Lectures in Classical Mechanics and Symplectic Geometry, Herbst 2003

Klaus Mohnke

Klaus Mohnke, Humboldt–Universität zu Berlin,Institut für Mathematik, Strasse Unter den Linden, 10099 Berlin, GERMANY
E-mail address: mohnke@mathematik.hu-berlin.de

Part 1

The Geometry of Newton's Mechanics

We study the equations of motion of massive points as derived from Newton's principles of mechanics and Galilei's existence of inertial coordinate systems. Mostly the forces will be conservative which leads to the definition of the total energy of the system and its conservation. The notions and laws in Newton's mechanics are invariant under the a choice of an intertial coordinate system or, in other words the action of Galilei's transformation group on the system. Each euclidean motion which leaves the potential of the force, or potential energy of the system invariant leads to a conservation law of the system. This idea is applied to the system of two celestial bodies from which we deduce Kepler's laws of the motion of planets around the sun.

CHAPTER 1

Celestial Mechanics

1. Kepler's Laws

We study the equations of motion of a mechanical system with a point-like object of mass m. We solve them explicitely in some cases.

1.1. Conservative Systems. The *degree of freedom* of a mechanical system is a priori given as the number of euclidean coordinates, e.g. of the position of a massive point. The corresponding differential equation is of the form

$$m\ddot{x} = F(x), \qquad x = x(t) \in \mathbb{R}^n.$$

Such a system is called *conservative*, if F is the negative gradient of a real function U:

$$F(x) = -\nabla U.$$

Conservative vector fields are characterized in the following way. Given a path $\gamma : [0,1] \to \mathbb{R}^n$ we denote by

$$\int_0^1 (F(\gamma(t)) \cdot \dot{\gamma}(t))dt$$

the *work* which has to be invested to treck the point along the path γ. Notice that this does not depend on the parameterization of γ and can be therefore denoted by

$$\int_\gamma F \cdot d\gamma.$$

Proposition 1. *A vector field is conservative if and only if the work along a path γ depends on its endpoints only.*

Proof. We fix some point $x_0 \in \mathbb{R}^n$ and set

$$U(x) := -\int_{x_0}^{x} F \cdot d\gamma$$

for some path connecting the two points.

EXERCISE 1. *Show that indeed*

$$F = -\nabla U.$$

On the other hand if $F = -\nabla U$ the main theorem of calculus shows

$$\int_0^1 \nabla U \cdot \dot{\gamma} dt = \int_0^1 \frac{d}{dt}(U(\gamma(t))dt = U(\gamma(1)) - U(\gamma(0)).$$

□

U is called the *potentential energy* of the system. It is only defined up to a constant.

COROLLARY 2. *A vector field F is conservative if and only if the work along each loop vanishes.*

EXERCISE 2. *Prove the above statement.*

EXERCISE 3. *Check whether the vector field $F_1 = x_2, F_2 = -x_1$ is conservative.*

EXERCISE 4. *Check whether the vector field*

$$F_1 = \frac{x_2}{x_1^2 + x_2^2} \qquad F_2 = \frac{-x_1}{x_1^2 + x_2^2}$$

is conservative on $\mathbb{R}^2 \setminus \{0\}$ or on $\mathbb{R}^2 \setminus \mathbb{R}_+$.

Systems with one degree of freedom are always conservative. Indeed, given

$$\ddot{x} = F(x)$$

for a real function in one variable we may define its potential

$$U(x) = -\int_{x_0}^{x} F(\xi) d\xi.$$

EXAMPLE 1. *For a point of mass m near the earth the acting force is given by mg, where $g = 9.8 \ldots \frac{m}{s^2}$ is the acceleration near the surface of the earth, which was already known to Galilei. Hence the potential energy is (up to a constant)*

$$U(h) = mgh.$$

Here h is the height of the point above the surface of the earth, the only free coordinate if we study the system of a freely falling massive point.

1.2. Conservation Laws.

1.2.1. *Conservation of Energy.* Denote by

$$E := \frac{m}{2}|\dot{x}|^2 + U(x)$$

the *energy* of the object.

PROPOSITION 3. *The energy of a massive point in a conservative system remains constant in time.*

PROOF.

$$\dot{E} = \frac{d}{dt}(\frac{m}{2}|\dot{x}|^2 + U(x)) = m\dot{x}\cdot\ddot{x} + \dot{x}\cdot\nabla U(x) = \dot{x}\cdot(m\ddot{x} - F(x)) = 0.$$

□

EXERCISE 5. *The motion of a charged particle of mass m and charge c in a magnetic field neglecting gravity is determined by*

$$m\ddot{x} = cV(x)\times\dot{x}.$$

By "$\times$" we denote as usual the vector product (also cross product) in $\mathbb{R}^3$. Show that the absolute value of the velocity $|\dot{x}|$ remains constant.

1.2.2. *Conservation of the Angular Momentum in a Central Field.* We consider a mechanical system for which the vector field F is invariant under isometries which leave the origin $0 \in \mathbb{R}^n$ fixed: $gF(x) = F(gx)$ for any such transformation g. Such a vector field is called *central field.* Let us introduce polar coodinates in the plane $(r,\varphi) \mapsto (r\cos\varphi, r\sin\varphi)$. Then a central field can be written in the form $\Phi(r)e_r$, where $e_r(x) := x/|x| = (\cos\varphi, \sin\varphi)$ is the *radial unit vector* at x.

EXERCISE 6. *Prove that each central field is conservative on the complement of the origin.*

Hence the equation of motion takes the form

$$m\ddot{x} = \Phi(r)e_r(x).$$

The vector product

$$M := mx\times\dot{x}$$

is called *angular momentum.*

PROPOSITION 4. *The angular momentum of a massive point in a central field remains constant.*

PROOF.

$$\frac{d}{dt}(x\times\dot{x}) = \dot{x}\times\dot{x} + x\times\ddot{x} = x\times\left(\frac{\Phi(r)}{m}\frac{x}{|x|}\right) = 0,$$

since $v\times v = 0$. □

EXERCISE 7. *Show that all possible orbits of massive points in a central field lie in a plane.*

We will learn later a general principle (Noether's Theorem) which states that each family of symmetries gives rise to a conservation law. The dimension of the preserved quantity is hereby equal to the dimension of the symmetry group.

EXERCISE 8. *Study for instance conservative axial symmetric fields. Prove that the potential of such is of the form $U = U(r, z)$ where (r, φ, z) are the cylindrical coordinates, the z–axis is the fixed under the isometries which leave the vector field invariant. The angular momentum M_z with respect to the z-axis is defined as*

$$M_z := e_z \cdot mx \times \dot{x}.$$

Show that this does not depend on the choice of the origin on the symmetry axis. How does it depend on the choice of a unit vector e_z which spans it? Prove that M_z is an invariant of motion in a system which is symmetric with respect to the z–axis.

1.2.3. *Kepler's Second Law.* The conservation of the angular momentum was first discovered by Johannes Kepler (1571–1630) through the observation of Mars. He formulated it as follows:
The area of the surface covered by the segments between 0 and $x(t)$ is the same for the any two equal time intervals.
If $S(t)$ is the surface covered between t_0 and t, then the statement reads

$$\frac{dS}{dt} = const.$$

Indeed, if Δt is small we find through geometric reasoning that

$$\Delta S := S(t + \Delta t) - S(t) = \frac{1}{2} r^2 \dot{\varphi} \Delta t + o(\Delta t)$$

with terms $o(\Delta t)$ which can be estimated by $const. |\Delta t|^2$. That is of size Δt. On the other hand we have with $e_\varphi = (-\sin\varphi, \cos\varphi)$

$$\dot{x} = \dot{r} e_r + r \dot{\varphi} e_\varphi$$

and hence

$$M = m(x \times \dot{x}) = m(x \times (\dot{r} e_r + r \dot{\varphi} e_\varphi)) = m r^2 \dot{\varphi} e_r \times e_\varphi.$$

In other words,

$$\frac{dS}{dt} = \frac{1}{2m} |M|.$$

1.3. The Orbits of the Planets. Using the conservation of the angular momentum we will describe all possible orbits of massive points in the gravitational field of a massive center like orbits of planets in a solar system or the moon and space ships in an orbit around the earth.
First, we reduce the system with three degrees of freedom to one with one degree.

PROPOSITION 5. *The distance r from a center of a central field changes like the height in a one-dimensional problem with potential*

$$V(r) = U(r) + \frac{|M|^2}{2mr^2}.$$

PROOF. Differentiating the identity $\dot{x} = \dot{r}e_r + r\dot{\varphi}e_\varphi$ and using $\dot{e}_r = \dot{\varphi}e_\varphi$ and $\dot{e}_\varphi = -\dot{\varphi}e_r$ we obtain

$$\ddot{x} = (\ddot{r} - r\dot{\varphi}^2)e_r + (2\dot{r}\dot{\varphi} + r\ddot{\varphi})e_\varphi.$$

Moreover,

$$\nabla U = \frac{\partial U}{\partial r}e_r$$

and hence

$$\begin{aligned} m(\ddot{r} - r\dot{\varphi}^2) &= -\frac{\partial U}{\partial r} \\ 2\dot{r}\dot{\varphi} + r\ddot{\varphi} &= 0. \end{aligned}$$

Now, conservation of angular momentum gives $\dot{\varphi} = M/mr^2$ and we obtain the one-dimensional problem

$$\begin{aligned} m\ddot{r} &= -\frac{\partial U}{\partial r} + mr\frac{|M|^2}{m^2r^4} \\ &= -\frac{\partial U}{\partial r} + \frac{|M|^2}{mr^3}. \end{aligned}$$

□

EXERCISE 9. *Show that the total energy of the new system* $\frac{m}{2}\dot{r}^2 + V(r)$ *is equal to the total energy of the original system.*

It follows that

$$\begin{aligned} \dot{r} &= \sqrt{2(E - V(r))/m} \\ r &= \int_0^t \frac{dr}{\sqrt{2(E - V(r))/m}} \\ \varphi(r) &= \int_{r_0}^r \frac{|M|}{r^2\sqrt{2m(E - V(r))}}dr. \end{aligned}$$

Notice that this is still true for *all* central fields. See also the interesting general discussion of the orbits in [?]. Amazingly, it turns out that only two kinds of central fields admit closed finite orbits: $U = ar^2, a \geq 0$ and $U = -k/r, k \geq 0$! The latter is the potential of a central gravitational field. Let us solve it in this case and deduce Kepler's observations.

1.3.1. *Kepler's First Law.* From the last formula above we obtain with the substitution $s = -|M|/r$

$$\begin{aligned}\int \frac{|M|}{r^2\sqrt{2m(E + k/r - |M|^2/2mr^2)}}dr &= \int \frac{1}{\sqrt{2mE - 2mks/|M| - s^2}}ds \\ &= \int \frac{1}{\sqrt{((2mE + m^2k^2/|M|^2 - (mk/|M| + s)^2}}ds.\end{aligned}$$

from which we deduce

$$\varphi = \arccos\Big(\frac{-\frac{|M|}{r} + \frac{mk}{|M|}}{\sqrt{2mE + m^2k^2/|M|^2}}\Big).$$

Hence with $p := |M|^2/mk$ and $e = \sqrt{1 + 2E|M|^2/mk^2}$ we have

$$r = \frac{p}{1 - e\cos\varphi}$$

which is the equation of an ellipse in polar coordinates centerd at one of the focal points for $e < 1$ which is the case for negative energy E. The quantity e is the excentricity while p determines its size. Thus we deduced Kepler's first law which states that the planets move along ellipses around the sun. The two half axes are given by

$$\begin{aligned}a &= \frac{p}{1 - e^2} \\ b &= a\sqrt{1 - e^2}\end{aligned}$$

1.3.2. *Kepler's Third Law.* The area of an ellipse is given by πab. Therefore, due to our proof of Kepler's second law, the time of circulation is equal to

$$T = \frac{2\pi abm}{|M|}.$$

Now a and b are determined by e and p which in turn are given by the constants of the system. We obtain

$$\begin{aligned}a &= -\frac{k}{2E} \\ b &= \frac{|M|}{\sqrt{-2mE}}\end{aligned}$$

Hence

$$
\begin{aligned}
T &= 2\pi \frac{mk}{-2E\sqrt{-2mE}} \\
&= \frac{2\pi a^{3/2}\sqrt{m}}{\sqrt{k}}. \qquad (1)
\end{aligned}
$$

We deduce Kepler's third law, that the square of the circulating time of any planet in the solar system is proportional to the cube of the size of the big half axis of its orbit.

1.4. The Phase Space. Let us wonder a little bit why we were so lucky to reduce the motion of a massive point with three degrees of freedom in a central field to a simple integration problem. For that it is convenient to canonically transform the system of second order differential equations governing the dynamics into a system of first order differential equations. Let $q \in \mathbb{R}^3$ denote the position vector while $p \in \mathbb{R}^3$ is the corresponding momentum. Then the equation of motion $m\ddot{x} = F(x)$ can be transformed into

$$
\begin{aligned}
m\dot{q} &= p \\
\dot{p} &= F(x).
\end{aligned}
$$

The new, larger set of coordinates $(p, q) \in \mathbb{R}^3 \times \mathbb{R}^3$ is called the *phase space* of the dynamical system. Its dimension is twice the degree of freedom of the mechanical system. Energy and angular momentum should then be understood as functions on the phase space. The conservation laws tell us that the motion $x(t)$ of the massive point can be canonically understood as a (parameterized) curve $(x(t), m\dot{x}(t))$ on level sets of these functions. A dimension count shows that the intersection of these sets should be two–dimensional.

We assume that the angular momentum, M, does not vanish. That reduces the system to one with two degrees of freedom. For a fixed position $q \neq 0$ the value of M defines a line of possible p's in the plane perpendicular to M which is parallel to the line defined by q. This defines a three–dimensional submanifold diffeomorphic to $(\mathbb{R}^2 \setminus \{0\}) \times \mathbb{R}$. The distance of these two lines is given by $|M|/r$ (remember $r = |q|$). On the other hand, $m\dot{r}$ is the length of the projection of $m\dot{x}$ onto the line generated by q. Hence,

$$
|p|^2 = m^2|\dot{r}|^2 + \frac{|M|^2}{r^2},
$$

and since the kinetic energy is given by $|p|^2/2m$ we obtain a one-dimensional problem

$$
\frac{m}{2}|\dot{r}|^2 + \frac{|M|^2}{2mr^2} + U(r) = E
$$

and the formula for the potential $V = V(r)$ as before. What is more important here is that we obtain

(1) two possible values for p when $V(r) < E$,
(2) one value if $V(r) = E$
(3) no solution if $V(r) > E$.

That means that in the (reduced) configuration space we see annuli or one possible complement of a disk around $0 \in \mathbb{R}^2$, or a punctured disk in which the point is moving. Above in the phase space sits the corresponding smooth surface without boundary which is either a torus or a cylinder in the latter two cases. What are the events these three cases correspond to?
We will see later that these systems are *integrable.* The angular momentum defines a three-dimensional foliation which satisfies a certain condition (Lagrangian) and is invariant under the flow on the phase space given by the equations of motion. Such systems can always be solved explicitely by the Arnold–Liouville Theorem, even though the calculation above shows that this is not an easy exercise of first-year calculus.

2. Systems with n massive points

To prove Kepler's Laws we studied the motion of a massive point in a central field. At the end of this section we will see that one can indeed reduce the system of two massive points with no other forces involved then their interaction, in particular no other massive points, to that.
We consider a system of n points with masses $m_j \quad j = 1, ..., n$ in $\mathbb{R}^3$. On each of them acts a force $F_j \in \mathbb{R}^3$. Then the motion of each of them satisfies the system of equations

$$m_j \ddot{x}_j = F_j.$$

One distiguishes between *inner*, F_j^i, and *outer forces* F_j^o. The outer force F_j^o is independent of the positions of the *other* points (e.g. the gravitational field of the sun for the earth–moon system) while the inner forces depend on them. They are typically (but not necessarily) sums of pairwise interactions $\sum_k F_{jk}^i$ between the jth and the kth massive point such that

$$F_{jk}^i = -F_{kj}^i$$

and the forces are parallel to the vector $x_j - x_k$. Examples are gravitation and electrostatic attraction.

2.1. Conservation Laws.

The *total momentum* of the system is given by

$$P = \sum m_j \dot{x}_j.$$

PROPOSITION 6. *The rate of change with respect to time of the total momentum is equal to the sum of outer forces.*

PROOF.

$$\dot{P} = \sum_j m_j \ddot{x}_j = \sum_j (F_j^o + \sum_k F_{jk}^i) = \sum_j F_j^o + \sum_{j,k} F_{jk}^i = \sum_j F_j^o.$$

□

Therfore, for a closed system ($F^o = 0$) the total momentum is conserved.
The *center of mass* of the system is given by

$$x := \frac{\sum_j m_j x_j}{m}$$

where $m := \sum_j m_j$ is its *total mass.* This is an affine quantity, i.e. it is independent of a choice of origin. It follows that the total momentum is exactly the momentum of a point at the center with mass m. It follows

COROLLARY 7. *The motion of the center of mass is the motion of a point whose mass is the total mass of the system with the sum of forces of the system acting on it. In particular, the center of mass in a closed system is inertial.*

Similarily, one defines the *total angular momentum* with respect to the reference point 0 to be

$$M = \sum_j x_j \times m_j \dot{x}_j.$$

PROPOSITION 8. *The total angular momentum of a closed system is conserved.*

PROOF. We compute

$$\dot{M} = \sum_j \dot{x}_j \times m_j \dot{x}_j + \sum_j x_j \times m_j \ddot{x}_j.$$

Now

$$\sum_j x_j \times m_j \ddot{x}_j = \sum_j x_j \times F_j = \sum_j x_j \times \left(\sum_k F_{jk}^i + F_j^o\right) = \sum_j x_j \times F_j^i.$$

$x_j \times F_j^i$ is called the *torque* of the force F_j^i with respect to 0. The sum of the torques of the two interactions $F_{jk} = -F_{kj}$ vanishes because $x_j \times F_{jk} + x_k \times F_{kj} = (x_j - x_k) \times F_{jk} = 0$ since the two vectors are parallel. □

COROLLARY 9. *From the proof follows that the rate of change of the total angular momentum is equal to the sum of torques of the outer forces.*

We assume that the interactions are homogenous in space, i.e. depend on the distance of the two points involved only. With respect to each of the two points that means that each of them is a central field and hence conservative. In particular,

$$F^i_{jk}(x_j, x_k) = -\frac{\partial U^i_{jk}(x_j - x_k)}{\partial x_j}.$$

Assume moreover that the outer forces have a potential

$$F^o_j(x_j) = -\frac{\partial U^o_j(x_j)}{\partial x_j}.$$

PROPOSITION 10. *The total energy of the system (sum of all kinetic energies and potential energies) is conserved.*

PROOF.

$$\begin{aligned}\dot{E} &= \sum_j m_j \dot{x}_j \cdot \ddot{x}_j + \sum_j \dot{x}_j \cdot \Big(\sum_k \frac{\partial U^i_{jk}(x_j - x_k)}{\partial x_j} + \frac{\partial U^o_j(x_j)}{\partial x_j}\Big) \\ &= \sum_j \dot{x}_j \cdot \big(m\ddot{x}_j - F_j\big) = 0.\end{aligned}$$

□

2.2. The Two-Body-Problem. Let us reexamine Kepler's problem. We put the origin in the center of the sun assuming it to be an inertial point. This is, of course, not correct since the gravitational force of the planet acts also on the much heavier sun. The equations of motions are

$$m_i \ddot{x}_j = -\frac{\partial U}{\partial x_j}$$

for $j = 1, 2$ and the potential for gravitation $U = U(|x_1 - x_2|)$.

