} \cdot X_t(\phi_t^X) \\ &= Y_t + (\phi_t^X)^* X_t.\end{aligned}$$

3.

$$\begin{aligned}\frac{d}{dt}(\psi \circ \phi_t^X \circ \psi^{-1}) &= T\phi \cdot X_t(\phi_t^X \circ \psi^{-1}) \\ &= \psi_* X_t(\psi \circ \phi_t^X \circ \psi^{-1}).\end{aligned}$$

4.

$$\frac{d}{dt}\phi^X_{h(t,x)} = \frac{\partial h}{\partial t}(t,x) X_{h(t,x)}(\phi^X_{h(t,x)}) =: Y_t(\phi^X_{h(t,x)}).$$

□

Form now on let all vector fields be time-independent. The *Lie derivative* of any tensor field τ in the direction of a vector field X is defined as

$$L_X\tau := \frac{d}{dt}\Big|_{t=0}(\phi_t^X)^*\tau.$$

The Lie derivative is related to flows as follows:

Lemma A.2. *Let X, Y be vector fields and ψ a diffeomorpism*

(i) $\psi_* X = \frac{d}{dt}\big|_{t=0}\psi \circ \phi_t^X \circ \psi^{-1}$*;*

(ii) $\psi_* X = X$ *if and only if* $\psi \circ \phi_t^X = \phi_t^X \circ \psi$*;*

(iii) $L_X Y = [X, Y]$*;*

(iv) $[X, Y] = 0$ *if and only if the flows commute,* $\phi_t^X \circ \phi_s^Y = \phi_s^Y \circ \phi_t^X$*.*

Proof. 1.
$$\frac{d}{dt}\Big|_{t=0}\psi\circ\phi_t^X\circ\psi^{-1} = T\psi\cdot X(\psi^{-1}) = \psi_* X.$$
2.
$$\begin{aligned}\frac{d}{ds}(\psi\circ\phi_t^X) &= T\psi\cdot X(\phi_t^X) = (\psi_*X)(\psi\circ\phi_t^X),\\ \frac{d}{ds}(\phi_t^X\circ\psi) &= X(\phi_t^X\circ\psi).\end{aligned}$$
If $\psi_*X = X$ these equations define the same flow.
3. By property 1 we get for a function f:
$$\begin{aligned}L_X Y\cdot f &= \frac{\partial^2}{\partial s\partial t} f\circ(\phi_t^X)^{-1}\circ\phi_s^Y\circ\phi_t^X\\ &= \frac{\partial^2}{\partial s\partial t}(f\circ\phi_s^Y\circ\phi_t^X - f\circ\phi_t^X\circ\phi_s^Y)\\ &= [X,Y]\cdot f.\end{aligned}$$
4. $[X,Y]=0$ implies $\frac{d}{dt}(\phi_t^X)^*Y = (\phi_t^X)^*L_XY = 0$, so $(\phi_t^X)^*Y = Y$. By property 2 this implies $\phi_t^X\circ\phi_s^Y = \phi_s^Y\circ\phi_t^X$. □

The following formula plays an important role in symplectic geometry:

Lemma A.3 (Cartan's formula)**.** *For a vector field X and a k-form α on a manifold,*
$$L_X\alpha = i_X d\alpha + d\, i_X\alpha.$$

Proof. Let ϕ_t be the flow of X and γ a k-chain. Then $\Gamma := \phi([0,t]\times\gamma)$ is a $(k+1)$-chain, and Stokes' Theorem reads
$$\int_\Gamma d\alpha = \int_{\phi_t(\gamma)}\alpha - \int_\gamma\alpha - \int_{\phi([0,t]\times\partial\gamma)}\alpha.$$
The t-derivative at $t=0$ and another application of Stokes' Theorem yield
$$\int_\gamma i_X d\alpha = \int_\gamma L_X\alpha - \int_{\partial\gamma} i_X\alpha = \int_\gamma L_X\alpha - \int_\gamma d\, i_X\alpha.$$
□

The formulae summarized below are easy consequences of the definitions and the results above.

Summary: formulae for Lie derivative and exterior derivative

Let X, X_i, Y be vector fields, σ, τ tensor fields, f a function and α a k-form.

(i) $L_X f = X \cdot f$.

(ii) $L_X Y = [X, Y]$.

(iii) $L_X i_\tau \sigma = i_{L_X \tau}\sigma + i_\tau L_X \sigma$.

(iv) Properties 1-3 determine the Lie derivative uniquely on arbitrary tensor fields. In particular,

$$L_X \alpha(X_1, \ldots, X_k) = X \cdot \alpha(X_1, \ldots, X_k) - \sum_{i=1}^{k} \alpha(X_1, \ldots, [X, X_i], \ldots, X_k).$$

(v) (Cartan's formula). $L_X \alpha = d\, i_X \alpha + i_x d\alpha$.

