Symplectic Geometry and Integrable Systems (MATH 538-003) Lecture notes

Anton Izosimov

Last updated: May 2, 2020

Contents

1	Introduction to planar billiards	3
2	Invariant area form of the billiard map and Poincaré's recurrence theorem	5
3	Billiards in disks	9
4	Billiards in ellipses	10
5	Geometric meaning of the Joachimsthal integral	13
6	The phase portrait of the billiard in an ellipse	16
7	Liouville integrable maps in 2D	18
8	Proof of the Arnold-Liouville theorem in 2D and Poncelet's closure theorem	20
9	Algebraic integrability of elliptic billiards I	22
10	Algebraic integrability of elliptic billiards II	24
11	Canonical Hamilton's equations	28
12	Symplectic vector spaces and symplectic manifolds	29
13	Vector fields and flows I	32
14	Vector fields and flows II	34
15	Hamiltonian vector fields	37
16	Poisson brackets	42
17	Liouville integrable systems on symplectic manifolds	44
18	Proof of the first part of the Arnold-Liouville theorem: topology	46

19 Proof of the second part of the Arnold-Liouville theorem: dynamics	48
20 Noether's principle, first examples of integrable systems	49
21 Further examples of integrable systems	53
22 Introduction to Poisson manifolds	53
23 Symplectic leaves	55
24 Symplectic leaves and Casimir functions	57
25 Lie groups and Lie algebras	59
26 Lie-Poisson brackets and their symplectic leaves	62
27 Lie-Poisson brackets and Casimirs	65
28 Lie-Poisson brackets from Poisson reduction	69
29 Reduction of Hamiltonian dynamics and rigid body motion	71
30 Free rigid body dynamics	74
31 Integrable cases of heavy rigid body motion	76
32 The Toda lattice	78
33 The Adler-Kostant-Symes scheme	80
34 Factorization theorem	84
35 Asymptotic behavior of the Toda lattice	86

Lecture 1: Introduction to planar billiards

Let $D \subset \mathbb{R}^2$ be a compact planar domain bounded by a smooth curve ∂D . The dynamics of a billiard ball (which we assume to be a point mass) in D is defined as follows: the ball is moving inside D along a straight line with constant velocity, until it hits the boundary. Upon hitting the boundary, the ball is reflected from it according to the law the angle of reflection is equal to the angle of incidence, after which it continues moving along a straight line with constant velocity until it hits the boundary again. Then it is again reflected according to the same law, and so on. See Figure 1.

We would like to understand the dynamics of the ball, i.e. find, if possible, its position and velocity as a function of time, as well of the initial position and velocity. To begin with, we need to describe the *phase space* of the ball, that is the space of all its possible positions and velocities. There are two possible approaches to defining this space for the billiard system. The first one is to consider all possible positions, including those strictly inside the domain D. However, in this approach, the dynamics is trivial on most of the phase space, because when the ball is strictly inside D, it is just moving along a straight line. The second approach is to only look at the positions of the ball at the boundary ∂D . Knowing the current position of the ball at the boundary, and also its current velocity, we can easily recover the subsequent dynamics of the ball, at least until it hits the boundary next time. So, we can only trace subsequent positions of the ball at the boundary, while all other positions are easily recovered. Also notice that since the magnitude of the velocity vector is preserved and does not affect the dynamics at all, we can assume that the velocity is given by a unit vector. Finally, notice that there is two kind of possible velocity vectors of the ball at the boundary, and has not yet been reflected, while inward velocities are velocities after reflection.

Remark 1.1. Let $\gamma(t)$ be a parametrization of ∂D , corresponding to its counter-clockwise orientation. We say that a vector v at $x = \gamma(t_0)$ is *inward* if the orientation of the frame $(\gamma'(t_0), v)$ is positive, and *outward* if the orientation of the frame $(\gamma'(t_0), v)$ is negative, see Figure 2. Also note that a velocity vector which is neither inward, nor outward, must be tangent to the boundary. Such velocities are only possible for billiards in non-convex domains, see Figure 3.

Since inward velocities and outward velocities are connected by the reflection law, it suffices to consider just one type of velocity vectors, for example the inward ones. So, we define the *phase space* of a billiard ball in D as the space of inward unit vectors attached at points of ∂D . In other words, an element of the phase space is a pair (x, v), where $x \in \partial D$, and $v \in T_x D$ is an inward tangent vector. Note that since ∂D is homeomorphic to a circle, while an inward tangent vector v is characterized by the angle $\alpha \in (0, \pi)$ it makes with the positive direction of ∂D (see Figure 4), our phase space can be viewed as the cylinder $S^1 \times (0, \pi)$. As coordinates on this cylinder, we can take the angle $\alpha \in (0, \pi)$, as well as any coordinate on $S^1 = \partial D$. In what follows, as coordinate on ∂D we will use the arc-length parameter t. So, the phase



Figure 1: Trajectory of a billiard ball in a planar domain D.



Figure 2: v is an inward velocity vector, and w is an outward velocity vector.



Figure 3: In a non-convex domain D, the velocity of a billiard ball may be tangent to the boundary.

space of the billiard in D is a cylinder with coordinates (t, α) , where t is a real number defined up to an integer multiple of the total length L of ∂D , and $\alpha \in (0, \pi)$.

Let M be the phase cylinder of the billiard in D. Define the *billiard map* $T: M \to M$ as follows: for an outward velocity vector v at a boundary point $x \in \partial D$, its image under T is the velocity of the billiard ball with initial velocity v immediately after it is reflected from the boundary for the first time, see Figure 5. Then the iterate T^k computes the velocity of the ball after it hits the boundary for the k'th time (see Figure 6). So, the dynamics of the billiard ball is pretty much determined by the behavior of the billiard map $T: M \to M$ and its iterates. For this reason, most (if not all) people studying billiards study properties of this map.

The first thing we would like to understand about this map T is how regular it is. This very much depends on the properties of the domain D and its boundary ∂D . In particular, if D is not convex, then T does not even have to be continuous. For example, in Figure 7, initially close vectors v and w are mapped by T to vectors that are far away from each other.

For convex domains with smooth boundary, a routine application of the implicit mapping theorem show that T is continuous, and, moreover, a diffeomorphism of the *open* phase cylinder $M = \partial D \times (0, \pi)$



Figure 4: An element of the phase space for the billiard in D is a point $x \in D$ and an inward tangent vector at $v \in T_x D$. The latter vector can be characterized by the angle $\alpha \in (0, \pi)$ it makes with the positive direction of ∂D .



Figure 5: Definition of the billiard map T.



Figure 6: Itertaions of the billiard map T.

to itself. If, moreover, the boundary of D has strictly positive curvature, then we have the following result:

Theorem 1.2 (see e.g. [9]). If the boundary of the domain D is smooth and has strictly positive curvature, then the corresponding billiard map $T: M \to M$ extends to a diffeomorphism of the closed phase cylinder $\partial D \times [0, \pi]$ to itself.

In other words, for sufficiently nice boundaries, the billiard map stays well-defined and smooth when the initial velocity vector tends to a vector parallel to the boundary.

Lecture 2: Invariant area form of the billiard map and Poincaré's recurrence theorem

Recall that the phase space M of a billiard in a domain D is the space of pairs (x, v), where $x \in \partial D$ is a point at the boundary of D, and $v \in T_x D$ is an inward unit tangent vector at x. Since ∂D is homeomorphic to a circle, and the space of unit inward tangent vectors at each point is homeomorphic to an open interval, the phase space M is topologically a cylinder. As coordinates on M, we take the arc



Figure 7: In a non-convex domain, the billiard map may be discontinuous.



Figure 8: The partial derivatives of the function $||\gamma(t') - \gamma(t)||$ are given by $\cos(\alpha')$ and $\cos(\alpha)$.

length parameter t on ∂D , and the angle α which the inward vector v makes with the positive direction of ∂D . Note that the coordinate t is defined modulo the total length L of ∂D , while α is well-defined real number in $(0, \pi)$. Consider the following 2-form on M:

$$\omega = \sin(\alpha) d\alpha \wedge dt.$$

Note that even though t is not quite a function (it is defined up to an additive constant L, so it is in a fact a map to a circle), its differential dt is a well-defined 1-form. So, the 2-form ω is well-defined as well. Furthermore, we have $d\alpha \wedge dt \neq 0$, because α and t are, by construction, coordinates on M. Finally, notice that $\sin(\alpha) \neq 0$, since $\alpha \in (0, \pi)$. So, ω is a non-vanishing 2-form on a 2-manifold M, i.e. it is an *area form* (a particular case of a *volume form*, that is a non-vanishing *n*-form on an *n*-dimensional manifold).

Theorem 2.1. The area form ω on M is preserved by the billiard map $T: T^*\omega = \omega$. In other words, for any domain $U \subset M$, we have

$$\int_U \omega = \int_{T(U)} \omega$$

provided that at least one of these integrals is well-defined.

The proof is based on the following lemma:

Lemma 2.2. Let γ be an arc length parametrized curve in \mathbb{R}^2 . Let also $||\gamma(t') - \gamma(t)||$ be the Euclidian distance between the points $\gamma(t)$ and $\gamma(t')$. Finally, let α and α' be the angles between the chord joining $\gamma(t)$ and $\gamma(t')$ and the arc of γ connecting those two points (see Figure 8). Then

$$\frac{\partial}{\partial t}||\gamma(t') - \gamma(t)|| = -\cos(\alpha), \quad \frac{\partial}{\partial t'}||\gamma(t') - \gamma(t)|| = \cos(\alpha'). \tag{1}$$

Proof of the lemma. The expression $||\gamma(t') - \gamma(t)||$ is a symmetric function in t and t', so its partial derivatives with respect to these variables should be the same, up to replacing t with t'. The reason for different signs of the partial derivatives in (1) is the asymmetric definition of the angles α and α' : α is the angle between the chord $\gamma(t)\gamma(t')$ and the positive direction of γ , while α' is the angle between the same chord and the negative direction of γ (see Figure 8). So, it suffices to establish the second of formulas (1). The first one then follows by symmetry.

To prove second of formulas (1), fix t and consider the function $g(x) = ||x - \gamma(t)||$, where $x \in \mathbb{R}^2$. The level set of this function through the point $\gamma(t')$ is a circle centered at $\gamma(t)$. Therefore, the gradient of g(x) at $\gamma(t')$ is orthogonal to that circle and hence collinear to the vector $\gamma(t') - \gamma(t)$, see Figure 9 (this can also be seen from the definition of the gradient as the direction of fastest increase). Furthermore, the derivative of the function $g(x) = ||x - \gamma(t)||$ at $x = \gamma(t')$ in the direction $\gamma(t') - \gamma(t)$ is clearly equal to

¹Here we distinguish between the derivative in the direction v, and the derivative along v. By definition, the derivative in the direction v is the derivative along v/||v||.

1, so the gradient of g(x) at $\gamma(t')$ is a unit vector positively collinear to $\gamma(t') - \gamma(t)$ (here we use that the magnitude of the gradient is equal to the derivative of the function in the direction of the gradient, and also that a function must increase, not decrease in the direction of its gradient).

Now we have that $||\gamma(t') - \gamma(t)|| = g(\gamma(t'))$, so the partial derivative of $||\gamma(t') - \gamma(t)||$ with respect to t' is equal to the derivative of the function g(x) along the velocity vector $\gamma'(t')$ of the curve γ at the point $\gamma(t')$. Therefore, we have

$$\frac{\partial}{\partial t'} ||\gamma(t') - \gamma(t)|| = \langle \operatorname{grad} g(\gamma(t')), \gamma'(t') \rangle,$$

where grad $g(\gamma(t'))$ is the gradient of g(x) at $x = \gamma(t')$. Furthermore, we already saw that grad $g(\gamma(t'))$ is a unit vector, while $\gamma'(t')$ is a unit vector because the curve γ is arc length parametrized. So, the inner product between these two vectors is equal to the cosine of the angle between them. To complete the proof, it suffices to notice that this angle is precisely α' (see Figure 9).

Proof of Theorem 2.1. Let (t', α') be the coordinates of the image of the point (t, α) under the billiard map T. Then t', α' are smooth functions of t, α , and hence smooth functions on M. Let also γ be the arc length parametrized boundary of the billiard table D. Then α' can be defined as the angle between the chord $\gamma(t)\gamma(t')$ and the negative direction of the curve γ , see Figure 10. Now, consider the function $f = ||\gamma(t') - \gamma(t)||$. Since t and t' are smooth functions on M, so is f. Furthermore, by Lemma 2.2, we have

$$df = \cos(\alpha')dt' - \cos(\alpha)dt.$$
⁽²⁾

Consider now the 1-form ξ on M given by

$$\xi = -\cos(\alpha)dt.$$

Notice that

$$d\xi = d(-\cos(\alpha)) \wedge dt = \sin(\alpha)d\alpha \wedge dt = \omega$$

Furthermore, by (2) we have

$$df = \xi - T^* \xi. \tag{3}$$

Any f satisfying this equation for some ξ such that $d\xi = \omega$ is called a *generating function* of the map T. A map possessing a generating function is automatically area-preserving. Indeed, taking the differential of both sides in (3), we get

$$0 = d\xi - dT^*\xi = d\xi - T^*d\xi = \omega - T^*\omega.$$

 \mathbf{SO}

$$T^*\omega = \omega,$$

as desired. Thus, the billiard map T has a generating function given by $||\gamma(t') - \gamma(t)||$ and hence is area-preserving, q.e.d.