PROPOSITION 11. *The dynamics of the difference $x = x_1 - x_2$ is the same as the dynamics of a point with mass*

$$m = \frac{m_1 m_2}{m_1 + m_2}$$

in the central field with potential $U(|x|)$.

PROOF. Multiply each of the two equations with the mass m_k of the other point and substract them from each other. □

On the other hand, the center of mass is an inertial point and could therefore be chosen to be the origin of an inertial coordinate system. Then the coordinate of x_1 in it is

$$\frac{m_2}{m_1 + m_2}(x_1 - x_2),$$

and similarily for x_2. Hence if $m_2 >> m_1$, x_1 moves like a point in a central field at 0 with mass m_1. However, the much heavier object will also orbit around the center of mass.

EXERCISE 10. *Determine the size of the big half-axis of the ellipse around the center of mass of the earth-moon system on which the center of the earth lies. Does the center lie inside or outside of the earth?*

2.3. The Three-Body-Problem. The phase space of the system with three massive points is 18–dimensional. If we choose the origin to coincide with the center of mass this will reduce it to 12 dimensions. However, the conservation law for the angular momentum with respect to the center of mass gives only three more dimensional reductions. We end up with 9–dimensional level-sets on which the orbits in the phase space have to stay. In order to be Lagrangian they had to be half–dimensional. It turns out that these are all conserved quantities besides the energy. The problem of three or more points with masses leads to non-integrable systems. On may attempt to describe them as perturbation of integrable systems of two bodies, where one introduces the force of the third far away massive point as a periodic time-dependent force which is small compared to the interaction of the two points under study.

Part 2

Variational Principles in Lagrangian Mechanics

We describe the equations of motion as critical points of a functional on the space of paths with fixed end points. The notions and laws of Lagrangian mechanics are invariant under an action of the group of diffeomorphisms of the configuration space and time. Each family of diffeomorphisms which leaves the functional invariant leads to a conservation law.

CHAPTER 2

Calculus of Variation

We introduce the general setting of functionals on Banach spaces and define their critical points. We study several examples.

1. Calculus on Banach Spaces

1.1. Derivatives. Recall that a continous real function $f : B \to \mathbb{R}$, often referred to as functional, on a Banach space B is differentiable in $x \in B$ if there is a linear functional $f'(x) : B \to \mathbb{R}$ such that

$$\lim_{\|h\| \to 0} \frac{f(x+h) - f(x) - f'(x)(h)}{\|h\|} = 0.$$

EXERCISE 11. *Prove that $f'(x)$ is uniquely determined and continous as linear functional.*

Let $B = \mathbf{C}^2([t_0, t_1]; \mathbb{R}^n)$ be the space of paths with two continous derivatives. Let $L = L(a, b, c)$ be a smooth function in three variables, often referred to as the *Lagrangian*. Then we define for $x \in B$ through

$$\mathcal{L}(x) := \int_{t_0}^{t_1} L(x, \dot{x}, t) dt$$

a differentiable functional on B. Indeed,

$$\begin{aligned} \mathcal{L}(x+h) - \mathcal{L}(x) &= \int_{t_0}^{t_1} [L(x+h, \dot{x} + \dot{h}, t) - L(x, \dot{x}, t)] dt \\ &= \int_{t_0}^{t_1} \Big[\frac{\partial L}{\partial x} h + \frac{\partial L}{\partial \dot{x}} \dot{h} + R(x, \dot{x}, h, t)\Big] dt \end{aligned}$$

where $\lim_{h \to 0} R/\|h\| = 0$. Hence, via partial integration we obtain

$$\begin{aligned} \mathcal{L}'(x)(h) &= \int_{t_0}^{t_1} \Big[\frac{\partial L}{\partial x} h + \frac{\partial L}{\partial \dot{x}} \dot{h}\Big] dt \\ &= \int_{t_0}^{t_1} \Big[\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}\Big] h dt + \Big(\frac{\partial L}{\partial \dot{x}} h\Big)\Big|_{t_0}^{t^1}. \end{aligned}$$

EXERCISE 12. *Consider the set of real functions, $B = C^2(B^n;\mathbb{R})$, on the unit ball in $\mathbb{R}^n$ with two derivatives. We are given a function $L : \mathbb{R} \times \mathbb{R}^n \times B^n \to \mathbb{R}$ and define the functional $\mathcal{L} : B \to \mathbb{R}$*

$$\mathcal{L}(x) := \int_{B^n} L(f, \nabla f, x) d^n x.$$

Show that $\mathcal{L}$ is differentiable and compute its derivative.

EXERCISE 13. Length of a curve. *Let $B = C^2([t_0,t_1];\mathbb{R}^n)$. Let $\Big(g_{ij}(x)\Big)_{ij}$, $x \in \mathbb{R}^n$ be a family of symmetric, positive definite $n \times n$–matrices, i.e. a Riemannian metric or structure. It defines a family of euclidean products via $g_x(v,w) = \sum_{ij} g_{ij}(x) v_i w_j$. They are usually considered to be inner products on the tangent space (the velocity vectors) at a point $x \in \mathbb{R}^n$, the location. Then the norm of such a velocity v is given by $\|v\|^2 = g_x(v,v)$. The length of a curve γ parametrized as a $\gamma \in B$ is defined to be*

$$l(\gamma) := \int_{t_0}^{t_1} \|\dot\gamma(t)\| dt.$$

Show that this definition is independent of the parametrization of γ.
Compute its derivative on the euclidean plane, i.e. $n = 2$ and $g_{ij} = \delta_{ij}$.

1.2. Critical Points. A *critical (or extremal) point $x \in B$* of the restriction of the differentiable functional $\Phi : B \to \mathbb{R}$ to an affine subspace $x_0 + B_0$, $x_0 \in B, B_0 \subset B$ is a point for which $\Phi'(x)|_{B_0} = 0$, i.e. $\Phi'(x)(h) = 0$ for every $h \in B_0$.

PROPOSITION 12. *A Kurve $x = x(t)$ is extremal for $\mathcal{L}(x) := \int_{t_0}^{t_1} L(x,\dot x,t)dt$ on the curves in $\mathbb{R}^n$ starting at $x(t_0) = x_0$ and ending at $x(t_1) = x_1$ iff*

$$f(t) = \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot x}\Big) - \frac{\partial L}{\partial x} = 0. \tag{2}$$

PROOF. Variations of x are characterized by $h(t_0) = h(t_1) = 0$. Hence, from the above calculations we obtain

$$\mathcal{L}'(x)(h) = \int_{t_0}^{t_1} \Big[\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot x}\Big) - \frac{\partial L}{\partial x}\Big] h dt,$$

and the sufficiency follows immediately. On the contrary, assume that some coordinate $f_i(t') = 2c \neq 0$ does not vanish at $t' \in (t_0,t_1)$. Since $f \in C^0([t_0,t_1];\mathbb{R}^n)$ $|f_i(t)| > c$ on some neighborhood $t \in [t' - 3\delta, t' + 3\delta] \subset [t_0,t_1]$ with the same sign for f_i. Choose some function $h \in C^2([t_0,t_1],\mathbb{R}^n)$ which vanishes for $|t - t'| > 2\delta$, $h_i = 1$ on $[t'-\delta, t'+\delta]$ and $h_j = 0$ for all $j \neq i$. Then

$$|\int_{t_0}^{t_1} f(t)\cdot h(t) dt| > 2\delta c$$

and x was not a critical point. $\square$

DEFINITION 1. *The equation(s) (2) are called* Euler–Lagrange equations.

EXERCISE 14. *Show that the extremality condition does not depend on the choice of coordinates on* $\mathbb{R}^n$. *Which quantities in the definition of the Lagrangian are transformed? Show that the extremal points for the length functional in the euclidean plane are lines. Compute the equation for a line in polar coordinates.*
(*) How does a Lagrangian, defined for curves, change under time transformations, i.e. diffeomorphisms acting on the parametrization of the paths?

2. Euler–Lagrange Equations

We introduce the functional whose critical points are described by the equations of motion in a conservative field.

2.1. Hamilton's Principle of Least Action. Let $U = U(x)$ be the potential of a conservative field.

PROPOSITION 13. *The motions of the conservative mechanical system with potential* $U = U(x)$ *are the same as the critical values of the functional* $\Phi(x) = \int_{t_0}^{t_1} L(x, \dot{x})dt$ *on differentiable paths with fixed end points, where the Lagrangian is given by the difference of kinetic and potential energy*

$$L(x, \dot{x}) := \sum_j \frac{m_j}{2}|\dot{x}_j|^2 - U(x).$$

PROOF. We compute

$$\frac{\partial L}{\partial \dot{x}_j} = m_j \dot{x}_j$$
$$\frac{\partial L}{\partial x_j} = -\frac{\partial U}{\partial x_j}.$$

Hence

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{x}_j}\Big) - \frac{\partial L}{\partial x_j} = m_j \ddot{x}_j + \frac{\partial U}{\partial x_j} = 0.$$

□

COROLLARY 14. *Let* $q_1, ..., q_{3n}$ *be arbitrary coordinates of the configuration space of* n *massive points. Then the motion in theses coordinates is described by the Euler–Lagrange equations.*

The Euler–Lagrange equations of a mechanical system are sometimes also called *Lagrange's equations*, q_i the *generalized coordinates*, $\dot{q}_i$ the *generalized velocities*, $\partial L/\partial \dot{q}_i$ the *canonical conjugated moments*, $\partial L/\partial q_i$ the *generalized forces*, $\int L(q, \dot{q}, t)dt$ the *action*.

EXERCISE 15. *Determine the elements of a Lagrangian system for a free massive point in $\mathbb{R}^3$, i.e. in a system without forces. What are the critical points of the functional? Compare that to the length functional.*
(*) Is that true for a general Riemannian structure on $\mathbb{R}^3$? Why?

EXERCISE 16. *Determine the elements of a Lagrangian system for a massive point in a central field in the plane in polar coordinates. What is the role of the two coordinates in the Lagrangian?*

In both examples we have coordinates on which the Lagrangian does not depend:

$$\frac{\partial L}{\partial q_j} = 0.$$

Such coordinates are called *cyclic*. A simple consequence of the equations is

PROPOSITION 15. *The canonical conjugated moment of a cyclic coordinated is conserved.*

PROOF.

$$\frac{dp_j}{dt} = \frac{\partial L}{\partial q_j} = 0.$$

□

3. The Legendre Transform

We describe a transformation which will serve to reformulate the second order system of Lagrangian equations into a first order system.

3.1. The Construction. Let $y = f(x)$ be a differentiable strongly convex function. Given a real number p we define a new function $g = g(p)$ as follows: We denote by $x = x(p)$ be the point on the graph whith the biggest y–distance from the line $y = px$. It can be determined as the maximum of the function $F(p, x) := px - f(x)$. We let $g(p) := F(p, x(p))$. Since f is strongly convex $x(p)$ is uniquely determined by $0 = \partial F(p, x)/\partial x = -f'(x) + p$.

EXAMPLE 2. *Let $f(x) = x^2$. Then $F(p, x) = px - x^2$, $x(p) = p/2$ and $g(p) = p^2/4$. Hence, if $f(x) = mx^2/2$ then $g(p) = p^2/2m$.*

EXERCISE 17. *Compute the Legendre transform of $f(x) = x^\alpha/\alpha$ for $\alpha > 1$.*

EXERCISE 18. *What is the Legendre transform of $f(x) + c$ if the Legendre transform of f is g.*

Let $f(x)$ be piecewise linear with one convex corner at x. The domain of its transform is the set of slopes p for which $x = x(p)$ is uniquely determined. The y-distance from the corner point to the line is obviously a linar function of the slope, hence the transform is the restriction of that function to the domain.
If f is two times differentiable with $f'' > 0$ then its Legendre transform g is two times differentiable and $g'' > 0$.

PROPOSITION 16. *The Legendre transform is involutiv, i.e. f is the Legendre transform of g.*

PROOF. Let $G(x,p) = xp - g(p)$ be the function to determine $p = p(x)$. For fixed p this is exactly the equation for the tangent to the graph of f with slope p. For x fixed we are looking for the maximum of that function for varying p. Geometrically this is y–coordinate of the intersection of the tangent with slope p with the vertical line through $(x,0)$. Since f is strongly convex it is clear that the maximum is attained for the tangent through $(x, f(x))$ where the value is obviously $f(x)$. $\square$

3.2. Young's Inequality. The definition of Legendre transform gives $F(x,p) = px - f(x) \leq g(p)$. Hence for f and g which are the Legendre transform of each other we deduce Young's inequality

$$px \leq f(x) + g(p).$$

The somewhat obvious statement is beter illustrated through examples.

EXAMPLE 3. *For $\alpha > 1$ the Legendre transform of $f(x) = \frac{x^\alpha}{\alpha}$ is $g(p) = \frac{p^\beta}{\beta}$ with $1/\alpha + 1/\beta = 1$. Hence*

$$px \leq \frac{x^\alpha}{\alpha} + \frac{p^\beta}{\beta}.$$

EXERCISE 19. *Derive Young's inequality for $f(x) = e^x$.*

3.3. Higher-dimensional Legendre transform. We suppose that $f(x), \quad x \in \mathbb{R}^n$ is a strongly convex smooth function, i.e. the Hessian of f, $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{ij}$, is positive definite. Then, as before, $x = x(p)$ is the extremal point of the function $F(x,p) = p(x) - f(x)$ for a fixed $p \in (\mathbb{R}^n)^*$, i.e. the point x which satisfies

$$\frac{\partial f}{\partial x} = p.$$

Then $g(p) := F(x(p), p)$ is a function on the dual space.

EXERCISE 20. *Show that these conditions define g uniquely and completely, provided that the differential df_x attains all values in $(\mathbb{R}^n)^*$ as x runs through $\mathbb{R}^n$. Show that all other statements and corollaries remain true.*

EXERCISE 21. *Compute the Legendre transform g of a positive definite quadratic form $f(x) = \sum g_{jk}x_jx_k$. Show that for the dual variable p it satisfies*

$$g(p) = f(x).$$

4. Hamilton's Equations

We describe the very special system of first order differential equations we obtain from the system of Lagrange equations after applying the Legendre transform, the *Hamilton equations.*

PROPOSITION 17. *The system of Lagrange equations for a Lagrange function $L = L(q, \dot{q}, t)$, $q \in \mathbb{R}^n$ which is convex in the $\dot{q}$-variable is equivalent to the system of $2n$ first order differential equations*

$$\dot{p} = -\frac{\partial H}{\partial q}$$
$$\dot{q} = \frac{\partial H}{\partial p}.$$

The function $H = H(p, q, t) = p\dot{q} - L(q, \dot{q}, t)$ is the Legendre transform of L as a function on $\dot{q}$. It is called the Hamilton function.

PROOF. Remember that the variable p was given as $p = \partial L/\partial \dot{q}$. Differentiating H we get

$$dH = \frac{\partial H}{\partial p}dp + \frac{\partial H}{\partial q}dq + \frac{\partial H}{\partial t}dt.$$

Using the identity $H = p\dot{q} - L$ with the identity for p we have

$$\begin{aligned} dH &= \dot{q}dp + pd\dot{q} - \frac{\partial L}{\partial q}dq - \frac{\partial L}{\partial \dot{q}}d\dot{q} - \frac{\partial L}{\partial t} \\ &= \dot{q}dp - \frac{\partial L}{\partial q}dq - \frac{\partial L}{\partial t}. \end{aligned}$$

(3)

Hence we obtain the desired system of differential equations together with the additional

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t},$$

which is true since H is the Legendre transform of L as a function in $\dot{q}$.

EXERCISE 22. *Show the inverse statement: Given a Hamiltonian function $H = H(p, q, t)$ which is strongly convex in p then the Hamilton equations transform into the Lagrange equations for L being the Legendre transform $L(q, \dot{q}, t) = p\dot{q} - H(p, q, t)$, where $\dot{q} = \partial H/\partial p$.*

$\square$

4.1. Conservation of Energy. We may formulate now a more general conservation law for Lagrangian systems.

PROPOSITION 18. *A solution of the Lagrange equations $q = q(t)$ satisfies*

$$\frac{dH(p(t), q(t), t)}{dt} = \left(\frac{\partial H}{\partial t}\right)(p(t), q(t), t),$$

where $p(t) = (\partial L/\partial \dot{q})(q(t), \dot{q}(t), t)$. H is also called the energy *of the Lagrangian system. In particular, if L (or H) are time-independent the energy of the Lagrangian system is conserved.*

PROOF.

$$\begin{aligned}\frac{dH(p(t), q(t), t)}{dt} &= \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial t} \\ &= \frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial q}\right) + \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} + \frac{\partial H}{\partial t}\end{aligned}$$

from which the statment follows. □

EXAMPLE 4. *Let $L = T - U$, where $U = U(q)$ is a potential and T is the kinetic term like*

$$T = \frac{1}{2}\sum_{jk} g_{jk}\dot{q}_j\dot{q}_k.$$

Then $H = E$, the total energy of the mechanical system. Indeed, the Legendre transform of T, T' satisfies $T'(p) = T(\dot{q})$. Hence, the Legendre transform of $L = T - U$ is equal to $H(p, q) = T(\dot{q}) + U(q)$.