(vi)

$$\begin{aligned} d\alpha(X_0, \ldots, X_k) = &\sum_{i=0}^{k} (-1)^i X_i \cdot \alpha(X_0, \ldots, \hat{X}_i, \ldots, X_k) \\ &+ \sum_{0 \le i < j \le k} (-1)^{i+j} \alpha([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_k), \end{aligned}$$

where $\hat{X}_i$ means the X_i is omitted.

Lemma A.4 (Relative Poincaré Lemma)**.** *Let M be a manifold that deformation retracts onto a closed submanifold $A \subset M$. Let α be a closed k-form on M such that $\alpha|_W = 0$. Then there exists a $(k-1)$-form β on M such that $\beta = 0$ along W and $d\beta = \alpha$.*

Proof. (cf. citeCdS, [11]). Let $\phi_s : M \to M$ be the deformation retraction such that $\phi_1 = \mathbb{1}$, $\phi_0 : M \to W$, and $\phi_s|_W = \mathbb{1}$. Define vectors $X_s\big(\phi_s(x)\big) := \frac{d}{ds}\phi_s(x)$, and the homotopy operator $P : \Omega^k(M) \to \Omega^{k-1}(M)$,

$$P\alpha := \int_0^1 \phi_s^*(i_{X_s}\alpha) ds.$$

Although X_s is not a vector field, the explicit form

$$(P\alpha)_x = \int_0^1 i_{X_s\left(\phi_s(x)\right)} \alpha_{\phi_s(x)} \circ T_x\phi_s \; ds$$

shows that P is well-defined. By Cartan's formula,

$$\begin{aligned} dP\alpha - Pd\alpha &= \int_0^1 \Big(\phi_s^* di_{X_s}\alpha - \phi_s^* i_{X_s} d\alpha\Big) ds \\ &= \int_0^1 \phi_s^* L_{X_s}\alpha\, ds = \int_0^1 \frac{d}{ds}\phi_s^*\alpha\, ds \\ &= \phi_1^*\alpha - \phi_0^*\alpha. \end{aligned}$$

Using $\phi_1 = \mathbb{1}$ and $\phi_0 = i \circ \pi$ where $i : W \to M$ is the inclusion and $\pi : M \to W$, we obtain the *homotopy formula*

$$dP\alpha - Pd\alpha = \alpha - \pi^* i^* \alpha.$$

moreover, $\psi_s|_W = \mathbb{1}$ implies that $P\alpha$ vanishes along W. If $d\alpha = 0$ and $\alpha|_W = 0$ the homotopy formula becomes $dP\alpha = \alpha$, so $\beta := P\alpha$ is the desired $(k-1)$-form. □

A.2 Frobenius' Theorem

Consider a distribution $\mathcal{F}$ of subspaces $\mathcal{F}_x \subset T_xM$ on a manifold M which is *smooth* in the sense that for every $x \in M$ there exist vector fields $X_1, \ldots, X_k$ tangent to $\mathcal{F}$ such that $X_1(x), \ldots, X_k(x)$ is a basis for $\mathcal{F}_x$. This implies that the dimension of $\mathcal{F}$ is a lower semicontinuous function on M.

Define the *leaves* of a distribution as follows: Two points are on the same leaf if they can be connected by a piecewise smooth path tangent to $\mathcal{F}$.

The distribution is called *integrable* if for every x the vector fields $X_1, \ldots, X_k$ can be chosen in such a way that their flows preserve $\mathcal{F}$. Note that we do not require $\mathcal{F}$ to have constant rank.

Problem A.1. Find the simplest example of an integrable distribution which is not of constant rank.

Theorem A.5 (General Frobenius Theorem)**.** *For a smooth distribution $\mathcal{F}$ the following statements are equivalent:*

(i) The flow of every vector field tangent to $\mathcal{F}$ leaves $\mathcal{F}$ invariant.

(ii) There exists a set $\mathcal{S}$ of vector fields tangent to $\mathcal{F}$ whose flows leave $\mathcal{F}$ invariant, and such that every $\mathcal{F}_x$ is spanned by vector fields in $\mathcal{S}$.

(iii) Each leaf $\mathcal{L}$ is an injectively immersed submanifold with $T_x\mathcal{L} = \mathcal{F}_x$ for all $x \in \mathcal{L}$.

Proof. $(i) \Longrightarrow (ii)$: obvious.