Figure 9: The gradient of the function $g(x) = ||x - \gamma(t)||$ at the point $x = \gamma(t')$.





Figure 10: The billiard map T takes the point (t, α) to the point (t', α') .

Theorem 2.3 (Poincaré's recurrence theorem for billiards). The billiard map $T: M \to M$ has the following recurrence property: for any open subset $U \subset M$, almost all points in U will eventually return, under the action of iterations of T, to the set U. In other words, for almost any $v \in U$ there exists a natural number n > 0 such that $T^n(v) \in U$.

Remark 2.4. As one can see from the proof, this theorem is true for any volume-preserving map of an *n*-dimensional manifold, provided that the total volume of the manifold is finite. Furthermore, the same result holds for any measure-preserving map of any measure space, again provided that the measure of the whole space is finite. This is the way this theorem is usually stated.

Proof of Theorem 2.3. Let $V \subset U$ be the set of points which do not return to U under the iterations of T. This means that V consists of points $v \in U$ such that $T^n(v) \notin U$ for any natural number n > 0. We need to show that the set V has area 0. To that end, notice that for any natural n > 0, the set $T^n(V)$ is disjoint from V. Indeed, by definition of V, the set $T^n(V)$ does not intersect U and hence V. So, we have that $T^n(V) \cap V$ is an empty set. Furthermore, since T is invertible, it follows that for any integer $l \ge 0$ the set $T^{n+l}(V) \cap T^l(V)$ is also empty (if $v \in T^{n+l}(V) \cap T^l(V)$, then it must be that $T^{-l}(v) \in T^n(V) \cap V$, which is impossible). And since n > 0 and $l \ge 0$ are arbitrary, it follows that the sets

$$V, T(V), T^2(V), \dots \tag{4}$$

are all pairwise disjoint. So, their total area cannot exceed the total area of the phase cylinder M, which is

$$\int_0^L \int_0^\pi \sin(\alpha) d\alpha dt = 2L,$$

i.e. twice the perimeter of the billiard table. At the same time, since T is area-preserving, all the sets (4) have the same area and thus may have finite total area only if each of them has area zero. Thus, the theorem is proved.

Remark 2.5. There is one step that we skipped in the proof: one actually needs to explain why the set V has well-defined area at all. This can be done by showing that V is a difference of two open sets and hence measurable.

Example 2.6. Consider a billiard table D of an arbitrary shape, and let $I \subset \partial D$ be a (possibly very small) open subset of the boundary. Consider billiard trajectories that start at points of I and make an angle with ∂D which is between 89° and 91°. Then, according to Poincaré's theorem (applied to the subset $U = I \times (89^\circ, 91^\circ) \subset M$), almost all of these trajectories will eventually hit I again, and, moreover, the angle of incidence will again be between 89° and 91°. Of course, it may (and, generally speaking, will) take a long long time for this event to occur: Poincaré's recurrence theorem is saying that for almost all

initial data v in our open subset U, there is n > 0 such that $T^n(v) \in U$, but it is saying nothing about what the value of n is.

Lecture 3: Billiards in disks

We now turn to studying concrete examples of billiards. We will start with the simplest case of billiards in disks. Clearly, it is sufficient to consider the case when D is a unit disk: billiards in disks of different radii behave in the same way. Let A be a point in the unit circle. Consider the billiard trajectory in the unit disk which starts at A and makes an angle α with the positively oriented unit circle (see Figure 11). Then this trajectory will meet the unit circle again at a point B, and the angle of incidence is also equal to α (because a chord intersects a circle at the same angle at both intersection points). Therefore, after the trajectory is reflected at B, it will again make an angle α with the positive direction of the unit circle. This means that billiard map T does not affect the α -coordinate on the phase cylinder M at all: $\alpha' = \alpha$. Further, denote the *t*-coordinate of A by t, and the *t*-coordinate of B by t'. Then the increment t' - t is equal to the measure of the arc of the circle that goes from A to B (in the counter-clockwise direction). This measure is equal to the angle $\angle AOB$ (where O is the center of the circle), which, in turn, is equal to 2α (by the tangent-chord theorem combined with the inscribed angle theorem). Therefore, in coordinates, the billiard map T in the unit disk is given by

$$\begin{cases} t' = t + 2\alpha \\ \alpha' = \alpha, \end{cases}$$

where addition in the formula for t' is understood modulo the total length of the unit circle, which is 2π . This is a linear map (at least if we forget about the periodicity of the *t*-coordinate), and the corresponding dynamics (i.e. the behavior of iterates of this map) is easy to understand. Namely, we have the following explicit formulas for the iterate T^n :

$$\begin{cases} t' = t + 2n\alpha, \\ \alpha' = \alpha, \end{cases}$$

where, again, addition in the formula for t' is understood modulo 2π . We, however, would like to understand what the dynamics looks like qualitatively. To that end, notice that since α is preserved by the map, we can just fix it and study the dynamics, of the t variable, which is given by

$$t \mapsto t + 2\alpha. \tag{5}$$



Figure 11: Billiard in a disk.



Figure 12: Periodic billiard trajectories in a disk.

Recall that t can be thought of a as a point in a circle, $t \in \mathbb{R}/2\pi\mathbb{Z}$. The map (5) is then a rotation of that circle by angle 2α . This rotation behaves differently depending on whether the number α/π is rational or irrational. If $\alpha/\pi = m/n$ is rational, then $2n\alpha = 2m\pi$, and thus the *n*'th power of the map (5) is the identity map. So, for $\alpha/\pi \in \mathbb{Q}$, the map (5) is periodic. In terms of the billiard, this means that any trajectory with such α eventually closes up. Figure 12 shows two 5-periodic trajectories corresponding to $\alpha = \pi/5$ and $\alpha = 2\pi/5$.

Now, consider the case $\alpha/\pi \notin \mathbb{Q}$. In this case, we can still approximate α/π with a rational, with arbitrary precision (the denominator of that rational may be large, though). So, even though the dynamics in this case is not periodic, it is close to periodic. Such dynamics is known as *almost periodic* or *quasiperiodic* (the periodic case can also be considered a particular case of the quasi-periodic one). Note that in this case $t + 2n\alpha \neq t + 2m\alpha$ modulo 2π unless m = n, which means that the orbit of any $t \in S^1$ under iterations of the map (5) is infinite. One can in fact quite easily show that any orbit is dense in the circle. Figure 13 shows what a piece of a non-periodic trajectory in a disk, corresponding to $\alpha/\pi \notin \mathbb{Q}$, looks like. We in fact generated this as a periodic trajectory with large period. In practice, one does not really see the difference between quasiperiodic dynamics and periodic dynamics with large period.

We conclude this discussion by noticing that Poincaré recurrence clearly holds for the billiard in a disk (as it should by Theorem 2.3). Moreover, since the dynamics is quasi-periodic, every point eventually returns to any its small neighborhood.

Lecture 4: Billiards in ellipses

We now turn our attention to billiards in ellipses. Most of the results we will obtain also apply to hyperbolas. The case of a hyperbola is, however, somewhat more sophisticated, because, in contrast to ellipses, hyperbolas do not bound compact domains. So we will stick with ellipses. That being said,



Figure 13: A non-periodic billiard trajectory in a disk.

hyperbolas will still be appearing from time to time in our study.

Consider an ellipse given in a Cartesian coordinate system (x_1, x_2) by

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1$$

As known from linear algebra, any ellipse can be written in this form in a suitable orthogonal coordinate system.

Given a point (x, v) in the phase cylinder M for the billiard in the ellipse, consider the function

$$J = -\left(\frac{x_1v_1}{a_1^2} + \frac{x_2v_2}{a_2^2}\right),\tag{6}$$

where (x_1, x_2) are components of x, and (v_1, v_2) are components of v.

Proposition 4.1. The function J is invariant under the billiard map $T: M \to M$, i.e. J(T(x,v)) = J(x,v).

Remark 4.2. A function invariant under a map is called a *conserved quantity*, a *first integral*, or just an *integral* of the map. This particular function J is known as the *Joachimsthal integral*. Note that this integral can also be rewritten as

$$J = -\frac{1}{2} \langle v, \operatorname{grad} f(x) \rangle,$$

where

$$f = \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2}$$

is the defining function of the ellipse. This explains the negative sign in the definition: since grad f(x) is an outward normal to the ellipse, and v is an inward vector, the angle between grad f(x) and v must be obtuse, and $\langle v, \text{grad } f(x) \rangle < 0$. Therefore, our function J is always positive. Of course, the function

$$-J = \frac{x_1 v_1}{a_1^2} + \frac{x_2 v_2}{a_2^2}$$

is also preserved by the billiard map, it is just somewhat more pleasant to work with positive functions than with negative ones.

Remark 4.3. Also notice that in the case of a unit circle the norm of the vector grad f(x) is equal to 2, so

$$J = -\frac{1}{2} \langle v, \operatorname{grad} f(x) \rangle = -\cos(\alpha + \pi/2) = \sin(\alpha)$$

where we also used that v is a unit vector, and that its angle with the outward normal grad f(x) is equal to $\alpha + \pi/2$, with α being the angle between v and the positive direction of the unit circle. So, in the circle case the conserved quantity J essentially coincides with the conserved quantity α which we found in the previous lecture (more precicely, $J = \sin(\alpha)$, but saying that α is preserved is more or less the same as saying that $\sin(\alpha)$ is preserved). So, we already know that Proposition 4.1 is true for the circle.

Proof of Proposition 4.1. Let $(x, v) \in M$ be a point in the phase cylinder, and let (x', v') = T(x, v) be its image under the billiard map. Then, by definition of the billiard map, the vectors v and v' make the same angle with the (tangent line to the) ellipse at x'. Since those are unit vectors, this is equivalent to saying that the vector v + v' is tangent to the ellipse at x', see Figure 14. Using also that the gradient grad f(x') is orthogonal to the ellipse at x', this gives

$$\langle v + v', \operatorname{grad} f(x') \rangle = 0,$$



Figure 14: To the proof of Proposition 4.1.

 \mathbf{SO}

$$J(x',v') = -\frac{1}{2} \langle v', \operatorname{grad} f(x') \rangle = \frac{1}{2} \langle v, \operatorname{grad} f(x') \rangle.$$

We need to show that this is equal to

$$J(x,v) = -\frac{1}{2} \langle v, \operatorname{grad} f(x) \rangle,$$

which is equivalent to proving that

$$\langle v, \operatorname{grad} f(x) + \operatorname{grad} f(x') \rangle = 0$$

Also notice that v is collinear to x' - x (by definition of the billiard map), so we need to show that

$$\langle x - x', \operatorname{grad} f(x) + \operatorname{grad} f(x') \rangle = 0.$$
 (7)

To show this, we will use that f is a homogeneous quadratic function. Write f in the form $f = \langle x, Ax \rangle$, where

$$A = \left(\begin{array}{cc} \frac{1}{a_1^2} & 0\\ 0 & \frac{1}{a_2^2} \end{array}\right)$$

Notice that grad f(x) = 2Ax. So, the left hand side of (7) can be rewritten as

$$2\langle x - x', Ax + Ax' \rangle = 2\left(\langle x, Ax \rangle - \langle x', Ax' \rangle + \langle x, Ax' \rangle - \langle x', Ax \rangle\right).$$

But $\langle x, Ax \rangle = f(x) = 1$, since x lies on the ellipse. Likewise, $\langle x', Ax' \rangle = 1$. Finally, notice that $\langle x, Ax' \rangle = \langle x', Ax \rangle$ since A is symmetric. So, we conclude that (7) indeed holds, as desired.

We now want to find a geometric interpretation of the Joachimsthal integral J. To that end, we will need a more geometric definition of an ellipse:

Definition 4.4. Let f_1, f_2 be two points in the Euclidian plane \mathbb{R}^2 , and let l > 0 be a positive real number. Then the set of points $\{x \in \mathbb{R}^2 \mid |xf_1| + |xf_2| = l\}$ is called an *ellipse with foci* f_1, f_2 . Here $|xf_i|$ stands for the Euclidian distance between x and f_i . Similarly, the set of points $\{x \in \mathbb{R}^2 \mid ||xf_1| - |xf_2|| = l\}$ is called a *hyperbola with foci* f_1, f_2 .

We now want to obtain an analytic description of ellipses and hyperbolas with given foci f_1 , f_2 . We will assume that $f_1 = (-a, 0)$, $f_2 = (a, 0)$. By doing so we do not loose any generality because any pair of points has such coordinates in a suitable orthogonal coordinate system.