EXAMPLE 5. *Let us consider the system with one degree of freedom. It satisfies*

$$m\ddot{q} = -\frac{\partial U}{\partial q}.$$

The kinetic energy is $T = m\dot{q}^2/2$ the potential energy $U = U(q)$. Hence the Hamiltonian is given by $H(p, q) = p^2/2m + U(q)$ and the Hamilton equations are

$$\dot{q} = p/m \tag{4}$$

$$\dot{p} = -\frac{\partial H}{\partial q} = -\frac{\partial U(q)}{\partial q}. \tag{5}$$

When we compare this to Newton's equation this is a good way to remember where to put the negative sign in Hamilton's equations. Without integrating the system we know the curves in the phase space parameterized by solutions are the level sets of $H = H(p, q)$ from conservation of energy.

EXERCISE 23. *Sketch the graph of a differentiable potential $U = U(q)$ with at least three critical points. Indicate the structure of the dynamics of the corresponding Hamiltonian system in the phase space.*

4.2. Cyclic Coordinates. Recall the definition of a *cyclic coordinate* q_1 of a Lagrangian system with Lagrangian $L = L(q, \dot{q}, t)$: L does not depend on q_1. This condition is obviously equivalent to the vanishing of the partial derivative $\partial H/\partial q_1 = 0$ for the corresponding Hamiltonian $H = H(p, q, t)$.

PROPOSITION 19. *If q_1 is a cyclic coordinate then the corresponding momentum p_1 is a first integral, i.e. a preserved quantity, of the Hamiltonian system. The dynamics of the remaining coordinates is as in a system with one less degree of freedom and a Hamiltonian*

$$H_c(p_2, ..., p_n; q_2, ..., q_n; t) := H(c, p_2, ..., p_n; q_1, q_2, ..., q_n; t).$$

PROOF. With $p' := (p_2, ..., p_n)$ and $q' := (q_2, ..., q_n)$ we obtain

$$\dot{q}' = \frac{\partial H}{\partial p'}, \quad \dot{q}_1 = \frac{\partial H}{\partial p_1}$$
$$\dot{p}' = \frac{\partial H}{\partial q'}, \quad \dot{p}_1 = 0$$

The last equation shows that $p_1 = const.$. Hence it enters the equations as a parameter only. Therefore we may solve it completely for (p', q'). Last we obtain

$$\dot{q}_1 = \left(\frac{\partial H}{\partial p_1}\right)(c, p'(t), q'(t), t).$$

□

Most explicit solutions of problems in mechanics are obtained that way.

COROLLARY 20. *Each system with two degrees of freedom which has a cyclic coordinate is integrable.*

5. Liouville's Theorem

Let $D \subset \mathbb{R}^n \times \mathbb{R}^n$ be a measurable set of the phase space. Suppose that Hamilton's equations for a given Hamiltonian $H = H(p, q)$ have a solution for each time in the interval $[t_0, t_1]$ for almost every initial value in this set, $(p(t_0), q(t_0)) \in D$. Then we define a map $g^t : D \to \mathbb{R}^n \times \mathbb{R}^n$ asking that $g^t(p, q)$ be the t–value of the solution of Hamilton's equations with initial value (p, q). This will be defined almost everywhere on D.

THEOREM 21. *The transformation of the phase space given by g^t is volume preserving, i.e. on subsets D characterized above we have*

$$(g^t)^*\mu|_D = \mu|_D$$

for the Lebesgue measure μ on $\mathbb{R}^{2n}$. In particular, for any such subset D we have

$$\mu(D) = \mu(g^t(D)).$$

PROOF. The transformation of the Lebesgue measure under a diffeomorphism $g : \mathbb{R}^m \to \mathbb{R}^m$ is governed by the Jacobian of that map:

$$g^*\mu = |\det\Big(\frac{\partial g_k}{\partial x_j}\Big)|\mu.$$

Now since $\frac{d}{dt}(\det(E + A(t))) = Tr\dot{A}$ for a family $A(t)$ with $A(0) = 0$, and $g^0 = Id$ we obtain

$$\frac{d}{dt}\Big(\det\big(\frac{\partial g_k}{\partial x_j}\big)\Big) = \frac{\partial}{\partial x_k}(\dot{g_k}) = \operatorname{div}(\dot{g}).$$

The right hand side of Hamilton's equations defines the *Hamiltonian vector field* on the phase space $X_H := (-\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p}) \in \mathbb{R}^n \times \mathbb{R}^n$. Clearly, we have

$$\operatorname{div}(X_H) = \sum_k \Big[\frac{\partial}{\partial p_k}\Big(-\frac{\partial H}{\partial q_k}\Big) + \frac{\partial}{\partial q_k}\Big(\frac{\partial H}{\partial p_k}\Big)\Big] = 0.$$

Therefore, $\operatorname{div}(\dot{g}) = 0$ and the measure is preserved. □

5.1. Poincare's Recurrence Theorem. There are several immediate consequences of the theorem. For instance, in a Hamiltonian system there are no sinks and sources, neither are there limit cycles. A more striking consequence is the following

THEOREM 22. *Let g be a volume preserving, continous function which maps an open set D of finite volume to itself. Then , given any point in D for each neighborhood $U \subset D$ there is a point $x \in U$, which returns to U, i.e. $g^n(x) \in U$ for some $n > 0$.*

PROOF. Consider all of the sets

$$U, g(U), g^2(U), \ldots$$

Since they have all the same volume two of them $g^k(U), g^l(U), \quad k < l$, say, must intersect. Indeed among $\{1, \ldots, m\}$ with $m > \operatorname{vol}(D)/\operatorname{vol}(U)$ we will find such a pair. Pick a point $x' \in g^k(U) \cap g^l(U)$, then $g^{-k}(x') \in U \cap g^{l-k}(U)$. □

REMARK 23. *This is a very general statement, which therefore must be very weak. Indeed, volume preserving maps can be very different in nature, i.e. either integrable and predictable or chaotic. Neither can it be used to produce periodic points (or closed orbits) since, in general, the period, n, could become arbitrarily large as we shrink U. Even for volume preserving flows on nice spaces, like the three-sphere, there are, in general, no closed orbits (see* [?]*). On the other hand it produces a paradox which can never be truely observed in nature. For instance, all particles of a gas which are in a bottle at one time will return there at some point. This is, of course, an event which needs a very large time to have some observable probability, certainly more time than the age of our solar system.*

5.2. Ergodic Systems. Let us discuss the recurrence theorem in some examples. Let D be a circle and and g be the rotation by $\alpha/\pi \notin \mathbb{Q}$. Such a transformation has only one periodic point. However, for any point $x \in \mathbb{S}^1$ the set $\{g^n(x)\}$ is everywhere dense in $\mathbb{S}^1$.
Consider the 2-torus $\mathbb{S}^1 \times \mathbb{S}^1$ and a flow $g^t(\phi_1, \phi_2) := (\phi_1 + t\alpha_1, \phi_2 + t\alpha_2)$ for some fixed pair of angles (α_1, α_2). Then if $\alpha_1/\alpha_2 \notin \mathbb{Q}$, the trajectory of $\{g^t(x)\}_{t\in\mathbb{R}}$ is everywhere dense in the torus.
For both examples holds a much stronger

THEOREM 24. *(1) For any measurable set $U \subset \mathbb{S}^1$ and any x*

$$\lim_{n\to\infty} \frac{\mu(\{k \leq n | g^k(x) \in U\})}{n} = \mu(U).$$

(2) For any measurable set $U \subset T^2$ and any x

$$\lim_{T\to\infty} \frac{\mu(\{0 \leq t \leq T | g^t(x) \in U\})}{T} = \mu(U).$$

We do not prove this theorem here. Orbits with such a property are called *ergodic*. A system is called *ergodic* if for all but a zero set of x the orbit through x is ergodic. Ergodic systems play a crucial role in statistical mechanics.

EXERCISE 24. *Consider the series $1, 2, 4, 8, 1, 3, 6, 1, 2, 5, 1, 2, 4, ...$ of first digits in the series 2^n. Does 7 occur at all? Which digit occurs more often, 7 or 8?*

CHAPTER 3

Lagrange–Mechanics on Manifolds

In an attempt to describe the different forces of nature (gravitation, electromagnetism, nuclear forces and the forces that bind quarks to constitute the elementary particles) as particular expression of one unified force theoretical physisists have abandoned Galilei's postulate, that the world could be described by a 4–dimensional euclidean space long time ago.

But even if we do assume that the world looks globally as it appears to our eyes even guarded by large telescopes it turns out that in the study of a system of massive particles which *a priori* lie on a certain subset of the configuration space, a *manifold*, it is convenient to reformulate the principles of mechanics to such a setting.

EXAMPLE 6. *A weight suspended from a fixed point on a rigid rod will remain in a fixed distance from that point equal to the length of the rod. Its theoretical full configuration space* $\mathbb{R}^3$ *is reduced to the sphere of that radius.*

EXAMPLE 7. *Assume that three massive points* A, B, C *are connected through rods of fixed length. The configuration space* $\mathbb{R}^{3\times 3}$ *is reduced as follows. The firts point* A *still has all three original degrees of freedom. Since* B *must stay in equal distance to* A *it will be completely described by points on a sphere of that radius once the position of* A *is given. Last, the point* C *has a fixed distance from both,* A *and* B*. Hence, it will stay on a circle of such points if the positions of* A *and* B *are fixed. However, the configuration space is* not *simply the product of these three. In order to describe it we observe that the vectors* $u := \vec{AB}/|\vec{AB}|$, $v := v'/|v'|$ *with* $v' = \vec{AC} - (\vec{AC}\cdot u)u$, *and* $w := u \times v$ *form an orthogonal oriented frame of* $\mathbb{R}^3$. *Indeed they define the columns of an orthogonal matrix* A *with* $A^{-1} = A^T$ *and* $\det(A) = 1$. *Such matrices form the group of* special orthogonal matrices $\mathbf{SO(3)}$. *The configuration space is therefore equal to* $\mathbb{R}^3 \times SO(3)$.

EXERCISE 25. *Show that* $SO(3)$ *is* not *homeomorphic (continously bijective) to* $\mathbb{S}^2 \times \mathbb{S}^1$. *Hint: Show that* $SO(3)$ *is the same as the real projective space* $\mathbb{RP}^3$. *Compare loops in each of the two spaces in question.*

1. Manifolds

A topological space is called a *Hausdorff* if any two points in it admit disjoint neighborhoods. It is called *locally compact* if any point has a neighborhood whose topology is generated by a countable basis of open sets with compact closure. A *differentiable manifold* M is a locally compact Hausdorff space which "looks locally like a euclidean space" everywhere. That means that for each point $x \in M$ there is an open set $U \subset \mathbb{R}^n$ and a homeomorphism (bijective continous map with continous inverse) $\varphi : U \to \varphi(U) \subset M$ mapping onto a neighborhod of x. The family of such *coordinate systems* satisfies the following condition: If (U, φ) and (V, ψ) are two such coordinates, then the transition from one into the other on the intersection

$$\varphi^{-1} \circ \psi : \psi^{-1}(\varphi(U) \cap \psi(V)) \longrightarrow \varphi^{-1}(\varphi(U) \cap \psi(V))$$

is a differentiable map – in the usually sense, meaning that all n functions are differentiable as funtions on $\mathbb{R}^n$. Note that since this is, of course, also is required for $\psi^{-1} \circ \varphi$, that means that the transition maps are diffeomorphisms. A family of such coordinate systems is called an *atlas* of M. Sometimes the convention is to require that the atlas contains *all* such coordinates, which is a rather uncomfortably huge set. One might also change the condition on the transitions and ask for: continous, k times differerentiable, real or complex analytic, or preserving a geometric structure such as spherical, euclidean or hyperbolic metric.

Exercise 26. *Show that the product of two manifolds $M \times N$ carries naturally the structure of a manifold.*

A submanifold of a manifold M is a subset $N \subset M$ whose induced topology is that of a manifold.

Example 8. *Consider $\{(\exp(it\alpha), \exp(it\beta)) \mid t \in \mathbb{R}\} \subset \mathbb{S}^1 \times \mathbb{S}^1$. This is a manifold if and only if the ratio $\alpha/\beta \in \mathbb{Q}$. Describe what prevents this to be a submanifold if this ratio is irrational?*

Quite often (abstract) manifolds appear as zero sets of sets of functions on a linear space. For instance the sphere of radius r of any dimension can be represented as the zero set of

$$f(x) = |x|^2 - r^2$$

for $x \in \mathbb{R}^{n+1}$. Indeed, we have the following

Proposition 25. *Let $f = (f_1, ..., f_m)$ consist of real functions on $\mathbb{R}^m$ such that the Jacobian (or differential) $df_x : \mathbb{R}^n \to \mathbb{R}^m$ is linear surjective for any x. Then the preimage $f^{-1}(b) \subset \mathbb{R}^n$ for any $b \in \mathbb{R}^m$ is a submanifold of dimension $n - m$.*

EXERCISE 27. *Represent $SO(3)$ as the zero set of a set of functions on a euclidean space and show that these satisfy the condition of the theorem.*

1.1. Functions and maps. A map $f : M \to N$ between two smooth manifolds is differentiable (continous) if for any pair $(x, f(x)$ and any chart (U, φ) around x and (V, ψ) around $f(x)$ the map represented in charts

$$\psi^{-1} \circ f \circ \varphi : U \longrightarrow V$$

is differentiable (smooth). This definition does not depend on the choice of coordinates.

EXERCISE 28. Lie groups. *Show that the maps*

$$(A, B) \in SO(3) \times SO(3) \mapsto (AB) \in SO(3)$$
$$A \in SO(3) \mapsto A^{-1} \in SO(3)$$

are differentiable.

1.2. Tangent and cotangent bundles. If $M^n \subset \mathbb{R}^N$, $n \leq N$ is a smooth submanifold there is a intuitive understanding what a tangent vector at x should be: Let $f = (f_1, ..., f_{N-n})$, be a regular function, such that $f^{-1}(0)$ is a neighborhood of x in M, then a tangent vector $v \in T_xM$ is characterized by the condition

$$df_x(v) = 0.$$

In other words $T_xM = \mathrm{Ker}(df_x)$ which is clearly a vector space inheriting the operations from $\mathbb{R}^n$. It is also not too hard to show that this could be taken a s a definition: it does not depend on f. On the other hand, it is not so obvious that T_xM could be intrinsicly defininded, i.e. independently from its realization as a submanifold. Indeed using our aquaintance with variational principles we should think of a tangent vector at x as a first variation of the point inside M. That means that a tangent vector v is given by a path $\gamma : [0, \epsilon)$, $\gamma(0) = x$ with the idea that "$v = \frac{d}{dt}|_{t=0}\gamma$". We should also think of tangent vectors as velocities. We say that two such paths γ, δ are equivalent if $\frac{\varphi^{-1}(\gamma(t))}{dt}|_{t=0}$ and $\frac{\varphi^{-1}(\delta(t))}{dt}|_{t=0}$ coincide in a chart (U, φ) around x. Suppose that $\varphi^{-1}(x) = 0 \in U \subset \mathbb{R}^n$. Hence, given two tangent vectors $v, w \in T_xM$ and the corresponding paths γ, δ starting at x we may define new paths $\alpha(t) := \varphi(\varphi^{-1}(\gamma(t)) + \varphi^{-1}(\delta(t)))$ and $\beta(t) := \varphi(\lambda\varphi^{-1}(\gamma(t)))$ which will be defined on $[0, \epsilon]$ for ϵ suffficiently small. Given a differentiable map $f : M \to N$ we immediately obtain a first variation, its *differential* $f_* : T_xM \to T_{f(x)}N$. Indeed if γ represents a tangent vector at x then $f \circ \gamma$ will represent a tangent vector at $f(x)$.

EXERCISE 29. *Show that the definition of T_xM and the operations do not depend on the coordinate chosen. Show that this defines a vector space structure on T_xM. Show that f_* is a linear map between vector spaces.*

The union of tangent spaces $\cup_x T_x M$ is called the *tangent bundle* of the manifold M. It carries a natural structure of a smooth manifold as follows. Let $\varphi : U \subset \mathbb{R}^n \to M$ be a smooth chart in M. Then its differential defines a chart in TM via

$$\varphi_* : TU \cong U \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n \to TM.$$

In fact, given two such charts (U, φ) and (V, ψ) we obtain the transition over $W := \psi^{-1}(\varphi(U) \cap \psi(V))$,

$$\varphi_*^{-1} \circ \psi_* : TW \subset \mathbb{R}^n \times \mathbb{R}^n \to T(\varphi^{-1}(\psi(W))) \subset \mathbb{R}^n \times \mathbb{R}^n.$$

Note that this function is linear in the second component of the product. The inverse φ_*^{-1} is a pointwise inverse of the linear maps $\varphi_{*x} : T_x U \cong \mathbb{R}^n \to T_{\varphi(x)} M$. Moreover,

$$\varphi_{*x}^{-1} = (\varphi^{-1})_{*\varphi(x)}.$$

We call the chart $(U \times \mathbb{R}^n, \varphi_*)$ the chart in TM *associated to* (U, φ).

EXERCISE 30. *Let $f : M \to N$ be a smooth map between manifolds M and N. Show smoothness of $f_* : TM \to TN$.*

EXERCISE 31. *(*) Show that $\cup_x (T_x M \times T_x M)$ carries the canonical structure of a smooth manifold (fibre product). Show that*

$$(v, w) \in \cup_x (T_x M \times T_x M) \mapsto (v + w) \in TM$$

is a smooth map between manifolds.

The cotangent space at each $x \in M$ is given as the vector space dual of the tangent space $T_x^* M := (T_x M)^*$. The union of these spaces $T^* M := \cup_x T_x^* M$ can be made into a smooth manifold.

EXERCISE 32. *Compute the transition functions of the associated charts in the cotangent bundle of a smooth manifold.*

Given a differentiable map $f : M \to N$ we define its pull-back $f^* : T^* N \to T^* M$ via $f^*(\alpha)(X) := \alpha(f_*(X))$ for $\alpha \in T^*_{f(x)} N$ and $X \in T_x M$.