(ii) $\Longrightarrow$ *(iii)*: Fix $x \in \mathcal{L}$ and let $X_1, \ldots, X_k$ be vector fields in $\mathcal{S}$ such that $X_1(x), \ldots, X_k(x)$ is a basis for $\mathcal{F}_x$. Define a map

$$\psi : \mathbf{t} = (t_1, \ldots, t_k) \mapsto \phi^1_{t_1} \circ \cdots \circ \phi^k_{t_k}(x),$$

where ϕ^i_t is the flow of X_i. This map has the following properties:

1. There exist open neighbourhoods V of x in M and $U = (-\varepsilon, \varepsilon)^k = \phi^{-1}(V)$ of $\mathbf{0}$ in $\mathbb{R}^k$ such that $\psi : U \to V$ is an embedding.

This holds because the differential at $\mathbf{0}$,

$$T_{\mathbf{0}}\psi(v_1, \ldots, v_k) = \sum_{i=1}^{k} v_i X_i(x)$$

is an injective map $\mathbb{R}^k \to T_x M$.

2. $\psi(\mathbf{t}) \in \mathcal{L}$ for all $\mathbf{t} \in U$.

Indeed, the path $x \to \phi^k_{t_k}(x) \to \phi^{k-1}_{t_{k-1}} \circ \phi^k_{t_k}(x) \to \cdots \to \psi(x)$ following the flow lines is a piecewise smooth path tangent to $\mathcal{F}$.

3. $T_{\mathbf{t}}\psi \cdot \mathbb{R}^k = \mathcal{F}_{\psi(\mathbf{t})}$ for all $\mathbf{t} \in U$.

The inclusion $\subset$ follows from

$$\begin{aligned} T_{\mathbf{t}}\psi \cdot \mathbf{e}_i &= \frac{\partial}{\partial t_i} \psi^1_{t_1} \circ \cdots \circ \phi^k_{t_k}(x) \\ &= \Big(\phi^1_{t_1} \circ \cdots \circ \phi^{i-1}_{t_{i-1}}\Big)_* X_i\Big(\phi^i_{t_i} \circ \cdots \circ \phi^k_{t_k}(x)\Big), \end{aligned}$$

which lies in $\mathcal{F}$ because $X_i \in \mathcal{F}$ and the flows preserve $\mathcal{F}$. For the other inclusion, just note that the flows preserve $\mathcal{F}$ and therefore the rank of $\mathcal{F}$ is constant ($= k$) along the leaf $\mathcal{L}$.

4. $\psi(U)$ is the path component of $\mathcal{L} \cap V$ containing x.

Clearly $\psi(U)$ is contained in the path component. For the converse, let y be a point in V that can be connected to x by a piecewise smooth path γ in V tangent to $\mathcal{F}$. Without loss of generality, suppose γ is smooth. Extend $\dot\gamma$ to a time-dependent vector field X on V tangent to $\mathcal{F}$ such that $\dot\gamma(t) = X\big(t, \gamma(t)\big)$. (This is possible by writing $\dot\gamma(t) = \sum_i f_i\big(t, \gamma(t)\big) X_i\big(\gamma(t)\big)$ and extending the functions f_i). The vector field X is tangent to the submanifold $\psi(U) \subset V$. So by the uniqueness theorem for ordinary differential equations, $\gamma(t)$ stays on $\psi(U)$, which proves that $y \in \psi(U)$.

(iii) $\Longrightarrow$ *(i)*: This is proved as Step 4 in *(ii)* $\Longrightarrow$ *(iii)*. $\square$

A distribution $\mathcal{F}$ satisfying the conditions of the General Frobenuis Theorem is called *integrable*. If $\mathcal{F}$ is integrable then it is *involutive*: If X, Y are vector fields tangent to $\mathcal{F}$, then $[X, Y]$ is tangent to $\mathcal{F}$. This follows directly from property (iii), or from property (i) as follows: Since the flow ϕ^X_t preserves $\mathcal{F}$, $\big((\phi^X_t)^* Y\big)_x \in \mathcal{F}_x$ for all t, and thus $[X, Y]_x = \frac{d}{dt}\big|_{t=0} \big((\phi^X_t)^* Y\big)_x \in \mathcal{F}_x$. If $\mathcal{F}$ has constant rank the converse is true:

Theorem A.6 (Classical Frobenius Theorem)**.** *A smooth distribution $\mathcal{F}$ of constant rank is integrable if and only if it is involutive.*

Proof. If $\mathcal{F}$ has constant rank, we can write it locally as $\mathcal{F} = \{\alpha_1 = \cdots = \alpha_l = 0\}$ for linearly independent 1-forms α_i. Suppose that $\mathcal{F}$ is involutive. Then for vector fields X, Y tangent to $\mathcal{F}$,

$$\begin{aligned} L_X\alpha_i(Y) &= (d\,i_X\alpha_i + i_X d\alpha_i)(Y) = d\alpha_i(X,Y) \\ &= X\cdot\alpha_i(Y) - Y\cdot\alpha_i(X) - \alpha_i([X,Y]) = 0. \end{aligned}$$