Figure 15: The foci of a horizontal/vertical ellipse.

Proposition 4.5. The equation of an ellipse/hyperbola with foci $f_1 = (-a, 0)$, $f_2 = (a, 0)$ and given parameter l > 0 is

$$\frac{x_1^2}{l^2/4} + \frac{x_2^2}{l^2/4 - a^2} = 1.$$
(8)

Remark 4.6. The reader may be confused by the fact that an ellipse and a hyperbola with the same parameters are described by the same equation. But there is no contradiction here. The point is that by the triangle inequality we must always have l > 2a for an ellipse and l < 2a for a hyperbola. So for any l there may exist either an ellipse, or a hyperbola with such l, but not both at a time.

The proof of Proposition 4.5 is a straightforward algebraic verification of the equivalence between equation (8) and Definition 4.4.

Lecture 5: Geometric meaning of the Joachimsthal integral

In this lecture we will reveal the geometric meaning of the Joachimsthal integral (6) for the billiard map in an ellipse. Consider once again an ellipse given by

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1.$$
(9)

Without loss of generality, we may assume that this ellipse is "horizontal", i.e. $a_1 \ge a_2$ (this can be always achieved by rotating the coordinate system if necessary). Furthermore, since we already studied the circle case in detail, we will assume that our ellipse is not a circle, which means that $a_1 > a_2$.

Proposition 5.1. The foci of the ellipse (9) are the points $(\pm a, 0)$, where $a = \sqrt{a_1^2 - a_2^2}$.

Remark 5.2. We assumed that $a_1 \ge a_2$ to have foci on the horizontal axis. For $a_1 \le a_2$, the foci are on the vertical axis (see Figure 15).

Proof of Proposition 5.1. Equation (9) coincides with (8) if we set $a = \sqrt{a_1^2 - a_2^2}$ and $L = 2a_1$.

It follows that the equations of conics (i.e. ellipses and hyperbolas) which are confocal with (i.e. have the same foci as) our ellipse (9) is

$$\frac{x_1^2}{l^2/4} + \frac{x_2^2}{l^2/4 - a_1^2 + a_2^2} = 1$$

This can be written in a more symmetric way if we define $\lambda = l^2/4 - a_1^2$. Then the above equation becomes

$$\frac{x_1^2}{a_1^2 - \lambda} + \frac{x_2^2}{a_2^2 - \lambda} = 1.$$
(10)



Figure 16: A confocal family. The solid dots are the foci.

This is a standard equation of a *confocal family*. For an arbitrary value of $\lambda \neq a_1^2, a_2^2$, this equation defines a conic, and all these conics have the same foci as the ellipse (9). And conversely, any conic confocal with the ellipse (9) is contained in the family (10). Figure 16 shows an example of a confocal family.

We will now be interested in the following question: given a line in \mathbb{R}^2 , how many, if any, conics from the family (10) are tangent to that line? The following lemma says that there is almost always exactly one, except for a couple of cases when there is none:

Lemma 5.3. Consider the line x + tv through the point $x = (x_1, x_2)$ with direction $v = (v_1, v_2)$. Then

- 1. This line is tangent to at most one conic from the family (10).
- 2. If such a tangent conic exists, then its parameter λ which distinguishes it in the family (10) is given by

$$\lambda = \frac{a_1^2 v_2^2 + a_2^2 v_1^2 - (x_1 v_2 - x_2 v_1)^2}{v_1^2 + v_2^2}.$$
(11)

- 3. Such tangent conic does not exist in the following two cases:
 - (a) The number λ given by formula (11) is equal to a_1^2 , in which case the line x + tv coincides with the minor axis of the ellipse (9).
 - (b) The number λ given by formula (11) is equal to a_2^2 , in which case the line x + tv is passing through one of the foci of the ellipse (9).

Proof. The line x + tv is tangent to the conic (10) when the equation

$$\frac{(x_1 + tv_1)^2}{a_1^2 - \lambda} + \frac{(x_2 + tv_2)^2}{a_2^2 - \lambda} = 1$$

for their intersection points has exactly one solution in terms of t; that is, when its discriminant is equal to 0. Equating the discriminant to 0 and solving for λ , we get formula (11). The expression on the right-hand side of (11), however, may be equal to a_1^2 or a_2^2 , which does not correspond to any conic in the family (10). In that case, there is no conic in the family (10) tangent to the line x + tv. In all other cases, such a conic exists, is unique, and corresponds to λ given by (11).

To complete the proof it now suffices to obtain a geometric interpretation of the cases $\lambda = a_1^2$ and $\lambda = a_2^2$. First assume that $\lambda = a_1^2$. Then (11) gives

$$(a_1^2 - a_2^2)v_1^2 = -(x_1v_2 - x_2v_1)^2.$$

Notice the left-hand side of this equation is non-negative (since $a_1 > a_2 > 0$), while the right-hand side is non-positive. Therefore, this equation holds if and only if both sides are equal to 0, which is equivalent to $v_1 = 0$ and $x_1 = 0$. But these two conditions together hold precisely when the line x + tv coincides with the vertical coordinate axis or, which is the same, with the minor axis of the ellipse (9).

Similarly, if $\lambda = a_2^2$, then

$$(a_2^2 - a_1^2)v_2^2 = -(x_1v_2 - x_2v_1)^2,$$

which is equivalent to

 $x_1v_2 - x_2v_1 = \pm av_2,$

with $a = \sqrt{a_1^2 - a_2^2}$. The geometric meaning of this equation is that the vectors $(x_1, x_2) - (\pm a, 0)$ and (v_1, v_2) are collinear, which is the same as to say that the line x + tv is passing through the one of the points (a, 0), (-a, 0). But those points are precisely the foci of the ellipse (9), hence the result.

We now show that Joachimsthal integral computed at a point (x, v) of the phase cylinder for the billiard in an ellipse is closely related to the parameter λ of the confocal conic to which the line x + tv is tangent:

Proposition 5.4. The quantity λ given by formula (11), regarded as a function on the phase cylinder M of the billiard in the ellipse (9), is related to the Joachimsthal integral J by the formula

$$\lambda = a_1^2 a_2^2 J^2. \tag{12}$$

Proof. For a unit vector v, formula (11) becomes

$$\lambda = a_1^2 v_2^2 + a_2^2 v_1^2 - (x_1 v_2 - x_2 v_1)^2 = (a_1^2 - x_1^2) v_2^2 + (a_2^2 - x_2^2) v_1^2 + 2x_1 x_2 v_1 v_2.$$
(13)

Furthermore, since the point (x_1, x_2) lies on the ellipse (9), we have

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1 \quad \Rightarrow \quad \frac{x_1^2 - a_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 0 \quad \Rightarrow \quad a_1^2 - x_1^2 = \frac{a_1^2 x_2^2}{a_2^2}.$$

Similarly, we have

$$a_2^2 - x_2^2 = \frac{a_2^2 x_1^2}{a_1^2},$$

so (13) can be written as

$$\lambda = \frac{a_1^2 x_2^2}{a_2^2} + \frac{a_2^2 x_1^2}{a_1^2} + 2x_1 x_2 v_1 v_2 = a_1^2 a_2^2 \left(\frac{x_1 v_1}{a_1^2} + \frac{x_2 v_2}{a_2^2}\right)^2 = a_1^2 a_2^2 J^2,$$

as desired.

Corollary 5.5. 1. Assume that a segment of a billiard trajectory in the ellipse (9) is tangent to some confocal conic (10). Then all segments of that trajectory are tangent to that conic.

- 2. Assume that a segment of a billiard trajectory in the ellipse (9) belongs to the minor axis. Then all segments belong to the minor axis.
- 3. Assume that a segment of a billiard trajectory in the ellipse (9) is passing through one of the foci. Then all segments are passing through one of the foci.

Proof. From (12) and preservation of J it follows that the function λ given by (11) is preserved by the billiard map. The above three statements are the particular cases of this result corresponding to, respectively, generic values of λ , $\lambda = a_1^2$, and $\lambda = a_2^2$.



Figure 17: A billiard trajectory along the minor axis of an ellipse.

Remark 5.6. In view of Proposition 5.4, Corollary 5.5 is equivalent to preservation of the Joachimsthal integral and thus can be viewed as a geometric form of the latter. Furthermore, all statements of Corollary 5.5 can be obtained geometrically. The second statement is particularly straightforward: since the minor axis is orthogonal to the ellipse, a billiard ball moving along that axis will continue doing so after any number of reflections (see Figure 17). Of course, the same is true for the major axis, but the major axis is not distinguished by any specific value of λ : its λ is the same as for any other line passing through one of the foci.

Also note that the third statement of the corollary can be strengthened as follows: if a segment of a billiard trajectory in the ellipse is passing through one of the foci, then the next segment must pass through *the other* focus. Indeed, a generic billiard trajectory through one of the foci is not orthogonal to the ellipse (see Figure 18) and hence cannot be reflected to the same focus. So, almost all trajectories through one of the foci get reflected to the other focus, and by continuity it must be true for all such trajectories.

The fact that a billiard trajectory through one of the foci is reflected to the other focus is known as the *optical property of the ellipse*. It can be reformulated by saying that all light rays starting at one of the foci get reflected, by an elliptic mirror, to the other focus. The optical property can be proved quite easily without using the Joachimsthal integral, see e.g. [18, Lemma 4.2]. See also [18, Theorem 4.4] for an independent geometric proof of the first result of Corollary 5.5.

Lecture 6: The phase portrait of the billiard in an ellipse

The phase cylinder M of the billiard in an ellipse

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1. \tag{14}$$

is foliated into level sets of the Joachimsthal integral

$$J = -\left(\frac{x_1v_1}{a_1^2} + \frac{x_2v_2}{a_2^2}\right).$$
 (15)

Each of these level sets is preserved by the billiard map T. Furthermore, according to the previous lecture each of these level sets can be seen as the set of velocity vectors $(x, v) \in M$ corresponding to billiard



Figure 18: A segment of a billiard trajectory through a focus of an ellipse.

trajectories tangent to one and the same conic confocal with the ellipse (14). This confocal conic is explicitly given by

$$\frac{x_1^2}{a_1^2 - \lambda} + \frac{x_2^2}{a_2^2 - \lambda} = 1,$$
(16)

where the constant λ is related to the value of the Joachimsthal integral by the formula

$$\lambda = a_1^2 a_2^2 J^2. \tag{17}$$

We would now like to understand the structure of level sets of the function J. To that end, we rewrite this function in terms of coordinates on M. As one of the coordinates, we take the angle α (the angle between the velocity vector v and the positively oriented ellipse), and as a second coordinate we take the parameter t on the ellipse (14) corresponding to the parametrization

$$x_1 = a_1 \cos(t), \quad x_2 = a_2 \sin(t).$$

Remark 6.1. Note that this is not an arc length parameter, and in fact one cannot write the arc length parametrization of an ellipse in terms of elementary functions.

Now, as in Lecture 4, we rewrite (15) as

$$J = -\frac{1}{2} \langle v, \operatorname{grad} f(x) \rangle, \tag{18}$$

where

$$f(x) = \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2}.$$

Further, since v is a unit vector, and grad f(x) is an outward normal to the ellipse, (18) can be rewritten as

$$J = -\frac{1}{2} ||\operatorname{grad} f(x)|| \cos(\alpha + \pi/2) = \frac{1}{2} ||\operatorname{grad} f(x)|| \sin \alpha = ||\operatorname{grad} (f(x)/2)|| \sin \alpha = \sqrt{\frac{x_1^2}{a_1^4} + \frac{x_2^2}{a_2^4}} \sin \alpha,$$

so in (t, α) coordinates we have

$$J = \sqrt{\frac{\cos^2 t}{a_1^2} + \frac{\sin^2 t}{a_2^2}} \sin \alpha$$

The level sets of this function are shown in Figure 19. This picture should be thought of as living on a cylinder, which is obtained by identifying the opposite sides t = 0 and $t = 2\pi$ of the colored rectangle. The solid dots are located at critical points $(0, \pi/2)$, $(\pi/2, \pi/2)$, $(\pi, \pi/2)$, $(3\pi/2, \pi/2)$ of the function J. They correspond to (initial velocity vectors of) billiard trajectories aligned with one of the axes of the ellipse. The colors are used to show two regions with qualitatively different behavior of billiard trajectories, as explained below.