EXERCISE 33. *Let $f : M \to \mathbb{R}$ be a smooth function on the manifold M. Show that its differential df_x defines a smooth map $df : M \to T^* M$ between manifolds. Let $\pi : T^* M \to M$ denote the natural projection which assigns to a cotangent vector its base point. Show that this is a smooth map between manifolds. Compute $\pi \circ df$.*

2. Lagrangian Dynamical Systems on Manifolds

We will now generalize the concept of a Lagrangian dynamical system to manifolds. Let M be a smooth manifold and $L : TM \to \mathbb{R}$ a differentiable function on its tangent bundle. The curve $\gamma : [t_0, t_1] \to M$ (or a loop $\gamma : \mathbb{S}^1 \to M$) is called a *motion in the Lagrangian system* (M, L) if it is an extremal of the functional

$$\mathcal{L}(\gamma) := \int_{t_0}^{t_1} L(\dot{\gamma})dt$$

with fixed boundary values (or $\mathcal{L}(\gamma) := \int_{\mathbb{S}^1} L(\dot{\gamma})dt$).
We saw earlier that if $M = \mathbb{R}^n$ the formulation of the variational principle for Lagrangian systems does not depend on the particular coordinates. That means for any choice $(q_1, ..., q_n)$ of local coordinates on an open subset $U \subset \mathbb{R}^n$, a curve $\gamma \subset U$ is extremal if

$$\delta\mathcal{L}_\gamma := \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0.$$

The function L has to be expressed in terms of the cordinates. In fact, slightly more is true

Exercise 34. *Prove that $\delta\mathcal{L}_\gamma(t)$ is a well-defined object on the manifold M, i.e. it does not depend on the choice of coordinates. Is it a tangent or a cotangent vector? Explain your answer.*

Proposition 26. *A curve $\gamma : [t_0, t_1]$ is a motion of the Lagrangian system (M, L) if and only if*

$$\delta\mathcal{L}_\gamma(t) = 0$$

for each t.

Proof. □

2.1. Geodesics. If $M \subset \mathbb{R}^N$ is a submanifold of the euclidean space we have a natural concept for the length l of a differentiable curve $\gamma : [t_0, t_1] \to M$:

$$l(\gamma) := \int_{t_0}^{t_1} |\dot{\gamma}(t)|dt$$

Notice that $|v| = \sqrt{v \cdot v}$ is determined from the scalar product. In local coordinates of M the scaler product looks much less "straight": Let (U, φ) be such a coordinate $U \subset \mathbb{R}^n$, and $\{e_1, ..., e_n\}$ the standard basis. Then

$$g_{ij}(x) := \varphi_{*x}e_i \cdot \varphi_{*x}e_j$$

defines a family of symmetric, positive definite matrices. This is the *Riemannian structure* induced on the submanifold expressed in local coordinates.

EXERCISE 35. *Compute the Riemann structure induced on* $\mathbb{S}^2 := \{x \in \mathbb{R}^3 \mid |x| = 1\}$ *in terms of stereographic and angular coordinates.*

hence if $\gamma \subset U$ we obtain

$$l(\varphi(\gamma)) = \int_{t_0}^{t_1} \sqrt{\sum_{ij} g(\gamma(t))_{ij}\dot{\gamma}_i\dot{\gamma}_j dt}$$

Thus Riemannian structures can be understood as *intrinsic* objects of manifolds:

DEFINITION 2. *Let M be a differentiable manifold. A* Riemannian structure $g = g(x)$, $x \in M$, *is a family of scalar products on the tangent spaces T_xM which, expressed in any chart (U, φ) of M, is given by a differntiable family of matrices.*

Extremal points of the length functional are called *geodesics.* A given geodesic $\gamma : \mathbb{R} \to M$ is indeed the shortest connection between $\gamma(t)$ and $\gamma(t+\epsilon)$ where $\epsilon > 0$ is small depending on γ and t.

EXERCISE 36. *Compute δl_γ in a coordinate chart. This is the differential equation for geodesics.*

EXERCISE 37. *Find the geodesics on $\mathbb{S}^2 \subset \mathbb{R}^3$ given above uisng stereographic and angular coordinates.*

2.2. Natural Systems. The interpretation of the scalar product in the context of submanifolds of $\mathbb{R}^N$ leads to a natural generalization of kinetic energy to Riemannian manifolds:

$$T(v) := \frac{1}{2} g_x(v, v),$$

where $v \in T_xM$ is a tangent vector. A *natural Lagrangian system* on a Riemannian manifold (M, g) is given by a Lagrangian of the form

$$\begin{aligned} L &: TM \to \mathbb{R} \\ L(v) &= T(v) - U(x), \end{aligned}$$

for a smooth function $U : M \to \mathbb{R}$ which serves as the potential.

The main approach to determine the motions of (M, L) is to express the system in local coordinates near the projection $\pi(v)$ of the initial point $v \in TM$. Second, one solves the equations within that neighbourhood until the location coordinates lie outside. Third one picks a point very close to the exit and repeats the procedure with initial values given by the location and the velocity at this point. The main calculation for a natural system is to determine the kinetic term.

EXERCISE 38. *Show that the motion* $\gamma : \mathbb{R} \to M$ *of a free massive point* $U \equiv 0$ *in a Riemannian manifold is given by the geodesic with the same initial point and direction in* M. *Show that the* velocity $|\dot\gamma|_{g(\gamma(t))}$ *is preserved.*

3. D'Alembert's Principle

In this section we derive the variational principle of a system described by the Lagrangian $L = L(x, \dot x)$ in $\mathbb{R}^N$ which is restricted to stay on a submanifold $M^n \subset \mathbb{R}^N$. This is called *D'Alembert's principle.* Let M be around the initial point described as the regular zero set of $N-n$ funtions $\mathbf{f} := (f_1, ..., f_{N-n})$. Then the restricted system can be locally interpreted as a Lagrangian system in $\mathbb{R}^N$ with constraints: $\gamma : [t_0, t_1] \to \mathbb{R}^N$

$$\mathcal{L}(\gamma) = \int_{t_0}^{t_1} L(\gamma(t), \dot\gamma) dt$$
$$\mathbf{f}(\gamma(t)) = 0 \quad \text{for all } t.$$

Variation of curves γ with fixed end points which satisfy the constraint lead to the following description of motions:

THEOREM 27 (D'Alembert's Principle). *The curve* $\gamma : [t_0, t_1] \to \mathbb{R}^N$ *is a critical point of the functonal* $\mathcal{L}$ *restricted to curves with the same end points satisfying the constraint if and only if*

(1) $\gamma \subset M$, *which is obvious from the constraint,*
(2) *for all* t *and* $\xi \in T_{\gamma(t)}$ *with* $\nabla f_i(\gamma(t)) \cdot \xi = 0$ *for all* i*:*

$$\Big(\frac{d}{dt}\frac{\partial L}{\partial \dot x} + \frac{\partial L}{\partial x}\Big) \cdot \xi = 0$$

The main point is that the differential $\delta\mathcal{L}$ vanishes in directions tangential to M. For the Lagrangian of a mechanical system of one massive point $L = T - U$ this can be interpreted as a force $m\ddot x + \partial U/\partial x$. The statement is that for a motion this force is perpendicular to the submanifold M or that the work of the force along each motion vanishes at any time. In fact for the spherical pendulum of the beginning of that chapter, the force on the massive point is provided by the rod and is parallel to it at any time.

EXERCISE 39. *Determine this force for the triangle of the second example and give a quantitative interpretation along the lines of the spherical pendulum.*

COROLLARY 28. *The restriction of a Lagrangian dynamical system* $L = L(x, \dot x)$ *on* $\mathbb{R}^N$ *to a subanifold* $M \subset \mathbb{R}^N$, $(M, L|_{TM})$ *is described as a variational principle of the original system restricted to curves in* M. *In fact, this holds for any restriction of a Lagrangian dynamical system* (M, L) *on a manifold to a submanifold* $N \subset M$.

EXERCISE 40. *Determine the geodesics of* $\mathbb{S}^2 \subset \mathbb{R}^3$ *as defined above from d'Alembert's principle.*

4. Noether's Theorem

We formulate the conservaton law for Lagrangian systems which generalizes the conservation laws of Newton mechanics: Each conserved quantity corresponds to a one–parameter group of diffeomorphisms of the configuration space which preserves the Lagrangian.

THEOREM 29. *Let* (M, L) *be a Lagrangian system. Assume for simplicity that it is time–independent. Let* $h_\tau : M \to M$, $\tau \in \mathbb{R}$ *and* $h_0 = id_M$ *be the identity which preserves the Lagrangian* $L : TM \to \mathbb{R}$,

$$(h_\tau)^* L(v) = L(h_{\tau *} v) = L(v)$$

for all $v \in TM$. *Then, the system of Lagrange—equations coresponding to* L *has a first integral* $I : TM \to \mathbb{R}$ *given in local coordinates* q *of* M *by*

$$I(q, \dot{q}) := \frac{\partial L}{\partial \dot{q}} \frac{dh_\tau(q)}{d\tau}\Big|_{\tau=0}.$$

EXERCISE 41. *Express the first integral without using local coordinates.*

PROOF. Although all involved quantities are global objects the most convenient way to prove the statement is to do it in local coordinates $(q, \dot{q}) \in TM|_U$. Let $\gamma : I \subset \mathbb{R} \to U$ be a solution in U of

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

Since L is invariant under the flow we have

$$\frac{d}{ds}[L(h_s(\gamma(t)))] = \frac{\partial L}{\partial q}\frac{\partial [h_s(\gamma(t))]}{\partial s} + \frac{\partial L}{\partial \dot{q}}\frac{\partial^2 [h_s(\gamma(t))]}{\partial s \partial t} = 0.$$

We express $\partial L / \partial q$ using the Euler–Lagrange equations to obtain

$$0 = \Big(\frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{q}}\Big)\frac{\partial [h_s(\gamma(t))]}{\partial s} + \frac{\partial L}{\partial \dot{q}}\frac{\partial^2 [h_s(\gamma(t))]}{\partial s \partial t} = \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}}\frac{dh_s(\gamma(t))}{ds}\Big).$$

□

EXAMPLE 9. *We consider a system of massive points* $(x_1, ..., x_n) \in \mathbb{R}^{3n}$ *of masses* m_i. *Let* $L = \sum_i m_i \frac{\dot{x}_i}{2} - U(x)$ *be its Lagrangian. We allow the points to satisfy a constraint* $x \in M \subset \mathbb{R}^{3n}$. *Assume that the system is invariant under translations along* $e_1 \in \mathbb{R}^3$. *Then the first coordinate of the center of mass is a linear function of time. Indeed,* $\frac{dh_s}{ds} \equiv e_1$ *and hence the first integral is*

$$P_1 = \sum_i \frac{\partial L}{\partial \dot{x}_i} \cdot e_1 = \Big(\sum_i m_i \dot{x}_i\Big) \cdot e_1,$$

the first component of the total moment of the system.

EXAMPLE 10. *Assume that the system of* n *massive points is invariant under rotations about the axis* e_1. *Then the angular momentum corresponding to this axis is preserved:*

$$M_1 = \Big(\sum_i m_i x_i \times \dot{x}_i\Big) \cdot e_1.$$

Notice that this is a generalization of these properties we have discuused before for systems without constraints.

EXAMPLE 11. *The system of a massive point suspended from a fixed point on a rod on the surface of the earth has a rotational symmetry with the axis being the perpendicular through the point. Choose the coordinate system* (x_1, x_2, x_3) *in* $\mathbb{R}^3$ *such that the* x_3*–axis coincides with the axis of rotation. The vector* $dh_s/ds|_{s=0}(x)$ *of this symmetry is given by* $e_3 \times x$ *with* x *understood as the position vector of the point on the sphere. Then the first integral is given by the angular momentum*

$$M_3 = m\dot{x} \cdot (e_3 \times x) = (m\dot{x} \times x) \cdot e_3.$$

CHAPTER 4

Mechanics of the Rigid Body

We exploit Noether's theorem on the connection between symmetries and conservation laws to derive and study the equations of motions for massive objects which have spatial dimensions.

1. Mechanical Description of a Rigid Body

A good model for a *rigid body* is given by a measurable function $\rho : \mathbb{R}^3 \to \mathbb{R}$ of compact support, the *density* distribution of the considered object. All physically relevant quantities will be computable from ρ. We may now fix three points on the body which do not lie on a line. Their relative position will not change, at least under reasonable circumstances. On the other hand, the position of these three points in space will completely determine the position of the body. Hence the configuration space of body is that of a triangle. We already discussed the latter to be $\mathbb{R}^3 \times \mathbf{SO(3)}$.

1.1. Celestial Mechanics Revisited. We drop the assumption that the planet is a dimensionless massive objects and assume instead that it is a massive body in a central field. The motions can be described by paths $(b(t), B(t)) \in \mathbb{R}^3 \times \mathbf{SO(3)}$. They describe for every point $x \in \mathbb{R}^3$ of the body its position in space $B(t)x + b(t)$. Hence potential and kinetic energy are given by

$$U(b, B) = \int \frac{\rho(x)}{|Bx + b|} d^3x$$
$$T(b, B) = \frac{1}{2} \int \rho(x) |\dot{B}x + \dot{b}|^2 d^3x$$

Assume now that $0 \in \mathbb{R}^3$ is the *center of mass*, which means that

$$\int x\rho(x) d^3x = 0.$$

Then

$$\int \rho(x)(\dot{B}x \cdot \dot{b}) d^3x = \dot{B}^T \dot{b} \cdot \int x\rho(x)) d^3x = 0,$$

and hence

$$T(b, B) = \frac{1}{2} \int \rho(x)(|\dot{B}x|^2 + |\dot{b}|^2) d^3x.$$

It follows that the Lagrangian $L = T - U$ is invariant under the action of $A \in \mathbf{SO(3)}$ on the configuration space:

$$(b, B) \mapsto (Ab, AB)$$

A is a rotation in $\mathbb{R}^3$ about an axis through 0 described by $a \in \mathbb{R}^3$. Denote by $A(t)$ the family of rotations with fixed axis and $A(0) = E_3$, the identity matrix, then

$$\dot{A}(0)x = a \times x$$

for any $x \in \mathbb{R}^3$. The relation between $\dot{A}(0)$ and a is as follows

$$\dot{A}(0) = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}, \quad a = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Applying Noether's theorem we find that the following quantity is preserved for all $a \in \mathbb{R}^3$

$$\begin{aligned} \frac{\partial L}{\partial \dot{b}}(a \times b) + \frac{\partial L}{\partial \dot{B}}(a \times B) &= \int \rho(x)(\dot{B}x \cdot (a \times Bx) + \dot{b} \cdot (a \times b))d^3x \\ &= (\int \rho(x)(Bx \times \dot{B}x + b \times \dot{b})d^3x) \cdot a \end{aligned} \tag{6}$$

The second sum in the integrand gives $M_0 = m(b \times \dot{b})$, which could be interpreted as the angular momentum of the center of mass with respect to the central point of the field and the total mass of the body attached to it. The first sum gives the *total angular momentum* of the rigid body with respect to its center of mass

$$\mu = \int \rho(x)(Bx \times \dot{B}x)d^3x.$$

EXERCISE 42. *Assume that the earth has the perfect shape of a round ball. Show that then both quantities M_0 and μ are conserved quantities.*

EXERCISE 43. *Show that the momentum of a free rigid body is preserved, i.e. its center of mass is non–accelerated.*

1.2. Inertia Tensor. We study the rotation of the rigid body separately assuming the center of mass is a resting point. In fact we may assume a situation in which $0 \in \mathbb{R}^3$ is a fixed point of the motion, i.e. the system has configuration space $\mathbf{SO(3)}$.

Let us describe the total angular momentum with respect to 0 in terms of the mass distribution of the rigid body. The angular velocity with respect to the coordinates which move with the body is expressed by $\Omega = B^{-1}\dot{B}$ which can be considered as a vector in $\mathbb{R}^3$. We have the following linear symmetric operator on $\mathbb{R}^3$:

$$A : v \mapsto \int \rho(x)(x \times v \times x)d^3x.$$

Then the angular momentum with respect to this coordinate system is given by $M := B^{-1}\mu = A\Omega$.
The eigenvectors of A are called the *main axes* of the body in 0. The set of vectors

$$\mathcal{E} := \{v \mid (Av \cdot v) = 1\}$$

is an ellipsoid and is called the *ellipsoid of inertia* of the rigid body. This set can be given a physical meaning. Assume the body rotates about an axis through 0 with angular velocity $\Omega \in \mathbb{R}^3$. The quantity $I_e = \int \rho(x)d(x)^2$ is the *inertia* of the body with respect to the axis given by Ω, $d(x)$ is the distance from the axis spanned by Ω. Hence the eigenvalues of the inertia tensor A are the inertia with respect to the main axes. If e is a unit vector given by a point on the unit sphere, then $e/\sqrt{I_e}$ is the corresponding point on the ellipsoid. That has therefore the form

$$I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2 = 1$$

where after a possible orthogonal transformation e_i, $i = 1, 2, 3$, are the main axes.

EXERCISE 44. *Show that the eigenvalues (I_1, I_2, I_3) of the inertia tensor of a rigid body satisfy the triangle inequality*

$$I_1 \leq I_2 + I_3, \quad I_2 \leq I_3 + I_1, \quad I_3 \leq I_1 + I_2.$$

The kinetic energy is given by $2T = (A\Omega) \cdot \Omega = M \cdot \Omega$ and is equal to

$$T = \frac{1}{2} I_e |\Omega|^2.$$

PROPOSITION 30 (Steiner). *Let I_0, I be the inertia of a rigid body with respect to two parallel axes, the one for I_0 going through the center of mass. Then*

$$I = I_0 + md^2$$

where m is the mass of the body and d is the distance of the two axes.

PROOF. Pick a coordinate system such that the axes are parallel to e_3, 0 is the center of mass, and the second axis is the line $(x_1, x_2) = (d, 0)$. Denote by $d_0(x)$ and $d(x)$ the distance from the two axes. Then

$$d_0(x) = |(x_1, x_2, 0)|, \quad d^2(x) = |(x_1, x_2, 0) - (d, 0, 0)|^2 = d_0^2(x) + d^2 - dx_1.$$

Integrating with respect to measure $\rho(x)d^3x$ gives the desired identity since $\int x_1\rho(x)d^3x = 0$. ∎

EXERCISE 45. *Let $\Omega \in \mathbb{R}^3$ be the angular velocity vector of a rotating rigid body. Show that the total angular momentum M is parallel to the normal to the ellipsoid of inertia at the intersection of the line through Ω with it.*

Rotational symmetries of the rigid body with respect to an axis will be reflected in the inertia ellipsoid. In particular, if the body is symmetric with respect to a rotation by an angle $2\pi/k$, $k > 2$, the ellipsoid will be symmetric with respect to such a rotation. It follows that the ellipsoid has to be round with respect to that axis.

EXERCISE 46. *Determine the ellipsoid of inertia for your favorite platonic body with respect to its center.*

The inertia tensor (or the ellipsoid) determines the kinetic energy of the rotation of the free rigid body and hence the motion completely. That is, given two rigid bodies with the same inertia ellipsoid and the same initial angular velocity they will follow the same motion.

2. Equation of Motion of a Rigid Body

We will describe the motion of a rigid body fixed in a resting point without outer forces.