This shows that $L_X\alpha_i$ vanishes on $\mathcal{F}$, so it can be written as

$$L_X\alpha_i = \sum_j f_{ij}\alpha_j$$

for unique functions f_{ij}. For the flow ϕ_t of X this implies

$$\frac{d}{dt}\phi_t^*\alpha_i = \phi_t^*\sum_j f_{ij}\alpha_j = \sum_j(\phi_t^* f_{ij})\phi_t^*\alpha_j.$$

So the vector $\alpha(t) := (\phi_t^*\alpha_1, \dots, \phi_t^*\alpha_l)_x$ satisfies the linear differential equation on $(T_x^*M)^k$,

$$\frac{d}{dt}\alpha(t) = A(t)\alpha(t),$$

where $A(t)$ is the matrix with components $(\phi_t^* f_{ij})(x)$. This equation has the solution

$$\alpha(t) = \exp\Big[\int_0^t A(s)ds\Big]\alpha(0),$$

so at the point x each $\phi_t^*\alpha_i$ is a linear combination of the α_j. This shows that $\phi_t^*\alpha_i$ vanishes on $\mathcal{F}$ for all i, so ϕ_t preserves $\mathcal{F}$, and condition (i) of the General Frobenius Theorem holds. $\square$

Problem A.2. Find an example of a distribution (not of constant rank) which is involutive but not integrable.

Bibliography

[1] V.I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd edition, Springer-Verlag, New York (1989).

[2] M. Audin, *The Topology of Torus Actions on Symplectic Manifolds*, Progress in Mathematics **93**, Birkhäuser Verlag, Basel (1991).

[3] T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*, Springer-Verlag, New York (1985).

[4] A. Cannas da Silva, *Lectures on Symplectic Geometry*, Lecture Notes 2000, http://math.berkeley.edu/ acannas.

[5] K. Cieliebak, *Pseudo-holomorphic curves and periodic orbits on cotangent bundles*, Journal de Mathématiques Pures et Appliquées **73**, 251-278 (1994).

[6] W. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston (1992).

[7] H. Hofer and E. Zehnder, *Symplectic Invariants and Hamiltonian Dynamics*, Birkhäuser, Basel (1994).

[8] R. Gompf, *A new construction of symplectic manifolds*, Annals of Mathematics **142**, 527-95 (1995).

[9] M. Gromov, *Pseudo holomorphic curves in symplectic manifolds*, Inventiones Mathematicae **82**, 307-47 (1985).

[10] V. Guillemin, *Moment Maps and Combinatorial Invariants of Hamiltonian T^n-spaces*, Progress in Mathematics **122**, Birkhäuser, Boston (1994).

[11] V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, Englewood Cliffs, N.J. (1974).

[12] V. Guillemin and S. Sternberg, *Birational equivalence in the symplectic category*, Inventiones Mathematicae **97**, 485-522 (1989).

[13] V. Guillemin and S. Sternberg, *Symplectic Techniques in Physics*, second edition, Cambridge University Press, Cambridge (1990).

[14] J. Marsden and T. Ratiu, *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics **17**, Springer-Verlag, New York (1994).

[15] D. McDuff and D. Salamon, *Introduction to Symplectic Topology*, Second edition, Oxford Mathematical Monographs, Oxford University Press, New York (1998).

[16] J. Moser, *On the volume elements on a manifold*, Trans. Amer. Math. Soc. **120**, 286-294 (1965).

[17] P. Ozsváth and Z. Szabó, *The symplectic Thom conjecture*, Preprint IAS Princeton (1998).

[18] K.F. Siburg, *Symplectic capacities in two dimensions*, manuscripta math. **78**, 149-163 (1993).

[19] S. Smale, *Diffeomorphisms of the* 2*-sphere*, Proceedings of the American Mathematical Society **10**, 621-626 (1959).

[20] W. Thurston, *Some simple examples of symplectic manifolds*, Proc. Amer. Math. Soc. **55**, 467-468 (1976).

[21] J. Weber, *Topology of* $Sp(2,\mathbb{R})$ *and the Conley-Zehnder index*, Preprint 1997, http://math.sunysb.edu/ joa.

[22] A. Weinstein, *Lectures on Symplectic Manifolds*, CBMS Conference Series **29**, American Mathematical Society, Providence, RI (1977).

[23] A. Weinstein, *Symplectic manifolds and their Lagrangian submanifolds*, Advances in Math. **6**, 329-346 (1971).

[24] A. Weinstein, *The local structure of Poisson manifolds*, J. Diff. Geom. **18**, 523-557 (1983).