Consider first the blue region. On the exterior boundaries of this region (given by $\alpha = 0$ and $\alpha = \pi$) we have J = 0, while the value of J on the interior boundary is equal to that at the point $(0, \pi/2)$ and hence is $1/a_1$. Therefore, by formula (17), we have $\lambda = 0$ at the exterior boundary and $\lambda = a_2^2$ at the interior boundary. And since there is no critical points in the interior of the region, in that interior we must have $0 < \lambda < a_2^2$. But then it follows from equation (16) that confocal conics tangent to billiard rays with initial conditions in the (interior of the) blue region are *ellipses*. Another important feature of the blue domain is that even though each level set of J in that domain consists of two connected components (corresponding



Figure 19: The phase portrait of the billiard in an ellipse.

to $\alpha < \pi/2$ and $\alpha > \pi/2$), the billiard map preserves each of those components individually. Indeed, if we take an initial velocity vector v with angle α close to 0, then its image under the billiard map will also have α close to 0. Therefore, components of level sets of J close to the lower boundary $\alpha = 0$ cannot be mapped to components close to the upper boundary $\alpha = \pi$. But then it follows by continuity that all components in the $\alpha < \pi/2$ region are mapped to themselves, and the same is true for the $\alpha > \pi/2$ region.

As for the green region, a similar analysis shows that it corresponds to trajectories tangent to confocal *hyperbolas* (except for the critical points $(\pi/2, \pi/2)$, $(3\pi/2, \pi/2)$ whose corresponding trajectories are aligned with the minor axis and hence are not tangent to any confocal conic). Another difference with the blue region is that for the green one two connected components of each *J*-level are *interchanged* by the billiard map. Indeed, this is obviously true for the critical points $(\pi/2, \pi/2)$, $(3\pi/2, \pi/2)$, $(3\pi/2, \pi/2)$ and thus is true for all components by continuity.

The curves separating the regions of different color are *separatrices*. Since at those curves we have $\lambda = a_2^2$, they correspond to billiard trajectories through the foci of the ellipse. Consider, for example, the upper-left separatrix. On this separatrix we have $t < \pi$ and $\alpha > \pi/2$, which means that the corresponding velocity vectors v are attached at the upper half of the ellipse and make an obtuse angle with its positive direction. From this it is easy to see that such vectors give rise to billiard trajectories whose initial segments pass through the right focus. Similarly, the lower-right separatrix also corresponds to the right focus, while the lower-left and the upper-right ones correspond to the right focus.

Lecture 7: Liouville integrable maps in 2D

The billiard map for an ellipse preserves an area form and a function. In this lecture, we will discuss this situation in an abstract setting:

Definition 7.1. Let M be a 2-dimensional manifold, and let $T: M \to M$ be an area-preserving diffeomorphism² (i.e. there is an area form ω on M such that $T^*\omega = \omega$). Assume also that T has a first integral (i.e. a smooth function $f: M \to \mathbb{R}$ such that $T^*f = f$) whose differential does not vanish almost everywhere on M. Then the map T is called *Liouville integrable, completely integrable*, or just *integrable*.

Here "almost everywhere" means "on an open dense subset". Note that the set where $df \neq 0$ is automatically open, so what is really required is that this set is dense. This is equivalent to saying that there exist no open subset $U \subset M$ such that the restriction of f to U is constant. In particular, this prohibits the situation when f is a constant function. Constant functions are first integrals for any map, so there is absolutely no reason to call the above map T integrable if it possesses a constant first integral.

One of the main general results on integrable systems is that integrability implies quasi-periodic dynamics. More precisely, under some natural assumptions, integrability implies quasi-periodic dynamics on regular level sets of f. A level set of a smooth function $f: M \to \mathbb{R}$ is called regular if it contains no critical points of f, i.e. zeros of df. Our assumption that $df \neq 0$ almost everywhere, combined with Sard's lemma, implies that the union of regular level sets is dense in M. So, integrability guarantees that the dynamic is relatively simple everywhere except possibly a nowhere dense set.

Theorem 7.2 (Arnold-Liouville theorem for maps in 2D). Let M be a 2-dimensional manifold, and let $T: M \to M$ be an integrable area-preserving diffeomorphism with first integral $f: M \to \mathbb{R}$. For $a \in \mathbb{R}$, let $M_a = f^{-1}(a)$ be the corresponding level set of f. Assume that M_a is regular, compact, and connected. Then M_a is diffeomorphic to a circle, and the dynamics of T restricted to M_a is quasiperiodic. In other words, there is a periodic coordinate $\phi: M_a \to \mathbb{R}/2\pi\mathbb{Z}$ such that in terms of ϕ the map $T|_{M_a}$ is a translation $\phi \mapsto \phi + c$.

Remark 7.3. The standard version of the Arnold-Liouville theorem is for differential equations (vector fields) instead of maps. We will discuss that version later on in the course. The version for maps is due to A. Veselov [20].

Remark 7.4. The fact that M_a is a circle is trivial, since any compact connected 1-dimensional manifold is (diffeomorphic to) a circle. The reason we emphasize this statement is to make connections with the multidimensional version of this theorem where a circle is replaced by a torus. In that multidimensional setting, the statement becomes not trivial at all: tori are very far from being the only compact connected manifolds in any dimension d > 1.

Remark 7.5. Instead of requiring that M_a is connected, we can assume that T preserves the given connected component M_a^0 of M_a . Then the conclusion of the theorem stays the same, with M_a replaced by M_a^0 . Furthermore, even if M_a^0 is not preserved by T, there still exists k > 1 such that M_a^0 is preserved by T^k (because M_a is compact and thus has only finitely many connected components). So, we can always apply the theorem for some iteration of T.

Also note that the compactness assumption is not really essential in our 2D case either, but it does become important in the multidimensional setting. Furthermore, it simplifies things even in 2D, so we include it.

To prove Theorem 7.2, we introduce the notion of a Hamiltonian vector field (which is also very important on its own right). Let M be a 2D manifold equipped with an area form ω . Then, for any $x \in M$, ω is a non-degenerate skew-symmetric bilinear form on $T_x M$ and thus can be viewed as an invertible linear map $\omega: T_x M \to T_x^* M$. Therefore, there is a well-defined linear map $\omega^{-1}: T_x^* M \to T_x M$.

²If M is compact, then any area-preserving smooth map of M to itself is automatically a diffeomorphism. This is, however, not true in the non-compact case. For example, $(x, y) \mapsto (x + 1, y)$ is an area-preserving map of the right-half plane to itself (with respect to the standard area form $dx \wedge dy$), but it is not surjective.

Definition 7.6. Let M be a 2-dimensional manifold M equipped with an area form ω . Let also $f: M \to \mathbb{R}$ be a smooth function. Then the associated Hamiltonian vector field X_f is defined by $X_f = \omega^{-1}(df)$. The function f itself is called the Hamiltonian corresponding to the field X_f .

Remark 7.7. A different notation for the Hamiltonian vector field associated with f is sgrad f, the skew gradient of f. Its definition is indeed similar to that of a gradient, with the metric replaced by the area form.

Example 7.8. Assume that $\omega = dx \wedge dy$. Then the components of X_f are $\frac{\partial f}{\partial y}$ and $-\frac{\partial f}{\partial x}$.

Proposition 7.9. The derivative of f along the corresponding Hamiltonian vector field X_f is 0.

Proof. This derivative is equal to $\langle df, X_f \rangle$, where \langle , \rangle is the pairing between covectors and vectors. But

$$\langle df, X_f \rangle = \langle \omega(X_f), X_f \rangle = \omega(X_f, X_f) = 0,$$

as desired.

Lecture 8: Proof of the Arnold-Liouville theorem in 2D and Poncelet's closure theorem

We are now in a position to prove Theorem 7.2. First note that the regular level set $M_a = f^{-1}(a)$ of fis automatically a circle simply because it is a compact connected 1-dimension manifold. So it suffices to construct a periodic coordinate $\phi: M_a \to \mathbb{R}/2\pi\mathbb{Z}$ such that in terms of ϕ the map $T|_{M_a}$ is a translation $\phi \mapsto \phi + c$. To that end, consider the Hamiltonian vector field X_f associated with f. By Proposition 7.9, the vector field X_f is tangent to the circle M_a and, therefore, can be restricted to it. Also note that X_f does not vanish in M_a since $X_f = \omega^{-1}(df)$ vanishes only at those points where df = 0, and there is no such points in M_a . So, X_f is a non-vanishing vector field on a circle. But then there must exist a periodic coordinate $\psi: M_a \to \mathbb{R}/L\mathbb{Z}$ on M_a such that $X_f|_{M_a} = \partial/\partial \psi$. Locally the existence of such a coordinate is a particular case of the general rectification theorem for vector fields, but in the case of the circle this is also true globally, and can be shown as follows. Let s be any periodic coordinate ψ such that X_f , written in terms of s, reads $v(s)\partial/\partial s$. We want to find a new coordinate ψ such that

$$\frac{\partial}{\partial \psi} = v(s)\frac{\partial}{\partial s}.$$

The latter expression can be rewritten as

$$v(s)\frac{\partial}{\partial s} = v(s)\frac{d\psi}{ds}\frac{\partial}{\partial\psi},$$

 \mathbf{SO}

$$\frac{\partial}{\partial \psi} = v(s) \frac{d\psi}{ds} \frac{\partial}{\partial \psi},$$

which means that ψ can be defined by

$$\psi = \int \frac{ds}{v(s)}.$$

The latter expression has a simple physical meaning. Define the distance traveled as the net change in the s coordinate. Then ds is an infinitesimal displacement, and ds/v(s) is the time it takes to travel the infinitesimal distance ds with speed v(s). So, the above integral can be interpreted as the total time it takes to get from a fixed point on the circle to a given one, when traveling with velocity prescribed by

our vector field $v(s)\partial/\partial s$. It is not a surprise that ψ has a meaning of time: since we want the velocity written in terms of the new coordinate to be equal to 1, the displacement and time should be the same.

The above formula for ψ also suggests a coordinate-free construction of that function. Notice that the form $\alpha = ds/v(s)$ takes the value 1 on the vector $X_f = v(s)\partial/\partial s$. So, this form can be defined in a coordinate-free fashion as the unique 1-form α such that at every point of M_a we have $\langle \alpha, X_f \rangle = 1$. In other words, at every point $x \in M_a$, $\{\alpha|_x\}$ is the basis of $T_x^*M_a$ dual to the basis $\{X_f|_x\}$ of T_xM_a . Then the function ψ can be constructed as

$$\psi(x) := \int_{x_0}^x \alpha,$$

where $x_0 \in M_a$ is any fixed point. This integral is well-defined up to the integral of α over the whole circle, so it is indeed a periodic coordinate (one also needs to check that $d\psi \neq 0$, but this follows from $d\psi = \alpha$). Furthermore, it is immediate from this construction that $X_f = \partial/\partial \psi$, because $\langle d\psi, X_f \rangle = \langle \alpha, X_f \rangle = 1$.

We are now in the following situation. On the circle M_a , we have a *constant* vector field $X_f = \partial/\partial \psi$. This vector field is preserved by the map T, because T preserves both the Hamiltonian f and the area form, and X_f is constructed out of these two ingredients. But the only maps that preserve a constant (non-zero) vector field on a circle are translations. The easiest way to see that is once again consider the 1-form $d\psi$ dual to the vector-field X_f . Since X_f is preserved by T, so is $d\psi$. But this means that

$$dT^*\psi = T^*d\psi = d\psi \quad \Leftrightarrow \quad d(T^*\psi - \psi) = 0 \quad \Leftrightarrow \quad T^*\psi = \psi + c,$$

where c is a constant. But this means that in terms of ψ the map T is $\psi \mapsto \psi + c$. This is almost what we need, the only problem is that ψ is L-periodic instead of 2π -periodic. This can be fixed by defining a new coordinate $\phi = 2\pi\psi/L$. In terms of this new coordinate, the map T is still a translation, by $2\pi c/L$. Thus, the Arnold-Liouville theorem for maps in 2D is proved.

Corollary 8.1. Let M be a 2-dimensional manifold, and let $T: M \to M$ be an integrable area-preserving diffeomorphism with first integral $f: M \to \mathbb{R}$. For $a \in \mathbb{R}$, let $M_a = f^{-1}(a)$ be the corresponding level set of f. Assume that M_a is regular, compact, and connected. Assume also that there is a natural number $k \geq 1$ and a point $x \in M_a$ such that $T^k(x) = x$. Then $T^k(y) = y$ for all points $y \in M_a$. In other words, if $T|_{M_a}$ has a periodic point, then all points in M_a are periodic with the same period.

Proof. By the Arnold-Liouville theorem, there is a 2π -periodic coordinate ϕ on M_a such that, in terms of this coordinate, the map $T|_{M_a}$ reads $\phi \mapsto \phi + c$. Then the map $T^k|_{M_a}$ reads $\phi \mapsto \phi + kc$, and since $T^k(x) = x$, we must have that kc is an integer multiple of 2π . But this means that $T^k|_{M_a}$ is the identity map, as desired.