2.1. Euler's Equations. Let $0 \in \mathbb{R}^3$ be the fixed point of the system. The total angular momentum of the body with respect to 0 is preserved. This lead to a descrption of the the total angular momentum and hence the angular velocity with respect to the moving cordinates.

PROPOSITION 31 (Euler's Equations). *For a rigid body fixed in* 0 *without outer forces we have*

$$\frac{dM}{dt} = M \times \Omega.$$

PROOF. We have $0 = \dot{\mu} = \dot{B}M + B\dot{M} = B(B^{-1}\dot{B}M + \dot{M}) = B(\Omega \times M + \dot{M})$. □

Via $M = A\Omega$ this equation can be alternatively considered as equation for $M = M(t)$ or $\Omega = \Omega(t)$. Let

$$\Omega = \Omega_1 e_1 + \Omega_2 e_2 + \Omega_3 e_3, \quad M = M_1 e_1 + M_2 e_2 + M_3 e_3$$

be the decompositions of the two vectors with respect to the main axes. Then $M_i = I_i \Omega_i$ and

$$\begin{aligned} \frac{dM_1}{dt} &= a_1 M_2 M_3 \\ \frac{dM_2}{dt} &= a_2 M_3 M_1 \\ \frac{dM_3}{dt} &= a_3 M_1 M_2 \end{aligned}$$

with

$$a_1 = \frac{I_2 - I_3}{I_2 I_3}, \quad a_2 = \frac{I_3 - I_1}{I_3 I_1}, \quad a_3 = \frac{I_1 - I_2}{I_1 I_2}.$$

In terms of Ω the equation reads

$$\begin{aligned} I_1 \frac{d\Omega_1}{dt} &= (I_2 - I_3)\Omega_2\Omega_3 \\ I_2 \frac{d\Omega_2}{dt} &= (I_3 - I_1)\Omega_3\Omega_1 \\ I_3 \frac{d\Omega_3}{dt} &= (I_1 - I_2)\Omega_1\Omega_2. \end{aligned}$$

Since the energy E and the total momentum μ are preserved, and $M = B^{-1}\mu$, the quantities

$$2E = \frac{M_1^2}{I_1} + \frac{M_2^2}{I_2} + \frac{M_3^2}{I_3}, \quad M^2 = M_1^2 + M_2^2 + M_3^2$$

are also preserved. Geometrically that means that M lies on the intersection of the ellipsoid $E = E_0$ with the sphere $M^2 = \mu_0^2$. Assume that $I_1 > I_2 > I_3$. Then for $M < \sqrt{2EI_3}$ or $M > \sqrt{2EI_1}$ the intersection is empty. For $M = \sqrt{EI_3}$ it consists of the two intersection points of the ellipsoid with the small half axes, for $\sqrt{2EI_3} < M < \sqrt{2EI_2}$ it consists of two loops whose centers are these two points, for $M = \sqrt{2EI_2}$ it consists of two circles intersecting at the two intersections of the middle half axes, while for $\sqrt{2EI_2} < M < \sqrt{2EI_1}$ it consists of two loops whose centers are the intersection of the big half axes with the ellipsoid which are exactly the intersection of the sphere with the ellipsoid for $M = \sqrt{2EI_1}$. Each of the 6 intersections of the half axes with the ellipsoid corresponds to a stationary point of the equation. That means that Ω and M and therefore ω and μ are parallel. Since μ is preserved, that means that the body rotates around a fixed axis in space. Since for a solution to Euler's equations $M = M(t)$ will stay on the curve cut out by the ellipsoid and the sphere that means that the stationary solutions $M = M_i e_i$ are stable for $i = 1, 3$. They are unstable for $M = M_2 e_2$.

2.2. Poinsot's Description of Motions. Let $B\mathcal{E}$ be the position of the ellipsoid of inertia in space. Then we have the following description

THEOREM 32. *The ellipsoid of inertia rolls without sliding on a fixed plane in space which is perpendicular to the total angular momentum μ.*

PROOF. There are two planes perpendicular to μ and tangential to $B\mathcal{E}$. All we need to show is that these touch the ellipsoid in the intersection points $\pm\xi = \omega/\sqrt{2T}$ of the axis of rotation and that their distance from 0 is fixed.

For the first notice that the normal to any point Ω of the (fixed) ellipsoid $\mathcal{E}$ was parallel to the total momentum M (see Exercise 45), hence the normal of $B\mathcal{E}$ at $\pm\xi$ is parallel to μ and the statement follows.
In order to lie in the ellipsoid ω has to be recaled by $\sqrt{2T}$ since $A\Omega \cdot \Omega = 2T$. Hence the distance of the plane from the origin is given by

$$\frac{\omega}{\sqrt{2T}} \cdot \frac{\mu}{|\mu|} = \frac{\omega \cdot \mu}{\sqrt{2T}|\mu|} = \frac{2T}{\sqrt{2T}|\mu|} = \frac{sqrt2T}{|\mu|}.$$

Hence it is constant ($T = E$ and $|\mu|$ are conserved).
The description of the motion follows from these observations, since the current axis of rotation is given by the line through 0 and the tangential point. If the ellipsoid slips the rotation of the ellipsoid will contradict the plane being tangential to it. $\square$

Study of conservative perturbations of the free system show that the description remains valid if the potential energy is very small in comparison to the kinetic energy of the rigid body (see [**?**]). For a planet in a solar system the part of the Lagrangian L which describes the action of the massive point at the center of mass with mass m

$$U - \frac{m}{2}|b|^2,$$

could be considered as a (time-dependent) conservative perturbation of the Lagrangian of the free system. It is very small compared to the original kinetic term. Hence, the the rotation of the earth around its axis is described by Poinsot. On the other hand, therefore Kepler's description of the orbit of the center of mass of the earth around the sun remains valid.

Part 3

Hamiltonian Dynamics

CHAPTER 5

Symplectic Manifolds

1. Preliminaries

1.1. Differential Forms. This chapter is mainly for reminding the reader of some important properties of differential forms and providing the

1.1.1. *Forms on* $\mathbb{R}^n$. Let $I = \{i_1, ..., i_k\} \subset \{1, ..., n\}$ be an ordered subset of indices. Then by

$$dx^I = dx^{i_1} \wedge ... \wedge dx^{i_k}$$

we denote the antisymmetric k–linear form on $\mathbb{R}^n$, which assigns to $(v_1, ..., v_k)$, $v_i = (v_i^1, ..., v_i^n) \in \mathbb{R}^n$, $i = 1, ..., k$ the determinant

$$\begin{vmatrix} v_1^{i_1} & \dots & v_k^{i_1} \\ \vdots & & \vdots \\ v_1^{i_k} & \dots & v_k^{i_k} \end{vmatrix}.$$

In particular in euclidean coordinates we have

$$dx^{i_1} \wedge ... \wedge dx^{i_k}(\frac{\partial}{\partial x_{i_1}}, ..., \frac{\partial}{\partial x_{i_k}}) = 1.$$

The wedge product of two such *monomials* dx^I, dx^J is given by

$$dx^I \wedge dx^J := \begin{cases} 0 & \text{if } I \cap J \neq \emptyset \\ (-1)^{\sigma(I,J)} dx^{I \cup J} & \text{else} \end{cases}$$

where $\sigma(I, J) \in \mathbb{N}$ is the number of interchanging neighboring pairs to order the set $\{i_1, ..., i_k, j_1, ..., j_l\}$. $\{dx^I\}_{|I|=k}$ is a basis of the vector space of k–linear, antisymmetric forms on $\mathbb{R}^n$ (short k–form) denoted by $\Lambda^k\mathbb{R}^n$. Therefore, we may linearily extend $. \wedge .$ to a bilinear operator assigning to any k–form α and l-form β a $(k + l)$–form $\alpha \wedge \beta$.
Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be a homomorphism between vector spaces. We define a homomorphism $A^* : \Lambda^k\mathbb{R}^m \to \Lambda^k\mathbb{R}^n$ via

$$(A^*\alpha)(v_1, ..., v_k) := \alpha(Av_1, ..., Av_k).$$

EXERCISE 47. *Show that* $A^*(\alpha \wedge \beta) = A^*\alpha \wedge A^*\beta$ *for any forms* α *and* β.

Also, any family of differential forms on $\mathbb{R}^n$ on $U \subset \mathbb{R}^n$ can be written as

$$\sum_{I:|I|=k} \alpha_I(x) dx^I.$$

We call it smooth if the coefficients $\alpha_I(x)$ depend smoothly on x. The space of such k–forms is denoted by $\Omega^k(U)$. The wedge product extends straightforward to this notion. If $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^m$ is a differentiable map, and $\alpha \in \Omega(V)$ then we define

$$(f^*\alpha)(x) := (df_x)^*(\alpha(f(x)).$$

The *exterior differential* of a form is given by $\mathbb{R}$–linear extension of

$$d(\alpha_I dx^I) = \sum_j \frac{\partial \alpha_I}{\partial x_j} dx^j \wedge dx^I.$$

EXERCISE 48. *Show that $d^2 = 0$. Show that for any smooth map $f : U \to V$ $d \circ f^* = f^* \circ d$. Prove that for $\alpha \in \Omega^k(U)$ and $\beta \in \Omega^l(U)$*

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta.$$

1.2. Stokes' Theorem. Let M^n be a smooth n–dimensional manifold. Then a smooth k–form $\alpha = \alpha(x)$ on M is a family of k–linear, antisymmetric forms $\alpha(x) \in \Lambda^k(T_x M)$ such that for any smooth chart $\varphi : U \to M$, $\varphi^*\alpha \in \Omega^k(U)$. Let $N \subset M$ be a smooth submanifold. Then for $\alpha \in \Omega^k(M)$ the restriction $\alpha|_N$ is given by $\alpha|_N := i^*\alpha$ where $i : N \to M$ is the inclusion. Let $\alpha = \alpha(x) dx^1 \wedge ... \wedge dx^k \in \Omega^k(U)$ for $U \subset \mathbb{R}^k$ and $K \subset U$ be a measurable subset then

$$\int_K \alpha := \int \alpha(x) d^k(x).$$

This can be extended to any *oriented* manifold M^k, i.e. one for which there is an atlas for which the differential of all transition functions has a positive determinant. If M is a compact $(k+1)$–manifold with smooth boundary ∂M we have

THEOREM 33.

$$\int_M d\alpha = \int_{\partial M} \alpha$$

1.3. Lie Derivatives. Let $X = (X_1, ..., X_n), Y = (Y_1, ..., Y_n)$ be smooth vector fields on a manifold $U \subset \mathbb{R}^n$. Their *Lie bracket* is defined via

$$[X, Y]_i(x) = d_x Y_i(X(x)) + d_x X_i(Y(x)).$$

This notion behaves well under diffeomorphisms $f : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^n$, namely $f_*[X, Y] = [f_*X, f_*Y]$. Hence we obtain the Lie bracket on smooth vector field on manifolds.

EXERCISE 49. *Show that* $[.,.]$ *satisfies Leibnitz' rule and Jacobi's identity*

$$[X, fY] = f[X,Y] + X(f)Y$$
$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

for vector fields X, Y, Z *and a smooth function* f.

One also calls $\mathcal{L}_X Y := [X,Y]$ the *Lie derivative* of Y along X. That Lie derivative extends to all tensors on the manifold, in particular for a k–form $\alpha \in \Omega^k(M)$ we define

$$\mathcal{L}_X \alpha(Y_1, ..., Y_k) := X(\alpha(Y_1, ..., Y_k)) - \alpha([X, Y_1], Y_2, ..., Y_k) - ... - \alpha(Y_1, ..., [X, Y_k]).$$

Let $\{\Phi_t\}_t$ be a family of diffeomorphisms on a manifold M. Denote by

$$X_t := \frac{d\Phi_t}{dt}$$

the corresponding (time–dependent)) vector field. Then for the action Φ_*^* on any tensor we have

$$\mathcal{L}_X = \frac{d}{dt}\Phi_*^*.$$

The contraction of a differential form $\alpha \in \Omega^k(M)$ along a vector field X is defined via

$$i_X \alpha(Y_1, ..., Y_{k-1}) = \alpha(X, Y_1, ..., Y_{k-1}).$$

THEOREM 34 (Cartan's Formula). *Let* X *be a vector field and* α *be a differential form then*

$$\mathcal{L}_X \alpha = i_X d\alpha + d(i_X \alpha).$$

PROOF. We first show the formula for $x \in M$ where $X(x) \neq 0$. In this case it is possible to choose coordinates around x such that $X = \partial/\partial x_1$. Then for $\alpha(x) = \sum_I \alpha_I dx^I$ we have

$$\mathcal{L}_{\partial/\partial x_1} \alpha = \sum_I \frac{\partial \alpha_I}{\partial x_1} dx^I.$$

It is straightforward to see that this is also the same as the expression given by Cartan's formula. On the other hand, the formula is obviously true if X vanishes in a neighborhood of x. The complement of these two sets is contained in their closure. Since both sides define continous forms the assertion follows for all points in M. $\square$

2. Symplectic Forms

Any bilinear form ω on a vector space V defines a homomorphism by

$$v \in V \mapsto \omega(v,.) \in V^*.$$

We call such a bilinear form *non–degenerate* if this is an isomorphism or, alternatively for finite–dimensional V, if the corresponding matrix $(\omega(v_i, v_j))_{ij}$ with respect to some basis $\{v_i\}_i$ of V has full rank.

Definition 3. *A* symplectic manifold *is a manifold M equipped with a non–degenerate, closed 2–form $\omega \in \Omega^2(M)$.*

First thing to note is that such a manifold has to be even–dimensional. Indeed, on the linear level, given any antisymmetric bilinear form ω on a vector space V, one can find a *standard basis* $\{v_i\}$ such that $\omega(v_{2i-1}, v_{2i}) = 1 = -\omega(v_{2i}, v_{2i-1})$ and all other coefficients of the matrix vanish.

Exercise 50. *Show that any two symplectic vector spaces (V_i, ω_i) of the same dimension $2n$ are isomorphic.*

Exercise 51. *Show that for a non-degenerate, anti–symmetric bilinear (short: symplectic) form on a vector space V, there exists a complex structure $J \in End(V)$, $J^2 = -E$, such that the bilinear form defined through $g(v, w) := \omega(v, Jw)$ is a euclidean inner product (symmetric and non–degenerate).*

Exercise 52. *Show the converse statement of the previous exercise: If g is an inner product of a vector space V and v admits a complex structure J which is an isometry,*

$$g(Jv, Jw) = g(v, w), \text{ for all } v, w \in V,$$

then the bilinear form defined through

$$\omega(v, w) := g(Jv, w)$$

is symplectic.

Exercise 53. *Show that for a symplectic form ω on an $2n$–dimensional vector space V, $\omega^n = \omega \wedge ... \wedge \omega$ defines a volume form, i.e. a non-vanishing $2n$–form.*

Hence, any symplectic manifold is orientable. For that we allow only charts (U, φ), such that $\varphi^*(\omega^n) = f(x)dx^1 \wedge ... \wedge dx^{2n}$ with $f(x)$ *positive*.

EXERCISE 54. *Show that this choice of coordinates defines an orientation on M, namely for any two charts $(U, \varphi), (V, \psi)$ in the atlas defined above, the determinant of the transition funtion is positive:*

$$\det\left(\frac{\partial(\psi \circ \varphi)_i}{\partial x_j}_{ij}\right) > 0.$$

Another

EXAMPLE 12 (Standard Structure on $\mathbb{R}^{2n}$). *Let ω be the symplectic form on $\mathbb{R}^{2n}$ such that standard basis $\{e_i\}_{i=1}^{2n}$ is a standard basis with respect to ω. Since all tangent spaces are canonical identified $T_x\mathbb{R}^{2n} \cong \mathbb{R}^{2n}$ we may define a differential form on $\mathbb{R}^{2n}$ denoted by ω_{st}. In standard* coordinates *$(x_1, ..., x_{2n})$ we have*

$$\omega_{st} = \sum_{i=1}^{n} dx_{2i-1} \wedge dx_{2i}.$$

EXAMPLE 13 (Kähler Manifolds). *A* Hermitian manifold *is a complex manifold M with a smooth Hermitian structure H on each tangent space considered as a complex vectopr space through the integrable almost complex structure $J \in End(TM)$, $J^2 = -1$. H decomposes into real and imaginary part $H = \Re H + i\Im H$. $g = \Re H$ is the underlying Riemannian structure on M while $\omega = \Im H$ is the non–degenerate skew–symmetric Hermitian form. (M, J, H) is called* Kähler manifold *if ω is closed, $d\omega = 0$. Hence any such Kähler manifold is symplectic. Any smooth algeraic variety $V \subset \mathbb{CP}^N$ is a Kähler manifold. The Kähler metric (and hence Kähler form) are given by restricting the corresponding structure on $\mathbb{CP}^N$ to V.*

EXAMPLE 14 (Cotangent Bundles). *The most important examples are provided by a canonical structure on cotangent bundles. Let L be any smooth manifold. Denote by $\pi : T^*L \to L$ the natural projection of a cotangent vector onto its base point. This is a smooth map between manifolds. There is a canonical 1–form $\theta \in \Omega^1(T^*L)$ which we call* Poincaré–Cartan–form *defined in the following way: let $\alpha \in T_x^*L$ be a cotangent vector and $X \in T_\alpha(T^*L)$ be a tangent vector at α. Note that $\pi(\alpha) = x$ and $\pi_*(X) \in T_xL$. Then*

$$\theta_\alpha(X) := \alpha_x(\pi_* X).$$

Pick any chart (U, φ) in L with coordinate $(q_1, ..., q_n) \in U$. Then the associated chart $(\mathbb{R}^n \times U, \Phi)$ with coordinates $(p_1, ..., p_n, q_1, ..., q_n)$ has the meaning

$$\Phi(p_1, ..., p_n, q_1, ..., q_n) = \sum_i p_i(\varphi^{-1})^* dq_i.$$

Then

$$\Phi^*\theta = \sum_i p_i dq_i.$$

Therefore $\Phi^*(d\theta) = d(\Phi^*\theta) = \sum_i dp_i \wedge dq_i$.

An endomorphism $A : V \to V$ of a vector space V equipped with a symplectic form ω is called *symplectic* if $A^*\omega = \omega$. Denote the set of these by $\mathbf{Sp}(2n)$. The composition of two such elements is again symplectic. Any such symlectic A also preserves the powers ω^k.

EXERCISE 55. *Show that for any* $A \in \mathbf{Sp}(2n)$

$$\det A = 1.$$

Hence any symplectic endomorphism A is invertible and $\mathbf{Sp}(2n)$ forms indeed a group.