Remark 8.2. As can be seen from the proof of the Arnold-Liouville theorem, the coordinate ϕ in which $T|_{M_a}$ is a rotation can be defined by an explicit formula involving integration. Furthermore, in terms of ϕ the dynamics of T, i.e. its iterate T^n , takes the form $\phi \mapsto \phi + nc$, where c can be computed as $\phi(T(x)) - \phi(x)$ for any $x \in M_a$. So, to compute $T^n(x)$, one needs to do the following. First, find the ϕ -coordinates of x and T(x) by performing integration. Then compute c and hence the ϕ -coordinate of $T^n(x)$. Finally, one needs to invert the formula expressing ϕ , in order to find $T^n(x)$ in original coordinates. So, an integrable system can be explicitly "solved" in terms of integration and computing inverse functions. A possibility to give a solution using just these operations is known as "solvability by quadratures". Quadrature is a historical term which means the process of determining area, i.e. integration. Thus, it follows from (the proof of) the Arnold-Liouville theorem that integrable maps in 2D are solvable by quadratures. This result is in fact usually included as one of the statements of the theorem.

From Corollary 8.1 applied to elliptic billiards, we get the following geometric result, known as Poncelet's closure theorem:



Figure 20: Poncelet triangles inscribed in \mathcal{E}_1 and circumscribed about \mathcal{E}_2 .

Theorem 8.3. Let $\mathcal{E}_1, \mathcal{E}_2$ be two ellipses. Assume that \mathcal{E}_2 lies in the interior of the domain bounded by \mathcal{E}_1 . Assume also that there exists an n-gon inscribed in \mathcal{E}_1 and circumscribed about \mathcal{E}_2 . Then there exists infinitely many such n-gons, and every point on \mathcal{E}_1 is a vertex of such an n-gon (see Figure 20 illustrating the case n = 3).

Proof. We will first prove the theorem under the assumption that \mathcal{E}_1 and \mathcal{E}_2 are confocal. In this case, an *n*-gon inscribed in \mathcal{E}_1 and circumscribed about \mathcal{E}_2 is a closed billiard trajectory in \mathcal{E}_1 . Indeed, any side of such an *n*-gon ca be viewed as a segment of a billiard trajectory. It is tangent to \mathcal{E}_2 , so by Corollary 5.5 all segments of the corresponding trajectory are tangent to \mathcal{E}_2 . So both the given polygon and the billiard trajectory are inscribed in \mathcal{E}_1 and circumscribed about \mathcal{E}_2 , and since they have a common segment, they must coincide.

Now, it suffices to prove that any other billiard trajectory in \mathcal{E}_1 which is tangent to \mathcal{E}_2 is closed with the same period. But this follows from Corollary 8.1, because by Proposition 5.4 trajectories tangent to the same confocal conic belong to the same level set of the Joachimsthal integral. Note also that even though this level set of the Joachimsthal integral is not connected, it lies in the blue region in Figure 19, which means that each connected component of the level set is preserved by the billiard map. Therefore, we can simply apply Corollary 8.1 with M being the upper part of the blue region (it is also not too hard to modify this argument to work for the green region as well).

To prove the theorem for arbitrary, not necessarily confocal, ellipses $\mathcal{E}_1, \mathcal{E}_2$, one can use a projective transformation taking them to confocal ones. For details, see e.g. [11, Section 3].

Lecture 9: Algebraic integrability of elliptic billiards I

We will now discuss the concept of algebraic integrability. Algebraic integrability is a different approach to establishing quasiperiodic dynamics, based on complex geometry. The advantage of this approach, as compared to the one provided by the Arnold-Liouville theorem, is that it explains the geometric meaning of the coordinate ϕ in which an integrable map is a rotation. It turns out that in order to reveal this meaning, one needs to complexify the phase space as well as the level sets of the first integral. We will demonstrate this approach by showing how it works in the elliptic billiard example. Our exposition is a variation on the proof of Poncelet's closure theorem by P. Griffiths and J. Harris [5].

Recall that the phase space M of the billiard in an ellipse \mathcal{E}_1 consists of pairs (x, v), where $x \in \mathcal{E}_1$ is a point on the ellipse, and $v \in T_x \mathbb{R}^2$ is an inward tangent vector at x. Clearly, for fixed x, the vector v is uniquely determined by the non-oriented direction that it defines. Moreover, almost all directions are realized in this way, except for the direction of the tangent line at x. So, one can regard v as an element of \mathbb{RP}^1 which is allowed to take any value except for one. For the sequel, it will be convenient to allow that value corresponding to the tangent line as well. This turns the phase space into the direct product $\mathcal{E}_1 \times \mathbb{RP}^1$, i.e. a torus. In terms of the phase cylinder M, adding tangent directions corresponds to identifying the opposite boundaries of M. One can easily see that both the billiard map, and its first integral are well-defined on this toric version of the phase space (for vectors v tangent to the ellipse, we set J = 0 and define the billiard map to be the identity).

The complexified version of the phase space can now be defined as the space of pairs (x, v), where x belongs to the complex ellipse, and $v \in \mathbb{CP}^1$ is a direction. A complex ellipse³ is defined in the complex plane \mathbb{C}^2 by the same equation as its real counterpart, i.e.

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1.$$
(19)

However, this set is not compact, which is inconvenient. In order to compactify it, we will consider its closure inside the complex projective plane \mathbb{CP}^2 . This closure is obtained by "homogenizing" the equation (19), i.e. substituting $x_i = z_i/z_0$, and multiplying by a suitable power of z_0 to make the equation polynomial. This gives

$$\frac{z_1^2}{a_1^2} + \frac{z_2^2}{a_2^2} = z_0^2.$$
(20)

We will denote the set of points in \mathbb{CP}^2 whose homogeneous coordinates satisfy this equation in the same way as the real ellipse, i.e. \mathcal{E}_1 . It can be viewed as a natural complexification of the real ellipse⁴. The complexified billiard phase space is thus $\mathcal{E}_1 \times \mathbb{CP}^1$. Furthermore, the complex conic \mathcal{E}_1 itself can in fact be identified with \mathbb{CP}^1 . In order to do that, one fixes any point $x_0 \in \mathcal{E}_1$ and assigns to every $x \in \mathcal{E}_1$ the direction of the line through x_0 and x (where the direction assigned to x_0 itself is the limiting position of the chord x_0x as $x \to x_0$, i.e. the tangent line to \mathcal{E}_1 at x_0). This gives a one-to-one correspondence $\mathcal{E}_1 \simeq \mathbb{CP}^1$, which is moreover a biholomorphic map between complex manifolds. Thus, the complexified billiard phase space is complex diffeomorphic to $\mathbb{CP}^1 \times \mathbb{CP}^1$, while from the real point of view it is the product of spheres $S^2 \times S^2$.

We now complexify a regular level set M_a of the Joachimsthal integral J. To that end, recall that by Proposition 5.4 such a level set can be described as the set of segments of billiard trajectories which are tangent to a given confocal conic \mathcal{E}_2 . In the complex setting, we use this property as the definition of M_a : M_a is the set of pairs $(x, v) \in \mathcal{E}_1 \times \mathbb{CP}^1$ such that the (complex) line through x with direction vis tangent to a fixed complex conic \mathcal{E}_2 confocal with \mathcal{E}_1 (confocality in the complex projective setting can be defined by the homogenized version of the equation (10)). This tangency condition is algebraic, so M_a is defined as a subset of $\mathbb{CP}^1 \times \mathbb{CP}^1$ satisfying an algebraic equation. In any affine chart in $\mathbb{CP}^1 \times \mathbb{CP}^1$, this condition can be written in the form f(z, w) = 0, and for generic conics $\mathcal{E}_1, \mathcal{E}_2$, the zero level set of fis regular. Therefore, M_a is a one-dimensional complex (and hence two-dimensional real) submanifold of $\mathbb{CP}^1 \times \mathbb{CP}^1$. In the next lecture, we will investigate the topology of that manifold.

³In the complex setting, there is no difference between ellipses and hyperbolas, because one can always change the sign of the second term in the left-hand side of (19) by means of a change of variables $x_2 \mapsto \sqrt{-1}x_2$. So, in what follows, instead of talking about complex ellipses and hyperbolas, we will just use the word *conics*.

⁴Note that equation (20) describes ellipses, hyperbolas, and also parabolas. Indeed, a parabola is defined by the equation $x_2 = x_1^2$, whose homogeneous form is $z_0 z_2 = z_1^2$. By making a change of variables $z_0 \mapsto z_0 - z_2$, $z_2 \mapsto z_0 + z_2$, we can rewrite this equation in the form (20). The same is actually true in the real projective setting: in \mathbb{RP}^2 ellipses, hyperbolas, and parabolas are all projectively equivalent to each other.



Figure 21: The complexified level set of the first integral is a double covering of \mathbb{CP}^1 ramified at four points.

Lecture 10: Algebraic integrability of elliptic billiards II

In the previous lecture we defined the complexified level set M_a of the Joachimsthal integral as the set of pairs $(x,v) \in \mathbb{CP}^2 \times \mathbb{CP}^1$ such that x belongs to a fixed conic $\mathcal{E}_1 \subset \mathbb{CP}^2$, and the line through x in direction v is tangent to another fixed conic \mathcal{E}_2 confocal with \mathcal{E}_1 . For generic conics \mathcal{E}_1 , \mathcal{E}_2 , the set M_a is a smooth manifold of complex dimension one and real dimension two. Also note that M_a is compact as a closed subset of a compact manifold. Furthermore, we claim that M_a is connected. To prove that consider the projection $\pi: M_a \to \mathcal{E}_1 \simeq \mathbb{CP}^1$ given by $(x, v) \mapsto x$. The number of preimages of a point $x \in \mathcal{E}_1$ under this projection is the number of tangent lines to \mathcal{E}_2 through x. One can easily show that the number of tangent lines to a given conic $\mathcal{E} \subset \mathbb{CP}^2$ through a given point $x \in \mathbb{CP}^2$ is exactly two, unless $x \in \mathcal{E}$. In the latter case, there is only one tangent line, that is the tangent to \mathcal{E} at x. Thus $\pi^{-1}(x)$ consists of two points if $x \notin \mathcal{E}_2$, and one point if $x \in \mathcal{E}_2$. But $x \in \mathcal{E}_2$ if and only if x is an intersection point of \mathcal{E}_1 , \mathcal{E}_2 , and for generic conics in \mathbb{CP}^2 there are exactly four such intersection points. So, there exist four points $x_1, x_2, x_3, x_4 \in \mathcal{E}_1$ such that $|\pi^{-1}(x_i)| = 1$, while for points $x \neq x_i$ we have $|\pi^{-1}(x)| = 2$. Since the mapping $\pi: M_a \to \mathbb{CP}^1$ is 2-to-1 everywhere except for four points, one says that M_a is a *double* covering of \mathbb{CP}^1 ramified at four points. The mapping π is schematically shown in Figure 21 (of course, M_a does not really have self-intersection points, because it is smooth; a better model for the behavior of the map π near the preimages of the points x_i is given by the map $z \mapsto z^2$ near the origin; one can indeed show that π takes this form in an appropriate coordinate system near $\pi^{-1}(x_i)$). It is clear from this figure that M_a is connected: any two points in M_a can be connected by a path through one of the ramification points $\pi^{-1}(x_i)$.

Thus, M_a is a compact connected manifold of real dimension two. Furthermore, M_a is orientable. Indeed, M_a is a complex manifold and thus has holomorphic transition functions. But any holomorphic function viewed as a map from $\mathbb{R}^2 \to \mathbb{R}^2$ has positive Jacobian.

Now recall that any compact connected orientable surface is diffeomorphic to a "sphere with g handles". The number g is known as the *genus* of the surface. For example, a surface of genus zero is a sphere, a surface of genus one is a torus, a surface of genus two is a pretzel etc. The genus is closely related to the notion of *Euler characteristic*. The latter can be computed by considering an arbitrary triangulation of the surface. For any triangulation, let v be the number of vertices, e be the number of edges, and f be the number of faces. Then the Euler characteristic is defined as $\chi = v - e + f$; this number does not depend on the triangulation. The Euler characteristic χ and the genus g are related by the formula $\chi = 2 - 2g$.

We will now compute the Euler characteristic, and hence the genus, of the surface M_a using that it is a

double covering of \mathbb{CP}^1 ramified at four points. To that end, connect each of the points $x_1, x_2, x_3, x_4 \in \mathbb{CP}^1$ to each other by an edge. This can be done without self-intersection, and gives a triangulation of \mathbb{CP}^1 which combinatorially look like a tetrahedron. It has four vertices, six edges, and four faces. Note that 4-6+4=2, which agrees with the fact that the Euler characteristic of $\mathbb{CP}^1 \simeq S^2$ is equal to 2. We now lift this triangulation to M_a by considering the preimages of vertices, edges, and faces under the mapping π . This gives a triangulation of M_a . The number of vertices of that triangulation is still four, because each of the points x_i has exactly one preimage under π . However, the number of edges and faces doubles, because points $x \neq x_i$ have two preimages. So, we obtain a triangulation of M_a with four vertices, twelve edges, and eight faces. Thus, the Euler characteristic of M_a is $\chi = 4 - 12 + 8 = 0$, which means that g = 1, i.e. M_a is a torus.