3. Hamiltonian Systems

A Hamiltonian system is a triple (M, ω, H) consisting of a symplectic manifold (M, ω) and a smooth time–dependent function on M $H : M \times \mathbb{R} \to \mathbb{R}$, the *Hamiltonian* or *Hamilton function.*

Since ω is a non–degenerate bilinear form on each tangent space T_x, it defines an isomorpism, denoted by the same letter $\omega_x : T_xM \to T_x^*M$ for each $x \in M$. Therefore, H defines a vector field X_H on M through

$$\omega_x(X_H(x), X) = -dH_x(X)$$

for any $X \in T_xM$. X_H is called the *symplectic gradient* of H.

The motions of a Hamiltonian system are now given by paths $\gamma : [0, T] \to M$ which satisfy Hamilton's equations

$$\dot\gamma(t) = X_H(\gamma(t), t) \text{ or equivalently} \tag{7}$$

$$\omega(\dot\gamma(t), v) = d^M H_{\gamma(t),t}(v) \text{ for all } v \in T_{\gamma(t)}M \tag{8}$$

The superscript in the exterior derivative d^M shall emphasize that the derivatives are taken in M–direction only.

EXAMPLE 15. *Let* $H = H(p, q, t)$ *be a function on* $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$. *Its symplectic gradient can be easily computed:*

$$X_H(p, q, t) = -\frac{\partial H(p, q, t)}{\partial q}\frac{\partial}{\partial p} + \frac{\partial H(p, q, t)}{\partial p}\frac{\partial}{\partial q}$$

Therefore Hamilton's equations for the system (R^{2n}, ω_{st}, H) *are*

$$\dot p = -\frac{\partial H}{\partial q}$$

$$\dot q = \frac{\partial H}{\partial p}. \tag{9}$$

Consider a massive point with mass m *and* n *degrees of freedom in a field with potential* $U = U(x)$. *It follows that this is described by the Hamiltonian system* $(\mathbb{R}^{2n}, \omega_{st}, H(p,q)$ *where*

$$H(p,q) = \frac{|p|^2}{2m} + U(q).$$

3.1. Conservation Laws. It is thus naturally to consider H as the *energy* of the Hamiltonian system. We have the following

PROPOSITION 35. *Assume that* $\gamma : [0,T] \to M$ *is a solution to Hamilton's equations. Then*

$$\frac{d}{dt} H(\gamma(t), t) = \frac{\partial H}{\partial t}(\gamma(t), t).$$

In particular, if H *does not depend on* t *it is preserved.*

PROOF. A straightforward calculaton shows

$$\frac{d}{dt}[H(\gamma(t),t)] = d^M H_{\gamma(t)}(\dot{\gamma}(t)) + \frac{\partial H}{\partial t}(\gamma(t)) = -\omega_{\gamma(t)}(X_H(\gamma(t),t), \dot{\gamma}(t)) = -\omega_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) = 0$$

since ω is an alternating 2–form. □

Similarily easy to prove but more subtle to formulate is a structural conservation law, stating that ω and any of its powers ω^k are *integral invariants*. Given an open subset $U \subset M$ assume that there exists a $T > 0$ such that for any $x \in U$ the initial value problem

$$\dot{\gamma}(t) = X_H(\gamma(t), t) \tag{10}$$

$$\gamma(0) = x. \tag{11}$$

has a unique solution $\gamma : [0,T] \to M$ on the whole interval. Then we define a family of maps $\{\Phi_t : U \to M\}_{t=0}^T$ by

$$\Phi_t(x) = \gamma(t)$$

for γ solving the previous equations. Notice that the above conditions are satisfied for $U = M$ if M is a closed manifold.

EXERCISE 56. *Show that*

$$\frac{\partial}{\partial t}\Phi_t(x) = X_H(\Phi_t(x), t)$$

PROPOSITION 36. Φ_t *preserves the symplectic structure*

$$\Phi_t^* \omega = \omega.$$

In particular, ω^k *is an integral invariant for the system. Namely, if* $c : U \subset \mathbb{R}^{2k} \to M$ *is any differentiable map,*

$$\int_c \omega^k = \int_{\Phi_t(c)} \omega^k$$

where $\int_c \omega^k := \int_U c^*(\omega^k)$.

PROOF. Since $\Phi_0 = Id_M$ the statement, obviously holds for $t = 0$. Now ω idoes not depend on t. Hence it is sufficient to show that

$$(\Phi_t^{-1})^* \frac{\partial}{\partial t}(\Phi_t^* \omega) = 0.$$

This is the Lie derivative. Hence we have to show

$$0 = \mathcal{L}_{X_H(t)}\omega = i_{X_H(t)} d^M \omega + d^M(i_{X_H(t)}\omega) = d^M(i_{X_H(t)}\omega)$$

since $d^M\omega = 0$. But due to the alternative formulation (8)

$$i_{X_H(t)}\omega = -d^M(H(.,t)),$$

i.e. this is an exact form. Since $(d^M)^2 = 0$ the claim follows. □

A map $\Phi : M \to M$ of a symplectic manifold (M, ω) with $\Phi^*\omega = \omega$ is locally a diffeomeoprhism, since $d_x\Phi$ is invertible for every x. It is called *symplectomorphism* though one usually asks it to be bijective. If the symplectomorphism is generated by a (possibly time–dependent) Hamiltonian $\Phi = \Phi_1$ and $\{\Phi_t\}_t$ as in the previous proposition, then it is called *Hamiltonian diffeomorphism.*

3.2. Legendre Transform on Manifolds. We will see that Hamiltonian systems describe Lagrangian mechanical systems, generalizing Example 15. Suppose that (M, L) is a Lagrangian system. Recall that the fibrewise derivative is a map

$$\frac{\partial L}{\partial \dot q} : T_qM \to T_q^*M.$$

Assume that this is a diffeomeorphism and that $L : TM \to \mathbb{R}$ is strongly convex on the fibres. Recall that the (fibre-wise) Legendre transform of L was then given by

$$H(p, q) := p\dot q - L(q, \dot q),$$

where $p \in T_q^*M$ and $\dot q \in T_qM$ are related through $p = \frac{\partial L}{\partial \dot q}(q, \dot q)$. Notice that naturally H is a function $H : T^*M \to \mathbb{R}$. With the canonical symplectic structure $d\theta$ as above we can formulate the following

PROPOSITION 37. *The dynamics of the Lagrangian system* (M, L) *and the associated Hamiltonian system* $(T^*, d\theta, H)$ *are equivalent.*

PROOF. All we need to do is to check that statement locally. So let us pick any chart (U, φ) in M. We have seen before that the Euler-Lagrange equations in any coordinate are given by

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

and that they are equivalent to the equations

$$\dot{p} = -\frac{\partial H}{\partial q}$$
$$\dot{q} = \frac{\partial H}{\partial p}$$

which were called Hamilton's equations. $H(p,q) = p\dot{q} - L(q,\dot{q})$ with $p = \partial L/\partial \dot{q}$ as before, only now expressed in the coordinates at hand. That these are the equations for the Hamiltonian system is clear from Example **??**, since $d\theta = \sum_i dp^i \wedge dq^i$ with respect to the coordinates. $\square$

3.3. Poisson Algebras. Given a symplectic manifold (M,ω) we construct a Lie bracket $\{.,.\}$ on the vector space of smooth functions such that the linear map $f \mapsto X_f$ is a homomorphism between Lie algebras, i.e.

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

DEFINITION 4. *For any two smooth functions f,g on M we define their* Poisson bracket *through*

$$\{f,g\} = \omega(X_f, X_g).$$

PROPOSITION 38. *The Poisson bracket defines a Lie structure on the space of functions. That means it satisfies the Leibnitz rule and is antisymmetric:*

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0$$
$$\{f,g\} = -\{g,f\}.$$

Moreover, it is a derivation, i.e. it satisfies (another) Leibnitz rule

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

REMARK 39. *The two Leibnitz rules can be rewritten in the following way*

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

The common structure of the relation becomes much clearer.

DEFINITION 5. *An algebra A with a Lie structure $[.,.]$ which induces a homomorphism $v \in A \mapsto [v,.] \in Der(A)$ into the derivations of A is called a Poisson algebra.*

One central algebraic question in quantum physics is the following

PROBLEM 40. *Let* $(A, \{.,.\})$ *be a Poisson algebra. Is there a deformation* $*_\hbar$ *of the algebraic structure*

$$a, b \in A \mapsto a *_\hbar b$$

of the form

$$a *_0 b = ab$$
$$(a *_\hbar b - ab)/\hbar|_{\hbar=0} = \{a, b\}?$$

To what extent are they unique? What are its representations as bounded operators on Hilbert spaces?

EXERCISE 57. *Show that this problem is well-posed, namely, an arbitrary deformation of the multiplicaton of that form is associative up to first order in* $\hbar$.

EXERCISE 58 (Heisenberg's Uncertainty Principle). *Compute the commutator of such a deformation (up to first order in* $\hbar$*).*

DEFINITION 6. *A manifold* M *equipped with a (commutative) Poisson structure* $\{.,.\}$ *on its smooth functions is called a* Poisson manifold. *Given a smooth function* f *on* M *the vector field* X_f *which is uniquely defined by*

$$X_f(g) = \{f, g\}$$

for all smooth functions g*, is called* Hamiltonian vector field.

3.4. Least Action Principles of the Canonical Equations. We will present functionals whose critical points can be described as solutions of the canonical equations.

3.4.1. *Least Action in Phase Space.* There is a direct description of solutions of a canonical system (9) as the critical points of a functional. This turns out to be very useful in finding solutions of the system (see [**?**]). Let $H : T^*M \times [t_0, t_1] \to \mathbb{R}$ be a time-dependent Hamiltonian, and $\gamma = (\gamma(t), t) = (p(t), q(t), t)$ a solution of (9) with $(p(t_i), q(t_i)) = (p_i, q_i)$ for $i = 0, 1$.

PROPOSITION 41. *(1) The curve* γ *in the extended phase space is a critical point of the functional*

$$\mathcal{L}(\gamma) := \int_\gamma \theta - Hdt = \int_{t_0}^{t_1} p(t)dq(t) - H(p(t)q(t), t)dt.$$

(2) Vice versa, any critical point of the functional $\gamma' = \gamma'(s) = (p'(t), q'(t), \mathbf{t}'(t))$ *with* $(q'(s_i), \mathbf{t}'(s_i)) = (q_i, t_i)$ *is a solution of (9).*

PROOF. Notice that for a variation $\gamma + \delta\gamma$ of a curve in the extended phase space as described in (2) we have with Stokes

$$\mathcal{L}(\gamma + \delta\gamma) - \mathcal{L}(\gamma) = \int\int dp \wedge dq - dH \wedge dt,$$

since $pdq - Hdt$ vanishes along $(q, t) = const.$ Hence $i_{\dot\gamma}(dp \wedge dq - dH \wedge dt) = 0$ which gives (9). $\square$

REMARK 42. *(1) The classical form $(T^*L, d\theta, H)$ of the Hamiltonian system is not necessary for the formulation of the last statement. In fact, solutions $\gamma : [t_0, t_1] \to M$ of (9) for a Hamiltonian system (M, ω, H) are exactly the critical points of the following functional. Let $\Gamma'(s) = (\gamma'(s), \mathbf{t(s)})$ be a curve in $M \times \mathbb{R}$ with the same beginning and end points, $\gamma'(t_i) = \gamma(t_i), \mathbf{t(t_i)} = \mathbf{t_i}$. As above, the functional will not depend on the particularr parameterization of Γ'. hence for convenience we choose the same interval as for the original curve γ. Assume that there is a homotopy between γ and γ'. This can be extended to a homotopy $\Phi : [0,1] \times [t_0, t_1] \to M \times \mathbb{R}$ with $\Phi(0, .) = \Gamma$ and $\Phi(1, .) = \Gamma'$. Then*

$$\mathcal{L}(\Gamma') := \int \Phi^*(\omega - dH \wedge dt)$$

is independent of the chosen parameterization and its critical points are the solutions of (9).
(2) In the quest for solutions of (9) this is applied to loops in M (provided that H is periodic) and paths which start and end on given Lagrangians. In the latter case the variations one has to allow generalize the variations with $(q(s_i), \mathbf{t(s_i)}$ fixed as in the above propositon.

3.4.2. *Maupertuis' Principle.* There is a very cunning variational principle for the solutions of (9) in the classical setting $(T^*M, d\theta, H)$ for an autonomous which is a little awkward to state. In the end it provides us with the following description

PROPOSITION 43. *Let g be a Riemannian structure on a manifold M. Let $U = U(q)$ be a smooth function on M, the potential, and consider the Hamiltonian system $(T^*M, d\theta, H)$ with $H(p, q) = T(p, q) + U(q)$ and $T = \frac{|p|_g^2}{2}$ the kinetic energy. Then, in $\{(p, q) \mid U(q) < h\}$, the trajectories of this system with total energy h are geodesics with respect to the conformal metric*

$$g' = (h - U(q))g.$$

We formulate and prove the variational principle behind this claim.

PROPOSITION 44. *Let $(T^*M, d\theta, H)$ be an autonomous Hamiltonian system. Then the solutions γ of (9) with $q(t_0) = q_0$ and $q(t_1) = q_1$ with (constant) eergy $h = H(\gamma)$ are exactly the critical points of the functional*

$$\int_\gamma \theta$$

on curves in $\Sigma := H^{-1}(h)$ with these restrictions on their end points.

PROOF. This is a direct consequence of Proposition 41 since $dH|_\Sigma = 0$. □

Notice that for the variation of $\int pdq$ only the q–coordinate has to be fixed while for $\int pdq - Hdt$ the time of the end points is fixed as well. The freedom of choice of the time parameter is in fact essential in the principle. Let $H(p,q)$ be strongly convex in p. Let $\gamma : [t_0, t_1] \to \Sigma$. Its momentum is given by $p = \frac{\partial L}{\partial \dot q}$, where $L(q, \dot q) = pq - H(p,q)$ as before. The functional $\int pdq$ can thus be understood as a functional on (parameterized) curves in the configuration space. The solutions of the canonical system (9) with energy $H = h$ are the critical points of this functional restricted to curves with

$$H(\frac{\partial L}{\partial \dot q}, q) = h.$$

Notice that on curves $\gamma = \gamma(t)$ the functional takes the form

$$\int_\gamma \theta = \int_{t_0}^{t_1} p\dot q dt = \int_{t_0}^{t_1} \frac{\partial L}{\partial \dot q}\dot q dt.$$

Remember tha in the case $L = T - U$ with $U = U(q)$ a potential, we have

$$\frac{\partial L}{\partial \dot q}\dot q = T.$$

COROLLARY 45. *A massive point which is forced to stay on a smooth manifold moves on geodesics.*

PROOF. To guarantee $H = h$ one has obviously to choose the parameterization $\gamma(t)$ to be such that $\gamma^* dt = d\gamma/\sqrt{2h}$, where dt is the Lebesgue measure on $\mathbb{R}$ and $d\gamma$ the measure on γ induced by the (induced) Riemannian metric on γ. Hence the action of γ is equal to

$$\int \frac{\partial L}{\partial \dot q}\dot q dt = \int \sqrt{2h} d\gamma = \sqrt{2h}\int d\gamma = \sqrt{2h}\, l(\gamma)$$

the latter denoting the length of γ since $\frac{\partial L}{\partial \dot q}\dot q = 2T = 2H$. □

PROOF OF PROPOSITION 43. Here we have

$$L = T - U \quad H = L + U \quad \frac{\partial L}{\partial \dot q}\dot q = 2T = \left(\frac{d\gamma}{dt}\right)^2 = 2(h - U).$$

Therefore,

$$\gamma^* dt = \frac{d\gamma}{\sqrt{2(h-U)}}$$

with the same notations as before and the rest of the argument repeated. $\square$

4. Symmetries

We study the question how Noether's theorem on the symmetries and conservation laws generalizes to the general setting of Hamiltonian systems.

4.1. Hamiltonian Diffeomorphisms. Let Φ_t be a flow on a symplectic manifold (M, ω) preserving its symplectic structure $\Phi_t^*\omega = \omega$. This is not necessarily the flow of a Hamiltonian vector field. Indeed, consider the 2–torus $(T^2, \omega = d\varphi \wedge d\psi)$, where $(\varphi, \psi) \in \mathbb{R}^2/\mathbb{Z}^2$ are its coordinates. Let $\Phi_t(\varphi, \psi) = (\varphi + t, \psi)$. Then

$$X = \frac{d}{dt}\Phi_t = \frac{\partial}{\partial \varphi}$$

is the corresponding vector field and

$$i_X \omega = d\psi$$

is not exact: if $d\psi = df$ for some smooth function on T^2 we would get a contradiction by integrating over the loop $\gamma(t) = (0, t)$, $t \in [0, 1]$:

$$1 = \int_\gamma d\psi = \int_\gamma df = \int_0^1 f' dt = f(1) - f(0) = 0.$$

EXERCISE 59. *Show that in the above example Φ_t cannot even be a Hamiltonian diffeomorphism unless $t \in \mathbb{Z}$, i.e. $\Phi_t = \Phi_0 = Id$.* Hint: *Average the coefficients of an assumed time–dependent Hamiltonian vector field over time.*

EXERCISE 60. *Let Φ_t be a flow on a manifold L.*
(1) Show that the induced family Φ_t^* defines a flow on the cotangent bundle T^*L which preserves its Poincaré–Cartan form and is therefore symplectic.
(2) Show that this flow is Hamiltonian and find the corresponding function.

4.2. Noether's Theorem. In the light of the previous discussion Noether's theorem for Lagarngian systems should be generalized as follows

PROPOSITION 46. *Consider a Hamiltonian system (M, ω, H). Let $f : M \to \mathbb{R}$ be a smooth function on M which Poisson commutes with the Hamiltonian H:*

$$\{f, H\} = 0.$$

Then f is constant on the orbits of the Hamiltonian flow, i.e. for any solution $\gamma : [0,T] \to M$ of Hamilton's equations

$$\frac{d}{dt} f(\gamma(t)) = 0.$$

PROOF. This is now a very easy calculation:

$$\frac{d}{dt} f(\gamma(t)) = df(\dot{\gamma}(t)) = df(X_H) = \omega(X_f, X_H) = \{f, H\} = 0.$$

□

EXAMPLE 16. *Let $U = U(q)$ be a potential of a force which is symmetric under rotations about some axis Ω. Consider the corresponding Hamiltonian system $H(p,q) = |p|^2/2m + U(q)$ of a massive point of mass m moving in the field. Denote by Φ_t the rotations and Φ_t^* the corresponding family on $T^*\mathbb{R}^3 \cong \mathbb{R}^3 \times \mathbb{R}^3$ simply given by $\Phi_t^*(p,q) = (\Phi_t(p), \Phi_t(q))$. Let $X = d\Phi_t^*/dt$ be the vector field corresponding to it. It is not hard to see that this is the Hamiltonian flow with the corrresponding function given by $M = \Omega \cdot (p \times q)$.*

5. Rigidity of Symplectic Structures

We show that two phenomena in symplectic geometry are basically due to the same underlying principle: Exact deformations of symplectic forms lead to symplectomorphic structures. In particular, similar to hyperbolic geometry for instance, symplectic structures cannot be deformed to ones which are not symplectomoprhic.
Moreover, unlike Riemannian and very much like flat, elliptic or hyperbolic structures, symplectic structures look locally all alike.

5.1. Moser's Trick. Let $\{\omega_\tau \in \Omega^2(M)\}_{\tau \in I}$ be a smooth family of symplectic forms on a certain interval $0 \in I \subset \mathbb{R}$. We assume that there exists a smooth family of 1–forms $\alpha_\tau \in \Omega^1(U)$ on an open subset $U \subset M$ such that

$$d\alpha_\tau = \frac{\partial \omega_\tau}{\partial \tau}. \tag{12}$$

There exists an associated time–dependent vector field X_τ on M such that

$$\omega(X_\tau, .) = -\alpha_\tau. \tag{13}$$

Assume that for any $x \in U$ there exists a unique solutuion to the initial value problem

$$\dot{\gamma}_x(t) = X_t(\gamma(t)$$
$$\gamma_x(0) = x.$$

PROPOSITION 47 (Moser's Trick). *The family of diffeomorphisms $\Phi_\tau : U \to \Phi_\tau(U)$ defined through $\Phi_\tau(x) := \gamma_x(t)$ satisfies*

$$\Phi_\tau^* \omega_\tau = \omega_0.$$

on U.

PROOF. We use Cartan's formula for the Lie derivative in the following short computation.

$$\frac{\partial \Phi_\tau^* \omega_\tau}{\partial \tau} = \Phi_\tau^* \mathcal{L}_{X_\tau} \omega_\tau + \Phi_\tau^* \frac{\partial \omega_\tau}{\partial \tau} = \Phi_\tau^*(d(-\alpha_\tau) + d(\alpha_\tau))) = 0.$$

□

EXERCISE 61 (Moser's trick for volume forms). *Let ω_i be two volume forms on a manifold M without boundary which coincide outside a compact set and have the same volume:*

$$\int_M (\omega_1 - \omega_2) = 0.$$

Show that there is a diffeomorphism $\Phi : M \to M$ such that

$$\Phi^* \omega_2 = \omega_1.$$

REMARK 48. *To apply Moser's trick we have to establish two things:*

(1) *Find the smooth family of 1–form α_τ on M satisfying equation (12).*
(2) *Show that the corresponding family of vector fields can be uniquely integrated over a fixed interval.*

In the following we apply this idea to the problem of deformations of symplectic structures

PROPOSITION 49. *Let $\{\omega_\tau\}_{\tau \in [0,1]} \in \Omega^2(\mathbb{R}^{2n})$ be a smooth family of symplectic structures on $\mathbb{R}^{2n}$ which coincide outside a compact set. Then they are all symplectomorphic. In particular, there exists a family $\{\Phi_\tau\}_\tau$ of compactly supported diffeomorphisms of $\mathbb{R}^{2n}$ such that*

$$\Phi_\tau^* \omega_\tau = \omega_0.$$

PROOF. Let $\beta_\tau := \partial \omega_\tau / \partial \tau \in \Omega_0^2(\mathbb{R}^{2n})$ be the compactly supported family of 2–forms. Using the proof of Poincaré's Lemma we will construct a family of copactly supported 1–forms $\alpha_t au$ such that

$$d\alpha_\tau = \beta_\tau.$$

Let $F_s(x) = sx$ be the homotopy of the identity $\mathrm{Id}_{\mathbb{R}^{2n}} \sim 0$ to the trivial map. It is obvious that $\partial F_s \partial s(x) = x$. Notice that $d\beta_\tau = 0$, and hence

$$\partial F_s^* \beta_\tau / \partial s = \mathcal{L}_x \beta_\tau = d(i_x \beta_\tau).$$

Thus we have

$$\beta_\tau = \int_0^1 \frac{\partial F_s^* \beta_\tau}{\partial s} ds = \int_0^1 F_s^*(\mathcal{L}_x \beta_\tau) ds = \int_0^1 F_s^*(d(i_x \beta_\tau)) ds$$
$$= \int_0^1 d(F_s^* i_x \beta_\tau) ds = d\Big(\int_0^1 F_s^* i_x \beta_\tau ds \Big).$$

Therefore we find

$$\gamma_\tau := \int_0^1 F_s^* i_x \beta_\tau ds \in \Omega^1(\mathbb{R}^{2n})$$

safisfying

$$d\gamma_\tau = \beta_\tau.$$

It is obvious that $\gamma_\tau(x) = 0$. Moreover, $\mathcal{L}_x \gamma_\tau = 0$ outside some sufficiently large ball about 0, meaning that the restriction $\gamma_\tau|_{S_R} \in \Omega^1(S^{2n-1})$ does not depend on R, once R is sufficiently large. Here we use the identification of the sphere of radius R, S_R, with the standard sphere by dilation about 0. On the other hand $d(\gamma_\tau|_{S_R}) = 0$, at least for R sufficiently large, since β_τ was compactly supported.

EXERCISE 62. *Use Poincaré's Lemma on the northern and on the southern hemisphere of S^{2n-1}, to construct a smooth family of functions f_τ on the sphere, whose differential is $\gamma_\tau|_{S_R}$.*

Let ϕ be a compactly supported function on $\mathbb{R}^{2n}$ which is equal to 1 near the origin, we define

$$\alpha_\tau := \gamma_\tau - d(\phi f_\tau).$$

Obviously, $d\alpha_\tau = d\gamma_\tau$. By construction, the new 1–form has compact support. □

REMARK 50. *There is a version of the above statement to any smooth manifold. Herer one has to ask that the de-Rham-cohomology-classes with compact support, $[\omega_\tau]$, is constsnt in τ. This is for example satisfied if M is compact and the usual de-Rham-cohomology-class $[\omega_\tau]$ is constant. See* [**?**] *for a discussion.*

Proposition 49 raises an interesting question:

PROBLEM 51. *Are there any exotic symplectic structures on $\mathbb{R}^{2n}$ which coincide with the standard form outside a compact set? The proposition only states that there is no possibility of* deforming *the standard structure. This statement is obvious for $n = 1$. It is already a non-trivial strong application of Gromov's (pseudo)holomorphic curves for $n = 2$. It is wrong if one drops the requirement for ω to coincide with the standard structure outside a compact set: the counterexamples are constructed using once again pseudoholomorphic curves. It is an open problem for $n > 2$ until today!*

5.2. Darboux' Theorem. Moser's trick can also be exploited to construct canonical charts of a symplectic manifold. Let (M, ω) be any symplectic manifold of dimension $2n$, say. Let (U, φ) be any coordinate around a given point x. Let $\Omega := \varphi^*\omega \in \Omega^2(U)$ be the pull-back of the symplectic form. Let A be the linear transformation $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$, such that $A^*\Omega_0 = \omega_0$, that is, a symplectic isomorphism between (T_0U, Ω_0) and the standard symplectic vector space. Denote by $\omega_0|_U$ the standard symplectic structure on $\mathbb{R}^{2n}$. Hence the pull-back under A considered as a diffeomeorphism, $A^*\Omega$ coincides with the standard structure, ω_0, at 0. Therefore, the family of closed 2–forms

$$\omega_\tau := (1-\tau)\omega_0 + \tau A^*\Omega, \quad \tau \in [0,1]$$

is constant (equal to ω_0) at 0 and thus symplectic. We may find a, possibly smaller, neighborhood $U' \subset U$ of 0 such that $\omega_\tau|U' \in \Omega^2(U')$ is symplectic for all $\tau \in [0,1]$.

$$\frac{\partial \omega_\tau}{\partial \tau} = A^*\Omega - \omega_0.$$

Since the right–hand–side is closed, using Poincaré's lemma we may find a 1–form $\alpha \in \Omega^1(U')$, $\alpha_0 = 0$ and

$$d\alpha = A^*\Omega - \omega_0.$$

The corresponding family of vector fields X_τ, given by $\omega_\tau(X_\tau, .) := \alpha$ constantly vanishes at 0. Therefore, $\gamma(\tau) \equiv 0$ is the unique solution of the corresponding differential equation

$$\begin{aligned}\dot\gamma(\tau) &= X_\tau(\gamma(\tau))\\ \gamma(0) &= 0.\end{aligned}$$

Finally, as long as the initial value $x \in U'$ stays close to 0, in some neighborhood $U'' \subset U$ containing 0 say, the corresponding equation with initial value $\gamma(0) = x$ has a unique solution $\gamma : [0,1] \to U'$. Therefore, we have a well–defined family of difeomorphisms

$$\begin{aligned}\Phi_\tau &: U'' \longrightarrow U'\\ \frac{d}{d\tau}\Phi_\tau &= X_\tau \circ \Phi_\tau\\ \Phi_0 &= \mathrm{Id}_{U''}.\end{aligned}$$

By Moser's trick we deduce that $\Phi_\tau^*\omega_\tau = \omega_0$. Therefore, $\Phi_1^*(A^*(\varphi^*\omega)) = \omega_0$ and we proved the following

THEOREM 52. *For any point $x \in M$ in a $2n$–dimensional symplectic manifold $(M\omega)$ there exists a symplectic chart (V, ψ), that is a diffeomeorhism $\psi : U \subset \mathbb{R}^{2n} \to M$, $x \in \psi(V)$, such that*

$$\psi^*\omega = \omega_0.$$

PROOF. Following the construction above we set $V := U''$, $\psi := \varphi \circ A \circ \Phi_1$ which satisfies the claim of the theorem. □

Therefore, locally, symplectic manifolds of the same dimension look all alike. In particular, there is no infinitesimal invariant in symplectic, like curvature in Riemannian geometry. Moreover, Darboux' theorem says that there exists a ball $B^{2n}(0,r) \subset \mathbb{R}^{2n}$ which can be symplectically embedded into a given symplectic manifold (M,ω). It turns out that (as long as M is connected) the size, r, does not depend on the point $x \in M$ where the center is mapped. The quantity

$$G(M,\omega) := \sup\{r \mid \exists \varphi : B^{2n}(0,r) \hookrightarrow M, \varphi^*\omega = \omega_0\}$$

is therefore a positive invariant of the symplectic manifold. In the light of Problem 51 one may ask the stronger, and therefore simpler, question

PROBLEM 53. *Does there exist an exotic symplectic structure ω' on $\mathbb{R}^{2n}$ which is equal to the standard structure outside some compact set, whose Gromov–width is finite*

$$G(\mathbb{R}^{2n},\omega') < \infty = G(\mathbb{R}^{2n},\omega_0)?$$

5.3. Symplectic Neighborhoods. Darboux' theorem has a generalization to neighborhoods of certain (real) submanifolds in symplectic manifolds due to Weinstein.

PROPOSITION 54. *Let $N_i \subset M_i$, $i = 1,2$ be two closed submanifolds of the symplectic manifolds $(M_i\omega_i)$. Assume that there exists an isomorphism ψ of the symplectic bundles $\pi_i : (TM_i|_{N_i},\omega_i) \longrightarrow N_i$ which covers a diffeomeorphism $\varphi : N_1 \to N_2$:*

$$(\psi^*\omega_2)|_{T_xN_1} = (\varphi^*(\omega_2|_{N_1}))_x,$$

and with

$$\varphi^*(\omega_2|_{N_2}) = \omega_1|_{N_1}.$$

Then there exists neighborhoods $U_i \supset N_i$ and a symplectomorphism $\Psi : U_1 \to U_2$, $\Psi^*\omega_2 = \omega_1$ with

$$\Psi_{*,x} = \psi_x$$

for any $x \in N_1$ and

$$\Psi|_{N_1} = \varphi.$$

PROOF. The proof is not much different from the proof of Darboux's theorem. One first constructs a diffeomorphism between neighborhoods of N_i, whose differential satisfies the required relation. For instance, this is provided by the exponential map of any Riemannian structure applied to the normal bundle of the N_i. This is already (linearly) symplectic aong N_1. Secondly, one interpolates the two structures and obtains a family of symplectic strcutures in a small neighborhood of N_1. Third one applies Moser's trick again. $\square$

The subtle thing with this statement is to interpret it in particular situations.

5.3.1. *Lagrangian Submanifolds.* Let (V, ω) be a symplectic vector space. For a linear subspace $W \subset V$ denote by $Ann_\omega(W) := \{v \in V \mid \omega(v, w) = 0 \text{ for all } w \in W\}$ the *annulator of* W.

DEFINITION 7. *A linear subspace $W \subset V$ of a symplectic vector space is called*

isotropic : *if* $Ann_\omega(W) \supset W$,
coisotropic : *if* $Ann_\omega(W) \subset W$,
Lagrange : *if* $Ann_\omega(W) = W$.

EXERCISE 63. *Show that if* $\dim W = 2n$ *any isotropic subspace has dimension not bigger than n, since ω in non-degenerate, while any coisotropic subspace has dimension not smaller than n.*

DEFINITION 8. *An immersion $\iota : N \hookrightarrow M$ into a symplectic manifold is called* (co)isotropic *or* Lagrangian immersion *if for any $x \in N$, $\iota_*(T_xN) \subset T_{\iota(x)}M$ is a (co)isotropic or Lagrangian subspace of $(T_{\iota(x)}M, \omega_{\iota(x)})$.*

EXERCISE 64. *Let $\Phi : M_1, \to M_2$ be a diffeomorphism between the two symplectic manifolds (M_i, ω_i) of the same dimension. Show that Φ is symplectic, i.e. $\Phi^*\omega_2 = \omega_1$, if and only if its graph, $\Gamma_\Phi := \{(x, \Phi(x)) \mid x \in M_1\}$ is a Lagrangian submanifold of $M_1 \times M_2$ with respect to the symplectic structure $pr_{M_1}^*\omega_1 - pr_{M_2}^*\omega_2$.*

Another natural class of Lagrangian submanifolds is given in the cotangent bundle T^*L of a differentialbel manifold L with respect to the canonical symplectic structure $d\theta$ introduced above. Indeed, the zero section, O_L, and the fibre, T_x^*L for any $x \in L$, are both Lagrangian, since the Poincaré–Cartan form θ vanishes on either, as a simple calculation shows, and hence $d\theta$ since restriction and exterior derivative commute by naturality. More generally, we have

EXERCISE 65. *Let $\alpha \in \Omega^1(L)$ be a 1–form on L. Denote by $\Gamma_\alpha := \{\alpha_x \in T_x^*L \mid x \in L\} \subset T^*L$ its graph. Show that Γ_α is Lagrangian if and only if $d\alpha = 0$.*

Since the differential *df* of a smooth function satisfies this condition one could think of Lagrangians as objects which generalize smooth functions (up to a constant) and symplectic diffeomorphisms. The construction of convenient coordinates in Hamiltonian systems of the form $(T^*L, d\theta, H)$ using *generating functions* relies on that fact.

We have seen before that for any point $0_x \in O_x \subset T^*L$, with base point $x \in L$, in the zero section of the cotangent bundle the tangent space to it splits naturally into

$$T_{0_x}(T^*L) \cong T_x^*L \oplus T_xL.$$

The symplectic structure $d\theta$ is given by the natural symplectic structure

$$d\theta_{0_x}((\alpha, v), (\beta, w)) = \alpha(w) - \beta(v).$$

For $x \in L$ there is a short exact sequence of bundles

$$0 \longrightarrow TL \xrightarrow{\iota} TM|_L \xrightarrow{\omega} T^*L \longrightarrow 0.$$

The first homomorphism is the natural inclusion while the second is given by

$$v \in T_xM \mapsto \omega_x(v, .).$$

If there was a splitting $TM|_L = TL \oplus E$ into Lagrangian subbundles we would obtain an isomorphsim of the bundles

$$\mathrm{pr}_E\omega^{-1} \oplus \iota : T^*L \oplus TL \longrightarrow TM|_L.$$

This is a symplectic homomorphism if and only if E is a Lagrangian subbundle. In order to construct it we need the following

LEMMA 55 (Compatible almost complex structures). *For any symlectic manifold* (M^{2n}, ω) *there exists an* almost complex structure J, *i.e. an fibrewise isomorphism of the tangent bundle* TM *which satisfies* $J^2 = -Id$, *which is* compatible *to* ω*:*

$$\omega(.,.) := \omega(J., J.),$$

and which is tamed by ω*:*

$$\omega(v, Jv) > 0, \text{ for all } v \in TM.$$

It is obvious that such a structure can be used to define a splitting into Lagrangian subbundles through

$$TM_L = TL \oplus J(TL).$$

PROOF OF LEMMA 55. We notice that the compatibility condition is equivalent to require that the bilinear form defined through $\omega(., J.)$ is symmetric. It is non–degenerate, of course. Fix for instance any Riemannian metric g on M. This exists on any smooth manifold provided we put conditions into their definition which assuer the existence of a smooth partition of the unit. From the non-degeneracy of ω and g we conclude that there is a unique isomorphism $A : TM \to TM$ with

$$g(.,.) = \omega(., A.).$$

It is not hard to check that the dual with respect to g satisfies

$$A^* = -A.$$

Hence, $-A^2$ is selfdual with respect to g and positive definite. Therefore the squareroot of $-A^2$ exists and yields an invertible matrix. We set

$$J := (-A^2)^{-\frac{1}{2}}A.$$

EXERCISE 66. *Show that A and $\sqrt{-A^2}$ commute. Hint: Express the latter as a converging power series, for which each element commutes with A.*

Hence $J^2 = -Id$. Using in addition the self-duality of the square root with respect to g we conclude the compatibility with ω:

$$\begin{aligned}\omega(Jv, Jw) &= \omega((-A^2)^{-\frac{1}{2}}Av, (-A^2)^{-\frac{1}{2}}Aw) = g((-A^2)^{-\frac{1}{2}}Av, (-A^2)^{-\frac{1}{2}}w)\\ &= g(-A^{-1}v, w) = \omega(w, -v).\end{aligned}$$

Last point to verify is that ω tames J. But

$$\omega(v, JJw) = \omega(v, (-A^2)^{-\frac{1}{2}}Aw) = g(v, (-A^2)^{-\frac{1}{2}}w).$$

The matrix $(-A^2)^{-\frac{1}{2}}$ is g-symmetric and positive definite. Hence $\omega(., J.)$ defines a Riemannian metric and the statement follows from positive definiteness of such. $\square$

5.3.2. *Hypersurfaces of contact type.* Let $\Sigma \subset M$ be a real hypersurface in a symplectic manifold (M, ω). The restriction $\omega|_\Sigma \in \Omega^2(\Sigma)$ necassarily degenerates. That means that the kernel $Ker(\omega|_\Sigma) := \{v \in T_x\Sigma \mid \omega_x(v, w) = 0 \text{ for all } w \in T_x\Sigma\}$ is non-trivial for each $x \in \Sigma$. Moreover, since ω is non-degenerate, $Ker(\omega|_\Sigma)$ is exactly one-dimensional. Since any field of tangential lines is integrable, we obtain a 1–dimensional foliation on Σ which we call the *characteristic foliation* of Σ. These are the flow–lines of an autonomous Hamiltonian $H : M \to \mathbb{R}$. Indeed, let H be such that, Σ is a regular level set: $\Sigma = H^c$ and $dH_x \neq 0$ for any $x \in \Sigma$. Then $\omega(X_H, v) = -dH(v) = 0$ for any $v \in T\Sigma$. Thus $X_H \in Ker(\omega)$, and since $X_H \neq 0$ along Σ it generates the line field. Therefore, the points in the system move along the leaves of the foliation.