We are now almost ready to describe the behavior of the complexified billiard map restricted to M_a . As a last preparatory step, we will need a description of holomorphic maps from a complex torus to itself. This description is based on the following fundamental result in complex geometry:

Theorem 10.1 (Uniformization theorem in genus one). Let X be a Riemann surface of genus one, i.e. a complex one-dimensional manifold which is real diffeomorphic to a torus. Then X is complex diffeomorphic to a surface of the form \mathbb{C}/L , where $L \subset \mathbb{C}$ is a full rank lattice, i.e. the set of integral linear combinations of two complex numbers which are linearly independent over \mathbb{R} .

We can now prove that there are very few holomorphic maps from a Riemann surface of genus one to itself (in fact, for higher genera there are even less):

Proposition 10.2. Assume that $\phi: \mathbb{C}/L \to \mathbb{C}/L$ is a holomorphic map from a Riemann surface of genus one to itself. Then ϕ is linear, i.e. $\phi(z) = \alpha z + \beta$.

Proof. The mapping ϕ can be viewed as a multivalued holomorphic function on \mathbb{C}/L , which is defined up to addition of complex numbers $z \in L$. Therefore, the derivative of ϕ is a genuine single-valued holomorphic function. But by the maximum principle there can be no non-constant holomorphic functions on a compact complex manifold. Thus ϕ' is constant, and ϕ is linear, as desired.

Remark 10.3. Not every function of the form $\alpha z + \beta$ defines a holomorphic map $\mathbb{C}/L \to \mathbb{C}/L$, because multiplication by α may not commute with the action of L on \mathbb{C} and hence may not descend to the quotient \mathbb{C}/L . In fact one can show that for generic lattices L the only bijective holomorphic maps $\mathbb{C}/L \to \mathbb{C}/L$ are of the form $z \mapsto \pm z + \beta$.

We could in principle use this remark to prove that the restriction of the complexified billiard map to M_a is a translation $z \mapsto z + \beta$, at least if the corresponding lattice L is generic. Indeed, the billiard map cannot be of the form $z \mapsto -z + \beta$, because the latter map is an *involution*, i.e. its square is the identity. However, the square of the billiard map is not identity for generic M_a , which means that the billiard map is of the form $z \mapsto z + \beta$ for almost all, and hence for all (by continuity), non-singular level sets M_a . Furthermore, this argument can be modified to work even for non-generic lattices, because even for such lattices any automorphism of \mathbb{C}/L which is not a translation has finite order. However, we will not pursue this approach, because in our setting it is easier to use the explicit description of the billiard map to show that it is a translation.

Proposition 10.4. The billiard map T for the ellipse, restricted to the complexified level set M_a of the first integral, is a translation relative to the group structure on M_a coming from the identification $M_a \simeq \mathbb{C}/L$.

Remark 10.5. The group structure on M_a is not completely unique, but it can be made unique by specifying a point $O \in M_a$ which is identified with the origin in \mathbb{C}/L . Indeed, if f,g are two different

biholomorphic maps from M_a to \mathbb{C}/L which take O to 0, then $f \circ g^{-1}$ is a biholmorphic automorphism of \mathbb{C}/L which preserves 0. Therefore, from the above classification of holomorphic maps of \mathbb{C}/L to itself it follows that $f \circ g^{-1}$ is of the form $z \mapsto az$. But any such map is a homomorphism of \mathbb{C}/L to itself, which guarantees that the pull-backs of the group structure from \mathbb{C}/L to M_a by means of f and g coincide⁵. Thus, any (compact and connected) Riemann surface of genus 1 with a distinguished point has a canonical group structure. A pair consisting of a Riemann surface of genus 1 and a distinguished point on it is called *an elliptic curve*. Sometimes this term is also used as a synonym for Riemann surface of genus 1 (without a distinguished point). Also note that the notion of a translation does not depend on the choice of the distinguished point (similarly to how the notion of translation in the affine space does not depend on the choice of the origin).

Proof of Proposition 10.4. We can think of M_a as the space of pairs (x, l), where x is a point in the conic \mathcal{E}_1 , and l is a line through x tangent to \mathcal{E}_2 . The billiard map $T: M_a \to M_a$ takes the pair (x, l) to the pair (x', l'), where x' is the second intersection point of l with \mathcal{E}_1 (along with x), and l' is the second tangent line from x' to \mathcal{E}_2 (along with l). Therefore, the billiard map can be written as a composition $\sigma_2 \circ \sigma_1$, where the maps $\sigma_{1,2} \colon M_a \to M_a$ are defined by $\sigma_1(x,l) = (x',l), \ \sigma_2(x',l) = (x',l')$. In other words, σ_1 interchanges the endpoints of a given tangent line to \mathcal{E}_2 , while σ_2 swaps the tangent lines to \mathcal{E}_2 passing through a fixed point. In particular, both maps σ_1 and σ_2 are involutions, i.e. $\sigma_1^2 = \sigma_2^2$ is the identity map. Also note that both σ_1 and σ_2 can be computed by solving a quadratic equation. For σ_1 , this is the equation determining the intersection points of a line and a conic. For σ_2 , this is the equation determining the tangent lines to a conic from a given point. In both cases, one of the roots of the corresponding quadratic equation is known, so the second root can be found by the Vieta theorem. This guarantees that the maps $\sigma_{1,2}: M_a \to M_a$ are holomorphic. So, the billiard map $T = \sigma_2 \circ \sigma_1$ is holomorphic as well. Thus, we can apply our classification of holomorphic maps $\mathbb{C}/L \to \mathbb{C}/L$ to T. it is, however, easier to first use that classification to describe the maps σ_1, σ_2 . Assume that $\sigma: \mathbb{C}/L \to \mathbb{C}/L$ is a holomorphic involution. Then $\sigma(z) = az + b$, and since it is an involution, we must have $a = \pm 1$. Furthermore, if a = +1, then $\sigma(z) = az + b = z + b$ is an involution if and only if $2b \in L$. To distinguish between involutions $\sigma(z) = -z + b$ and $\sigma(z) = z + b$, where $2b \in L$, one can count the fixed points of σ . In the $\sigma(z) = z + b$ case there is no fixed points unless $b \in L$, in which case σ is the identity. In the case $\sigma(z) = -z + b$ there is at least one fixed point, $\frac{1}{2}b$ (in fact, one can show that there are four fixed points). So, any non-trivial involution $\sigma \colon \mathbb{C}/L \to \mathbb{C}/L$ with fixed points must be of the form $z \mapsto b - z$. Now observe that the involutions $\sigma_{1,2}$ have fixed points. For σ_1 , the fixed points are of the form (x, l), where $x \in \mathcal{E}_1 \cap \mathcal{E}_2$, and l is the tangent line to \mathcal{E}_2 at x. For σ_1 , the fixed points are of the form (x, l), where *l* has a property of being tangent to both \mathcal{E}_1 and \mathcal{E}_2 . Thus, we have $\sigma_1(z) = b_1 - z$, $\sigma_2(z) = b_2 - z$. So, $T(z) = \sigma_2(\sigma_1(z)) = z + b_2 - b_1$, as desired.

Remark 10.6. As we know from Lecture 6, the real part of M_a consists of two circles. We thus have two closed curves on the torus M_a . According to Lecture 6 the billiard map may either preserve each of these circles (if we are in the blue region) or interchange them (if we are in the green region). This situation can be understood by looking at the following model example (in fact, one can prove that this is exactly what is happening). Consider the lattice L spanned by 1 and $\tau\sqrt{-1}$, where $\tau \neq 0$ is real. Then this lattice is invariant under complex conjugation. Therefore, one has a well-defined complex conjugation operation on \mathbb{C}/L (in this case one also says that the complex curve \mathbb{C}/L is endowed with a *real structure*). The set of real points in \mathbb{C}/L consists of points that are invariant under complex conjugation. This set is a union of two circles, one of which is the image of the real axis in \mathbb{C} , and the second one consists of numbers whose imaginary part is $\frac{1}{2}\tau\sqrt{-1}$. Then we have two kinds of translations in \mathbb{C}/L which preserve the set of the

⁵We define a pull-back of the group structure by $f: M_a \to \mathbb{C}/L$ as the unique group structure on M_a such that f is an isomorphism relative to that structure. Explicitly, the addition in this group structure is defined by $x+y=f^{-1}(f(x)+f(y))$.

real points: translation by a real number, and translation by a number whose imaginary part is $\frac{1}{2}\tau\sqrt{-1}$. Translations of the first type clearly preserve the components of the real part of \mathbb{C}/L , while translations of the second type interchange them.

Remark 10.7. Both the Arnold-Liouville theorem, and the complex geometric approach provide a coordinate ϕ defined modulo 2π and such that the billiard map written in this coordinate is a translation. We claim that the coordinates coming from these two approaches coincide. Indeed, assume that ϕ, ψ are two such coordinates. Then the billiard map must preserve two non-vanishing 1-forms $d\phi$ and $d\psi$. The ratio of these 1-forms is then a well-defined smooth function on the circle invariant under the billiard map. But if the billiard map is a rotation by an angle which is not a rational multiple of 2π , then its orbits are dense and it does not admit non-constant invariant functions. So $d\psi$ should be a constant multiple of $d\phi$, and since have integral 2π over the whole circle, they must coincide. Therefore, ϕ and ψ are the same up to an additive constant. Note this is also true for level sets on which the rotation angle is a rational multiple of 2π , by continuity.

Exercises.

- 1. Explain why any line and any conic in the complex projective plane either intersect at two distinct points, or are tangent to each other.
- 2. This exercise suggests a useful method or proving results in projective geometry, based on projective duality. Recall that the *dual projective plane* is defined as the set of straight lines in the given projective plane. More concretely, every line l is defined in homogeneous coordinates by an equation of the form $a_0z_0 + a_1z_1 + a_2z_2 = 0$, and a_0, a_1, a_2 are, by definition, the homogeneous coordinates of l in the dual projective plane. Show that straight lines passing through a given point form a straight line in the dual projective plane, and all lines in the dual plane are of this form (thus points of the dual plane can be identified with lines in the initial one, and vice versa). Then show that the set of tangents to a given conic forms a conic in the dual projective plane (*the dual conic*). Using this and the result of the previous exercise, prove that there always exist two distinct tangent lines to a given conic through a given point, provided that the point does not lie on the conic.
- 3. Consider the level set of the the Joachimsthal integral that corresponds to trajectories through the foci of the ellipse (this set is the boundary between two different colors in Figure 19). Show that the complexification of that level set is homeomorphic to two spheres attached to each other at two points.
- 4. Prove that each of the four pieces obtained by removing critical points from that level set admits a coordinate relative to which the billiard map T is a translation. Show that for any x in any of these four pieces the sequence $T^n(x)$ converges to a critical point. What geometric behaviour of the billiard trajectories does this correspond to?
- 5. For the billiard in a circle C define the complexified level set of the first integral as the set of pairs (x, l), where x is a point in the circle C and l is a line through x tangent to a fixed circle concentric with C. Prove that in this case *each* such level set is homeomorphic to two spheres attached to each other at two points.
- 6. Let p(x) be a cubic polynomial without multiple roots. Prove that the closure of the complex curve $y^2 = p(x)$ in \mathbb{CP}^2 is diffeomorphic to a torus.

Lecture 11: Canonical Hamilton's equations

In this lecture we discuss Hamilton's canonical formalism, which can be viewed as an ancestor of modern symplectic geometry. Let $q = (q_1, \ldots, q_n)$ be a point of mass m in \mathbb{R}^n moving in the potential force field with potential $U(q) = U(q_1, \ldots, q_n)$. Then, by definition of the potential, the force field is given by $F = -\operatorname{grad} U(q)$, and Newton's second law gives

$$mq_i'' = -\frac{\partial U}{\partial q_i}$$

where prime stands for the time derivative. We want to rewrite these equations as first order equations. To that end, introduce the momenta $p_i = mq'_i$. Then we get

$$\begin{cases} p'_i = -\frac{\partial U}{\partial q_i}, \\ q'_i = \frac{1}{m} p_i. \end{cases}$$

This can be further rewritten by introducing the function

$$H = U + \frac{m}{2} \sum_{i=1}^{n} (q'_i)^2 = U + \frac{1}{2m} \sum_{i=1}^{n} (p_i)^2,$$

which has the physical meaning of total energy and is also known as the *Hamiltonian*. In terms of this function, the above system can be rewritten as

$$\begin{cases} p'_i = -\frac{\partial H}{\partial q_i}, \\ q'_i = \frac{\partial H}{\partial p_i}. \end{cases}$$

These equations are known as *canonical Hamilton's equations*. In vector form, these equations can be rewritten as

$$x' = \pi dH(x),$$

where x is the column vector $(p,q)^t = (p_1, \ldots, p_n, q_1, \ldots, q_n)^t$, and π is the matrix

$$\pi = \left(\begin{array}{cc} 0 & -\mathrm{Id} \\ \mathrm{Id} & 0 \end{array}\right).$$

Note that π takes a covector dH(x) to a vector x'. So, for every $x \in \mathbb{R}^n$, the matrix π should be thought of as a map from the cotangent space at x to the tangent space at x. We can, therefore, also regard π as a bilinear form on cotangent vectors. Explicitly, this form can be written as

$$\pi = -\sum_{i=1}^n \frac{\partial}{\partial p_i} \wedge \frac{\partial}{\partial q_i}.$$

This form is skew-symmetric. Bilinear skew-symmetric forms on the dual space are also known as *bivectors* (we can also say that π is a *bivector field*, because it is a bivector defined on every tangent space). This particular bivector π entering Hamilton's equation is called the *Poisson tensor*.

One can also rewrite Hamilton's equation in terms of the symplectic structure. The latter is defined as $\omega = \pi^{-1}$. It is a map from the tangent space to the cotangent space, and thus can be viewed as a 2-form. By inverting the matrix of π we get

$$\omega = \left(\begin{array}{cc} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{array}\right),$$

$$\omega = \sum_{i=1}^{n} dp_i \wedge dq_i.$$

Note, in particular, that this form is closed. This property is very important and is included as one of the requirements in the abstract definition of a symplectic structure: a symplectic structure on a manifold is a closed non-degenerate skew-symmetric 2-form. We will be discussing the notion of an abstract symplectic structure, as well as of an abstract Poisson structure, in the next few lectures. We will also return to the above example of motion in the potential force field and see that this corresponds to a canonical symplectic structure on the cotangent bundle.

Lecture 12: Symplectic vector spaces and symplectic manifolds

Definition 12.1. A symplectic vector space is a vector space V endowed with a symplectic form ω , i.e. a non-degenerate skew-symmetric bilinear form.

Example 12.2. There exist no symplectic structures on one-dimensional vector spaces. Indeed, let V be a one-dimensional vector space, and let v be a basis vector in V. Then by skew-symmetry we have $\omega(v, v) = 0$, so ω is identically zero and hence degenerate.

Example 12.3. Any two-dimensional vector space V admits a symplectic structure. Such a structure is uniquely defined by $\omega(v_1, v_2) = 1$, where v_1, v_2 is any fixed basis in V. By skew-symmetry we also have $\omega(v_2, v_1) = -1$, $\omega(v_1, v_1) = 1$, $\omega(v_2, v_2) = 0$, so the matrix of ω in the basis v_1, v_2 is

$$\Omega = \left(\begin{array}{cc} 0 & 1\\ -1 & 0 \end{array}\right).$$

More generally, one has the following:

Proposition 12.4. A finite-dimensional vector space V admits a symplectic structure if and only if V is even-dimensional.

Proof. Let V be a symplectic vector space, and let Ω be the matrix of the symplectic form written in some basis. Then Ω is skew-symmetric, so

$$\det \Omega = \det(-\Omega^t) = (-1)^{\dim V} \det \Omega^t = (-1)^{\dim V} \det \Omega.$$

From the non-degeneracy of the symplectic structure we have that $\det \Omega \neq 0$, so it follows that $(-1)^{\dim V} = 1$, which means that V is even-dimensional. Conversely, if V is even-dimensional, then we can define a symplectic structure ω by taking a basis and considering a bilinear form which in this basis is given by the following block-diagonal matrix:

$$\Omega = \begin{pmatrix}
0 & 1 & & \\
-1 & 0 & & \\
& 0 & 1 & \\
& & -1 & 0 & \\
& & & \ddots \end{pmatrix}$$
(21)

It turns out that the above example of a symplectic form on an even-dimensional vector space is in fact universal in the following sense:

Theorem 12.5 (Linear Darboux theorem). In any finite-dimensional symplectic vector space there exists a basis in which the matrix of the symplectic form is (21).

Remark 12.6. By rearranging basis vectors v_1, \ldots, v_{2n} as $v_1, v_3, \ldots, v_{2n-1}, v_2, v_4, \ldots, v_{2n}$ we can also rewrite the above matrix as

 $\left(\begin{array}{cc}
0 & \mathrm{Id} \\
-\mathrm{Id} & 0
\end{array}\right),$ (22)

which is an already familiar to us symplectic structure of canonical Hamilton's equations.

This theorem can be proved by induction, based on the notion of a symplectic orthogonal complement, or, which is the same, skew-orthogonal complement. If W is a subspace of a symplectic vector space (V, ω) , then its orthogonal complement is defined by

$$W^{\perp} = \{ v \in V \mid \omega(v, w) = 0 \,\forall \, w \in W \}.$$

In contrast to the orthogonal complement in an inner product space, the symplectic orthogonal complement may not be an actual complement, i.e. in general it is not true that $V = W \oplus W^{\perp}$. The reason for that is that W and W^{\perp} may have non-trivial intersection. For example, any vector is symplectic orthogonal to itself, so any one-dimensional subspace is contained in its own symplectic orthogonal complement. However, even though W and W^{\perp} are in general not complementary to each other, their dimensions always are:

Lemma 12.7. For any subspace W of a symplectic vector space V, we have $\dim W^{\perp} = \dim V - \dim W$.

Proof. Consider the symplectic structure ω as a linear map $V \to V^*$. Then $\omega(W)$ consists of those covectors which annihilate W. So, by linear algebra we have $\dim \omega(W) = \dim V - \dim W$. But ω is non-degenerate, so $\dim \omega(W) = \dim W$, which proves the desired equality. \Box

Exercise 12.8. Deduce the linear Darboux theorem from this lemma.

The following subspaces in a symplectic vector space are of particular interest:

Definition 12.9. A subspace W of a symplectic vector space (V, ω) is called *isotropic* if the restriction of ω to W is zero, i.e. if $\omega(w_1, w_2) = 0$ for any $w_1, w_2 \in W$.

Example 12.10. Any one-dimensional subspace is isotropic.

Example 12.11. Assume that the symplectic form is given, in some basis v_1, \ldots, v_{2n} by the matrix (22). Then the subspace spanned by any subset of v_1, \ldots, v_n is isotropic, and so is the subspace spanned by any subset of v_{n+1}, \ldots, v_{2n} .

Proposition 12.12. The dimension of an isotropic subspace of a symplectic vector space V may not exceed $\frac{1}{2} \dim V$.

Proof. A subspace W is isotropic if and only if it is contained in its orthogonal complement W^{\perp} . So, if W is isotropic, then

 $\dim W \le \dim W^{\perp} = \dim V - \dim W,$

which implies

$$2\dim W \le \dim V,$$

q.e.d.

Definition 12.13. An istropic subspace W of a symplectic vector space V is called Lagrangian if it has maximal possible dimension $\frac{1}{2} \dim V$.

Lagrangian subspaces are also maximal isotropic in the sense that any isotropic subspace can be extended to a Lagrangian subspace. Indeed, if we have a non-Lagrangian isotropic subspace W, then by dimension reasons W is strictly contained in W^{\perp} , which implies that there exist vectors not in W which are orthogonal to all of W. Adding any such vector to W increases its dimension, and by doing so several times we will finally make W Lagrangian.

Example 12.14. Any one-dimensional subspace of a two-dimensional symplectic space is Lagrangian.

Example 12.15. Assume that the symplectic form is given, in some basis v_1, \ldots, v_{2n} by the matrix (22). Then the subspace spanned by v_1, \ldots, v_n is Lagrangian, and so is the subspace spanned by v_{n+1}, \ldots, v_{2n} .

We now turn our attention to symplectic manifolds.

Definition 12.16. A symplectic manifold is a manifold M endowed with a symplectic form (or symplectic structure) ω , i.e. a closed non-degenerate differential 2-form.

Since the tangent space of a symplectic manifold at every point is a symplectic vector space, it follows that there exist no odd-dimensional symplectic manifolds. So, the simplest example of a symplectic manifold is a two-dimensional one:

Proposition 12.17. Any orientable two-dimensional manifold admits a symplectic structure.

Proof. The closedness condition of a 2-form on a 2-manifold is vacuous, so a symplectic form in 2D is the same as an area form, i.e. a non-vanishing differential 2-form. Such a form can be constructed, for example, using a partition of unity subordinate to any oriented atlas (i.e. an atlas whose transition maps are orientation-preserving), or using a Riemannian metric. \Box

Clearly, orientability condition cannot be omitted, because an area form defines an orientation. So, there is no symplectic structures on non-orientable 2D manifolds. More generally, we have the following:

Proposition 12.18. Any symplectic manifold is orientable.

Proof. Let M be a 2n-dimensional manifold with symplectic structure ω . Then from non-degeneracy of ω it follows that the 2n-form

$$\wedge^n \omega = \underbrace{\omega \wedge \dots \wedge \omega}_{n \text{ times}}$$

is non-vanishing and hence a volume form. Therefore, M is orientable.

One of the most important examples of a symplectic manifold is the cotangent bundle:

Proposition 12.19. The cotangent bundle of any manifold M has a canonical symplectic structure.

Proof. We will first construct a 1-form λ on T^*M , and then define a symplectic structure ω by $\omega = d\lambda$. To define a 1-form we should describe its value on any tangent vector v. Assume that $v \in T_{(q,p)}T^*M$, where (q,p) is a point in T^*M , i.e. $q \in M$ and $p \in T^*_qM$. To define $\lambda(v)$, consider the projection $\pi: T^*M \to M$ given by $(q,p) \mapsto q$. Then $\lambda(v)$ is defined by

$$\lambda(v) = p(d\pi(v))$$

This is known as the *tautological* 1-form on the cotangent bundle, or *Liouville* 1-form. Let us compute this form in coordinates. Take any local chart (q_1, \ldots, q_n) on M. Then this chart also defines coordinates

on every cotangent space (which is within this chart). Those coordinates (p_1, \ldots, p_n) are, by definition, coordinates of a covector in the basis dq_1, \ldots, dq_n . Then $(p_1, \ldots, p_n, q_1, \ldots, q_n)$ form a local coordinate system on T^*M . In these coordinates, the projection π is given by

$$(p_1,\ldots,p_n,q_1,\ldots,q_n)\mapsto (q_1,\ldots,q_n),$$

so the differential of the projection maps $\partial/\partial p_i$ to 0 and $\partial/\partial q_i$ to $\partial/\partial q_i$. We now take any $v \in T_{(q,p)}T^*M$. This vector can be written as

$$v = \sum_{i=1}^{n} dp_i(v) \frac{\partial}{\partial p_i} + \sum_{i=1}^{n} dq_i(v) \frac{\partial}{\partial q_i}.$$

Therefore,

$$d\pi(v) = \sum_{i=1}^{n} dq_i(v) \frac{\partial}{\partial q_i},$$

and

$$\lambda(v) = p(d\pi(v)) = p\left(\sum_{i=1}^{n} dq_i(v) \frac{\partial}{\partial q_i}\right) = p_i(v) dq_i(v),$$

so in (p,q) coordinates the form λ reads

$$\lambda = \sum_{i=1}^{n} p_i dq_i.$$

We now define the form ω by

$$\omega = d\lambda = \sum_{i=1}^{n} dp_i \wedge dq_i.$$

This 2-form is automatically closed (since it is exact) and from the coordinate representation it is clear that ω is non-degenerate. Therefore, ω is indeed a symplectic structure. It is canonical in a sense that its construction can be performed without coordinates and does not require any additional structures on M.

Corollary 12.20. The cotangent bundle of any manifold has a canonical volume form and is orientable.

Proof. The volume form is defined as $\wedge^n \omega$.

Lecture 13: Vector fields and flows I

Let M be a manifold, and let v be a vector field on M.

Definition 13.1. A parametrized smooth curve $\gamma(t)$ is called an *integral trajectory* of v if

$$\frac{d\gamma}{dt} = v(\gamma(t))$$

for any t for which $\gamma(t)$ is defined.

In other words, at any point of an integral trajectory its velocity vector coincides with the value of the vector field v at that point.

Proposition 13.2. For any $x \in M$ there exists a unique integral trajectory $\gamma(t)$ of v defined for $t \in (-\varepsilon, \varepsilon)$, where $\varepsilon > 0$, such that $\gamma(0) = x$.

Proof. Let x_1, \ldots, x_n be local coordinates near the point x. Then the sought integral trajectory can be written in these coordinates as a vector function $x_1(t), \ldots, x_n(t)$, and the condition on that function to define an integral trajectory of v is equivalent to the system of ODEs

$$\dot{x}_i = v_i(x_1, \dots, x_n),$$

where v_i 's are components of the vector field v in coordinates x_1, \ldots, x_n . Furthermore, the requirement $\gamma(0) = x$ specifies an initial condition for this system of ODEs, so the desired statement follows from the existence and uniqueness theorem for ODEs.

As can be seen from this proof, vector fields can be thought of as systems of ODEs on manifolds. "Solving" a vector field means finding its integral trajectories. In local coordinates this is equivalent to solving actual systems of ODEs.

Definition 13.3. A vector field v is called complete if any its integral trajectory $\gamma(t)$ is defined for any $t \in \mathbb{R}$.