In order to get control over the behavior of a dynamical system, one often tries to find and analyze its periodic orbits. For an autonomous hamitonian system we have seen above that its orbits lie on level sets of the Hamiltonian function.

PROBLEM 56. *Let $\Sigma \subset M$ be a smooth hypersurface in a symplectic manifold. How many compact leaves has its characteristic foliation?*

In general, ther will be no compact leaves at all. However, there is a class of hypersurfaces which always seems have a compact leaf.

DEFINITION 9. *A hypersurface $\Sigma \subset M$ in a symplectic manifold is called of* contact type *if there exists a vector field X in a neighborhood of Σ which is*

- *everywhere transversal to Σ, in particular non-vanishing,*
- *the flow Φ_t of X conformally expands ω: $\Phi_t^*\omega = e^t\omega$, or, equivalently, the Lie derivative satisfies $\mathcal{L}_X\omega = \omega$.*

CONJECTURE 57 (A. Weinstein). *The characteristic foliation of any closed hypersurface of contact type in a symplectic manifold has at least one closed leaf.*

Two examples should motivate the restriction of the general quest for closed leaves.

EXAMPLE 17. *Consider the vector field of dilations of $\mathbb{R}^{2n}$ $X := \sum_k (p_k\partial/\partial p_k + q_k\partial/\partial q_k)$. It is not hard to see (using Cartan's formula) that the Lie derivative of the standard symplectic form $\omega = \sum_k dp_k \wedge dq_k$ satifies*

$$\mathcal{L}_X\omega = \omega.$$

Hence, for any smooth compact starshaped domain $\Omega \subset \mathbb{R}^{2n}$ its boundary $\partial\Omega$ is of contact type. It is an old result of Rabinowitz that such a hypersurface always admits at least one closed orbit.

EXAMPLE 18. *Let L be any (closed) manifold. Let $H : T^*L \to \mathbb{R}$ be a Hamiltonian which is star–shaped on each fibre, i.e. $H|_{T_q^*L}$ is strictly monotone along all rays starting at 0_q. For instance, Hamiltonians of the form $H = T + U$, where T is the kinetic energy with respect to some Riemannian metric and U a smooth potential, as considered before, are of this type. It is not hard to check that $\mathcal{L}_X d\theta = d\theta$ for $X = \sum p_k\partial/\partial p_k$ in local coordinates, and θ is the canonical Poincaré-cartan form. Hence the level sets $H^{-1}(E) \subset T^*L$ are of contact type, as long as the total energy $E > \min_{0_L}(H)$.*

Notice that the quest for closed leaves in the second example contains the one for closed geodesics – a holy grail in global Riemannian geometry.
Since ω is closed we find that

$$\mathcal{L}_X\omega = d(i_X\omega).$$

Hence the restriction of the 1–form, $\alpha := i_X\omega|_\Sigma$ has the following three remarkable properties. First,

$$\alpha \wedge (d\alpha)^{n-1} = \alpha \wedge (\omega^{n-1}|_\Sigma) = i_X(\omega^n)/n|_\Sigma$$

is a volume form, since X is transversal to Σ, and ω^n is a volume form on M. Secondly, the restriction of $d\alpha$ to the distribution $\xi := Ker(\alpha)$ defines a symplectic structure on each of the vector spaces. This is true, since both, X and $Ker(\omega|_\Sigma)$ annihilate the subspace ξ. Since it is of codimension 2, and ω is non–degenrate, they span the annihilator. Therefore,

$\omega|_\xi = d\alpha|_\xi$ is necessarily non-degenerate and thus symplectic. Third, ξ is *maximally non-integrable*. That means the maximal dimension of submanifolds in Σ which are everywhere tangent to ξ is $n-1$.

EXERCISE 67. *Show that these three proerties are equivalent.*

Such nowhere vanishing 1–forms on an odd–dimensional manifold are called *contact forms*, the distribution defined by such a form is called a *contact structure*

EXERCISE 68. *Let $\alpha \in \Omega^1(\Sigma)$ be a contact form on a smooth manifold of odd dimension. Show that for any positive smooth function f on Σ, $f\alpha$ is also a contact form.*

From the second condition follows that $d(e^t\alpha)$ is a symplectic form on the cylinnder $\Sigma \times \mathbb{R}$, were $t \in \mathbb{R}$ corresponds to the second factor in the product. Using the flow of the transverse vector field x it is not hard to show that a neighborhood of $\Sigma \subset M$ is symplectomorphic to $(\Sigma \times (-\epsilon, \epsilon), d(e^t\alpha))$, where Σ is mapped identically to $\Sigma \times \{0\}$.

EXERCISE 69. *Show that this statement follows almost trivially from the symplectic neighborhood theorem.*

5.3.3. *Further examples.* We will formulate the symplectic neighborhood theorem for closed isotropic and for symplectic submanifolds. Once again, Weinstein's Theorem will not immediately provide us with the desired form. To obtain this from the immediately obvious formulation one has to work a little bit.

EXERCISE 70. *Let $L \subset M$ be a closed isotropic submanifold of a symplectic manifold (M, ω). Choose a compatible almost complex structure J on M tamed by ω. Show that this induces a complex structure on the vector bundle of even rank $Ann(L)/TL$. On the other hand, let $E \to L$ be any complex vector bundle equipped with a hermitian metric. Let r be the corresponding real-valued norm–function on the total space E. Let $\alpha \in \Omega^1(E)$ be the 1–form associated to a hermitian connection. Consider the pull-back of this bundle and all its data under the projection $T^*L \to L$. Show that for a sufficiently small neighborhood U of the zero section $0_L \subset T^*L$ the closed 2-form $\omega_\alpha := d(r^2\alpha) + \pi_E^* d\theta$, where as before, θ refers to the Poincaré–Cartan form on T^*L. Show that there are sufficiently small neighborhoods of $L \subset M$ and the zero section of $0_E \subset \pi_L^*(E)$ which are symplectomorphic with respect to ω and ω_α. In particular, the symplectic structure near L only depends upon the isomorphism class of the complex bundle $Ann_\omega(L)/TL$.*

EXERCISE 71. *Let $C \subset M$ be a closed symplectic submanifold of $(M, \omega$, i.e. $\omega_C := \omega|_C$ is a symplectic form on C. Show that there exists a compatible and tamed almost complex structure on M such that C becomes J–holomorphic. Show that $Ann_\omega(C)$ thus becomes* *a J–complex bundle with a Hermitian structure defined by ω – the normal bundle of C. With r and α on a Hermitian structure as before we define the symplectic structure $\omega_\alpha = d(r^2\alpha) + \pi_E^*\omega_C$ on a complex vector bundle $E \to C$. Show that there are neighborhoods of $C \subset M$ and $0_E \subset E$ which are symplectomorphic with respect to that structure. In particular, there are neighborhoods of two closed symplectic submanifolds $C_i \subset M_i$ of symplectic manifolds (M_i, ω_i) if and only if $(C_i, \omega_i|_{C_i})$ are symplectomorphic and the complex bundles $Ann_{\omega_i}(C_i)$ are isomorphic.*

Remember that Moser's isotopy theorem for surfaces shows that two symplectic surfaces are symplectomorphic if and only if their volumes coincide. For symplectic surfaces in symplectic 4–manifolds the condition on the complex bundle (the normal bundle) simply means that the algebraic self-intersection number of the C_i must coincide. See [**?**] for both examples.

6. Huygen's Principle

The basic notions of Hamilton's description of classical dynamical systems originate from the generalization of Fermat's variational principle in optics to Hamilton's least action principle in classical mechanics. We will not go into detail here and the reader is referred to [**?**]. Instead we will show how to reformulate the canonical equations (a system of ordinary diffferential equations) as a partial differential equation of a real-valued time-dependent function on the configuration space – the so-called Hamilton–Jacobi equations. The solutions provide first integrals of the system, thus turning out to be useful to solve it. This is demonstrated at the end of this section.

6.1. The Action Function.

6.2. The Hamilton–Jacobi Equation.

6.3. Generating Functions. We will introduce a convenient representation for symplectomorphisms on $\mathbb{R}^{2n}$ – called generating functions. To illustrate its importance we will determine appropriate transformations of the Hamiltonian system of the gravitational field with two fixed centers of mass and show that this is completely integrable.
Let $\Phi = \Phi(p,q) = (P(p,q), Q(p,q))$ be a symplectomorphism of $\mathbb{R}^{2n}$. Then the canonical symplectic form $\omega = \sum_k dp_k \wedge dq_k$ is preserved:

$$\omega = \Phi^*\omega,$$

with $\Phi^*\omega = \sum_k dP_k \wedge dQ_k$. Hence, $d(pdq - PdQ) = 0$ and there exists a smooth function $S = S(p,q)$ such that

$$pdq - PdQ = dS.$$

Let us assume that (Q,q) are local coordinates in a neighborhood of a point (p_0,q_0), i.e. the Jacobian determinant of $(p,q)\mapsto(Q,q)$ does not vanish

$$0\neq \det(\frac{\partial(Q,q)}{\partial(p,q)} = \det\frac{\partial Q}{\partial p}.$$

In this case, S can be expressed as a function $S_1=S_1(Q,q)$:

$$S(p,q)=S_1(Q,q).$$

By the definition of S we have

$$\frac{\partial S_1(Q,q)}{\partial q}=p$$
$$\frac{\partial S_1(Q,q)}{\partial Q}=-P.$$

DEFINITION 10. *The function S_1 described above is called a* generating function *of the symplectomorphism Φ.*

Generating functions of the above kind can be characterized as follows:

PROPOSITION 58. *Let $S_1=S_1(Q,q)$ be a function in a neighborhood of $(Q_0,q_0)\in\mathbb{R}^n\times\mathbb{R}^n$. If S_1 satisfies*

$$\det(\frac{\partial^2 S_1}{\partial Q\partial q})\neq 0$$

then it is the generating function of a symplectomorphism.

PROOF. Using the above relation we consider the equation

$$\frac{\partial S_1(Q,q)}{\partial q}=p.$$

By the implicit function theorem and the assumption on S_1 we may solve this equation for $Q=Q(p,q)$ in a neighborhood of the point (q_0,p_0), where

$$p_0=\frac{\partial S_1(Q,q)}{\partial q}|_{Q_0,q_0}.$$

Finally we consider $P'(Q,q):=-\partial S_1(Q,q)/\partial Q$ and set $P(p,q):=P'(Q(p,q),q)$. The map

$$\Phi(p,q):=(P(p,q),Q(p,q))$$

is a symplectomorphism since

$$pdq-PdQ=\frac{\partial S_1(Q,q)}{\partial q}dq+\frac{\partial S_1(Q,q)}{\partial Q}dQ=dS_1(Q,q).$$

□

Not every symplectomorphism of $\mathbb{R}^{2n}$ can be generated by a function $S_1 = S_1(Q,q)$, e.g. the identity! The way out is to allow any set of coordinates $(P_{i_1},...,P_{i_k};Q_{j_1},...,Q_{j_l})$ for a partition $\{i_1,...,i_k,j_1,...,j_l\} = \{1,...,n\}$ which together with $\{q_1,...,q_n\}$ form a set of coordinates of $\mathbb{R}^{2n}$ in a neighborhood of (p_0,q_0). One can show that among the 2^n possible partitions there is one such choice. Since $pdq - PdQ = dS$ we find $pdq + \sum_r Q_{i_r}dP_{i_r} - \sum_s P_{J_s}dQ_{j_s} = d(\sum_r P_{i_r}Q_{i_r} + S)$. Hence with

$$S_3(P_i,Q_j,q) := \sum_r P_{i_r}Q_{i_r} + S(p,q)$$

we compute $p = \partial S_3/\partial q, Q_{i_r} = \partial S_3/\partial P_{i_r}, P_{j_s} = \partial S_3/\partial Q_{j_s}$. Vice versa, any given partition and function $S_3 = S_3(R,q)$ of $2n$ variables and $\det(\partial^2 S_3/\partial R\partial q) \neq 0$ defines a symplectomorphism analogous to what we did before.

Exercise 72. *Compute the symplectomorphism for the generating function $S_3 = Pq$.*

7. Hamilton–Jacobi Equations

Given a Hamiltonian system $(\mathbb{R}^{2n}, \omega_0, H = H(p,q))$ we would like to find new symplectic coordinates (P,Q), such that $H = H(Q)$ depends on Q only. Then Hamilton's equations take the form $\dot{Q} = 0$, $\dot{P} = \partial H/\partial Q$ and we find solutions $Q(t) = Q_0$ and $P(t) = P_0 + t\partial H/\partial Q|_{Q(0)}$. To find such a transformation we assume that it is given by a generating function $S = S(Q,q)$. Since $p = \partial S/\partial q$ it has to satisfy

$$H(\frac{\partial S(Q,q)}{\partial q}, q) = H(Q).$$

This is (a special form) of the Hamilton–Jacobi equations. Notice that it is a partial differential equation, whose solutions provide first integrals of the ordinary Hamilton equations.

Proposition 59. *If a family of solutions $S(Q,q)$ of the Hamilton–Jacobi equations exists, which satisfies $\det(\frac{\partial^2 S}{\partial Q, \partial q}) \neq 0$, one can explicitely solve Hamilton's equations. The functions $Q(p,q)$ are are first integrals.*

The statement "explicitly" is a slight lie. In fact the solution can be expressed via $\Phi^{-1}(Q_0, P_0 + t\partial H/\partial Q|_{Q(0)})$. We will see in the illustrating example below that it is a computationally non-trivial task to determine S and Φ^{-1}. The first integrals are of more significance here. Namely, $Q_1,...,Q_n$ all Poisson commute and define a maximal system of such functions. Such Hamiltonian systems are called *integrable.*

7.1. Elliptic Coordinates for Two Centers of Mass. We study the motion of a planet in a gravitational field of two suns of the same mass. To make life considerably easier we restrict ourselves to a situation where the centers of the two suns and the planet span a plane which is constant in time. There we introduce confocal coordinates for the position of the planet: if r_1, r_2 are the two distances from the center we denote by

$$\xi = r_1 + r_2, \quad \eta = r_1 - r_2.$$

The level sets of ξ are ellipses, while those of η are hyperbolas. Notice that they are perpendicular to each other. Indeed, denote by ρ_i the unit radial vector fields with respect to the two centers. Then $\rho_1 \pm \rho_2$ are the gradient vectors of ξ and η, respectively. Hence they are are orthogonal to the levels. It remains to check their scalar product vanishes:

$$(\rho_1 + \rho_2) \cdot (\rho_1 - \rho_2) = 0.$$

Hence the metric tensor on the cotangent vectors expressed in coordinates (ξ, η) is diagonal.
On the other hand, the lengths, of these gradients can be readily computed:

$$|\rho_1 \pm \rho_2|^2 = 2 \pm 2\rho_1 \cdot \rho_2.$$

The latter is equal to $\cos 2\alpha$ where 2α is the angle at the planet of the triangle it forms with the two centers. Hence

$$\begin{aligned} |\rho_1 + \rho_2|^2 &= 4\cos^2 \alpha \\ |\rho_1 - \rho_2|^2 &= 4\sin^2 \alpha. \end{aligned}$$

These are exactly the coefficients of the tensor.
If $2c$ is the distance of the two suns we find

$$4c^2 = r_1^2 + r_2^2 - 2r_1r_2\cos 2\alpha.$$

From that we deduce with a little bit of trigonometry expressions

$$\begin{aligned} \cos^2 \alpha &= \frac{(r_1 + r_2)^2 - 4c^2}{4r_1r_2} \\ \sin^2 \alpha &= \frac{4c^2 - (r_1 - r_2)^2}{4r_1r_2}. \end{aligned}$$

Hence the Hamiltonian of the system is expressed as

$$\begin{aligned} H &= p_\xi^2 \frac{(r_1 + r_2)^2 - 4c^2}{2r_1r_2} + p_\eta^2 \frac{4c^2 - (r_1 - r_2)^2}{2r_1r_2} - \frac{k}{r_1} - \frac{k}{r_2} \\ &= 2p_\xi^2 \frac{\xi^2 - 4c^2}{\xi^2 - \eta^2} + 2p_\eta^2 \frac{4c^2 - \eta^2}{\xi^2 - \eta^2} - \frac{4k\xi}{\xi^2 - \eta^2}. \end{aligned}$$

Multiplying by $(\xi^2 - \eta^2)$ the Hamilton–Jacobi equation takes the form

$$(\frac{\partial S}{\partial \xi})(\xi^2 - 4c^2) + (\frac{\partial S}{\partial \eta})(4c^2 - \eta^2) = K(\xi^2 - \eta^2) + 4k\xi.$$

This can be easily separated into

$$(\frac{\partial S}{\partial \xi})(\xi^2 - 4c^2) - 4k\xi - K\xi^2 = L$$

$$(\frac{\partial S}{\partial \eta})(4c^2 - \eta^2) + K\eta^2 = -L.$$

We arrive at a formula for S:

$$S(\xi, \eta; L, K) := \int \sqrt{\frac{L + K\xi^2 + 4k\xi}{\xi^2 - 4c^2}} d\xi + \int \sqrt{\frac{-L - K\eta^2}{4c^2 - \eta^2}} d\eta.$$

EXERCISE 73. *Express solutions of the Hamiltonian system with the help of the just found generating function S in confocal coordinates (ξ, η).*