As we know from the theory of ODEs, not every vector field is complete. For example, the vector field defined by the differential equation $\dot{x} = x^2$ (that is $v = x^2 \partial/\partial x$) on \mathbb{R} is not complete. Indeed, its integral trajectories are given by x(t) = 1/(c-t) and hence are not defined for every value of t.

Theorem 13.4. Any vector field on a compact manifold is complete.

The proof is quite straightforward and can be found in most differential geometry textbooks.

Definition 13.5. Let v be a complete vector field on M. The flow of v is a family of transformations $\phi_t \colon M \to M$ defined as follows. Let $\gamma(t)$ be an integral trajectory of v with $\gamma(0) = x$. Then $\phi_t(x) = \gamma(t)$.

One can also define the flow of an incomplete vector field, but that is, in general, only defined locally (i.e. in a neighbourhood of a given point $x \in M$) and only for $t \in (-\varepsilon, \varepsilon)$. The definition of the flow ϕ_t can also be expressed by saying that

$$\frac{d}{dt}\phi_t(x) = v(\phi_t(x)).$$

In order to be equivalent to the above definition, this equation needs to be supplemented by the condition $\phi_0 = id$. This equation is also sometimes written as

$$\frac{d}{dt}\phi_t = v \circ \phi_t.$$

Proposition 13.6. The flow ϕ_t is a diffeomorphism for every t.

Proof. Smoothness of ϕ_t follows from smooth dependence of solutions of ODEs on initial conditions. Furthermore, ϕ_t^{-1} is also smooth since $\phi_t^{-1} = \phi_{-t}$.

Along with the property $\phi_t^{-1} = \phi_{-t}$, we also have that $\phi_t \circ \phi_s = \phi_{s+t}$ and $\phi_0 = id$. In other words, the diffeomorphisms ϕ_t (considered for all values of t) form a group under composition. This group is isomorphic to \mathbb{R} with isomorphism given by $t \mapsto \phi_t$. Due to these properties, the flows of vector fields are also called 1-parametric groups of diffeomorphisms.

We now want to discuss how flows of vector fields act on different geometric objects on M, in particular on tensor fields. Let ξ be a tensor field. Then one can consider the pull-back $\phi_t^*\xi$. It is again a tensor field, of the same type as ξ . This in particular allows us to compute the derivative $d/dt (\phi_t^*\xi)$. This derivative is defined point-wisely: for any x we evaluate the tensor field $\phi_t^*\xi$ at that particular point x. That gives a family of tensors at x depending on the parameter t. This family can be differentiate with respect to t, which produces a new tensor of the same type. **Definition 13.7.** The *Lie derivative* of a tensor field ξ with respect to the vector field v is defined by

$$L_v\xi = \left.\frac{d}{dt}\right|_{t=0} \left(\phi_t^*\xi\right),\,$$

where ϕ_t is the flow of v.

The Lie derivative $L_v \xi$ is a tensor field of the same type as ξ .

Proposition 13.8. The Lie derivative $L_v\xi$ vanishes if and only if ξ is invariant under the flow, i.e.

$$\phi_t^* \xi = \xi. \tag{23}$$

Proof. Assume that ξ is invariant under the flow. Then differentiating (23) with respect to t at t = 0 gives $L_v \xi = 0$. Conversely, assume that $L_v \xi = 0$. To establish (23), we show that

$$\frac{d}{dt}\left(\phi_t^*\xi\right) = 0.$$

for every t. We have

$$\left. \frac{d}{dt} \right|_{t=t_0} \left(\phi_t^* \xi \right) = \left. \frac{d}{dt} \right|_{t=0} \left(\phi_{t+t_0}^* \xi \right).$$

Using the group property of the flow, this can be further rewritten as

$$\frac{d}{dt}\Big|_{t=0} \left(\phi_{t_0}^* \phi_t^* \xi\right) = \phi_{t_0}^* \left(\frac{d}{dt}\Big|_{t=0} \left(\phi_t^* \xi\right)\right) = \phi_{t_0}^* \left(L_v \xi\right) = 0,$$

as desired.

Lecture 14: Vector fields and flows II

In the case when ξ itself is a vector field $\xi = w$, the condition $L_v w = 0$ is also equivalent to the fact that the flows of v and w commute:

Proposition 14.1. Let v, w be vector fields on a manifold M. Then $L_v w = 0$ if and only if the flow $\{\phi_t \mid t \in \mathbb{R}\}$ of v commutes with the flow $\{\psi_t \mid t \in \mathbb{R}\}$ of w, which means that

$$\phi_t \circ \psi_{t'} = \psi_{t'} \circ \phi_t \quad \forall t, t' \in \mathbb{R}.$$

Proof sketch. As we know, $L_v w = 0$ is equivalent to the fact that $\phi_t^* w = w$ for any $t \in \mathbb{R}$. The latter in turn is equivalent to ϕ_t preserving integral trajectories of w, i.e. if $\gamma(t)$ is an integral trajectory of w, then so is $\phi_t(\gamma(t))$. Let us, for example, show that if $\phi_t^* w = w$ for any $t \in \mathbb{R}$ then ϕ_t preserves integral trajectories. Let $\gamma(t)$ be an integral trajectory of w. We want to show that $\phi_t(\gamma(t))$ is also an integral trajectory. To that end, observe that the derivative

$$\frac{d}{dt}\phi_t(\gamma(t))$$

is equal to the push-forward of the tangent vector $\gamma'(t) = w(\gamma(t))$ by the map ϕ_t , which is the same as the pull-back by γ_{-t} . Also using that $\phi_t^* w = w$ for any t, we get

$$\frac{d}{dt}\phi_t(\gamma(t)) = (\phi_{-t}^*w)(\gamma(t)) = w(\gamma(t)),$$

which means that $\phi_t(\gamma(t))$ is indeed an integral trajectory of w.

It now suffices to show that ϕ_t preserves integral trajectories of w if and only if it commutes with the flow of w. This is almost a tautology. Indeed, equation $\phi_t \circ \psi_{t'} = \psi_{t'} \circ \phi_t$ means that

$$\phi_t(\psi_{t'}(x)) = \psi_{t'}(\phi_t(x)) \quad \forall x \in M.$$
(24)

But $\psi_{t'}(x)$ considered as a function of t' (for fixed x) is an integral trajectory of w, and all integral trajectories of w are, by definition, of this form. Similarly, $\psi_{t'}(\phi_t(x))$ is also an integral trajectory of w (for fixed x and t). So, if the flows commute, then ϕ_t maps integral trajectories of w to integral trajectories of w. Conversely, if ϕ_t maps integral trajectories of w to integral trajectories of w, then it must map a trajectory starting at x to the trajectory starting at $\phi_t(x)$. But these trajectories are given by $\psi_{t'}(x)$ and $\psi_{t'}(\phi_t(x))$, which gives (24).

We will now discuss hot to compute the Lie derivative $L_v \xi$ for particular types of tensors. We start with functions:

Proposition 14.2. If f is a smooth function, then $L_v f$ is equal to the directional derivative of f along v:

$$L_v f = df(v). \tag{25}$$

Proof. We have

$$(L_v f)(x) = \left. \frac{d}{dt} \right|_{t=0} (\phi_t^* f)(x) = \left. \frac{d}{dt} \right|_{t=0} f(\phi_t(x)) = df\left(\left. \frac{d}{dt} \right|_{t=0} \phi_t(x) \right) = df(v(x)),$$

q.e.d.

Corollary 14.3. For a function f and vector field v, the following conditions are equivalent:

- 1. The directional (equivalently, Lie) derivative of f along v vanishes: $L_v f = df(v) = 0$.
- 2. The function f is preserved by the flow of v.
- 3. The function f is constant along integral trajectories of v.

Proof. The first two conditions are equivalent by Proposition 13.8. The last two conditions are equivalent because

$$(\phi_t^* f)(x) = f(\phi_t(x)),$$

so f is preserved by the flow if and only if it is constant along curves of the form $\phi_t(x)$ (for fixed x). But such curves are, by definition, exactly the integral trajectories of v.

Formula (25) can also be rewritten as

$$L_v f = i_v df, \tag{26}$$

where the operation i_v is defined for an arbitrary k-form as follows: if ξ is a k-form, then $i_v \xi$ is a (k-1)-form given by

$$i_v \xi(w_1, \dots, w_{k-1}) = \xi(v, w_1, \dots, w_{k-1}).$$

For k-forms the formula for the Lie derivative is similar to (26) but has one more additional term:

Proposition 14.4 (Cartan's magic formula). Let ξ be a differential form, and v be a vector field. Then

$$L_v\xi = i_v d\xi + di_v\xi.$$

In other words, for differential forms we have

$$L_v = i_v d + di_v.$$

Proof sketch. Using the fact that pull-backs commute with wedge products, one establishes the product rule for the Lie derivative: $L_v(\alpha \wedge \beta) = (L_v\alpha) \wedge \beta + \alpha \wedge (L_v\beta)$. The right hand side of Cartan's formula can also be shown to obey this rule, which reduces the proof to the case of $\xi = dx_i$, where x_i is a local coordinate. But since pull-backs also commute with the exterior derivative, it follows that $L_v d = dL_v$, and thus

$$L_v(dx_i) = d(L_v x_i) = d(i_v dx_i),$$

which coincides with Cartan's formula in the case $\xi = dx_i$.

We now establish a formula for the Lie derivative of a vector field. To that end, we will need the notion of the *commutator (Lie bracket)* of vector fields:

Definition 14.5. The Lie bracket of two vector fields v, w is the unique vector field [v, w] such that

$$L_{[v,w]}f = L_v L_w f - L_w L_v f$$

for any function f.

Proposition 14.6. Such a vector field [v, w] exists and it is indeed unique.

Proof. If there were two such vector fields u and u' then we would have $L_u f = L_{u'} f$ for any f, i.e.

$$df(u-u')=0.$$

But any cotangent vector at any point is the differential of a suitable function, so all cotangent vectors vanish on u - u', which means that u - u' = 0 and u = u'. This proves uniqueness. As for existence, it is sufficient to verify it locally, because then by the uniqueness part so-defined local vector fields can be patched into a global one. In local coordinates we have

$$L_{v}L_{w}f - L_{w}L_{v}f = \sum_{i,j} v_{i}\frac{\partial}{\partial x_{i}} \left(w_{j}\frac{\partial f}{\partial x_{j}}\right) - w_{i}\frac{\partial}{\partial x_{i}} \left(v_{j}\frac{\partial f}{\partial x_{j}}\right).$$

All terms containing second derivatives cancel out, and we end up with

$$L_v L_w f - L_w L_v f = \sum_{i,j} v_i \frac{\partial w_j}{\partial x_i} \frac{\partial f}{\partial x_j} - w_i \frac{\partial v_j}{\partial x_i} \frac{\partial f}{\partial x_j},$$

which is the derivative of f along the vector field

$$\sum_{i,j} \left(v_i \frac{\partial w_j}{\partial x_i} - w_i \frac{\partial v_j}{\partial x_i} \right) \frac{\partial}{\partial x_j}.$$

So, the vector field [v, w] indeed exists and it is given in local coordinates by the formula

$$[v,w] = \sum_{i,j} \left(v_i \frac{\partial w_j}{\partial x_i} - w_i \frac{\partial v_j}{\partial x_i} \right) \frac{\partial}{\partial x_j}.$$
Example 14.7. In \mathbb{R}^2 we have

$$\left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right] = 0,$$

since

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y}f = \frac{\partial}{\partial y}\frac{\partial}{\partial x}f$$

for any function f.

Proposition 14.8. For any vector fields v, w we have

$$L_v w = [v, w].$$

Proof sketch. From the definition of the Lie derivative it easily follows that for any vector field v and any 1-form α we have

$$L_v(i_w\alpha) = i_{L_vw}\alpha + i_w(L_v\alpha)$$

which can also be rewritten as

$$i_{L_v w} \alpha = L_v(i_w \alpha) - i_w(L_v \alpha).$$

Taking $\alpha = df$ we get

$$i_{L_vw}df = L_v(i_w df) - i_w(L_v df) = L_v(L_w f) - i_w(L_v df)$$

Using Cartan's magic formula, the latter term can be rewritten as

$$i_w(di_v df) = L_w L_v df,$$

 \mathbf{SO}

$$i_{L_vw}df = L_v(L_wf) - L_w(L_vdf),$$

which means that $L_v w = [v, w]$, as desired.

Corollary 14.9. [v, w] = 0 if and only the flows of v and w commute.

Lecture 15: Hamiltonian vector fields

We now discuss the notion of a Hamiltonian vector field on a symplectic manifold (cf. Lecture 7).

Definition 15.1. A vector field v on a symplectic manifold (M, ω) is a *Hamiltonian vector field* if there exists a smooth function H on M (called the *Hamiltonian* of the vector field v) such that

$$i_v\omega = dH$$

Explicitly, the Hamiltonian vector field with Hamiltonian H is given by $\omega^{-1}dH$, where ω is regarded as an operator from the tangent space to the cotangent space. The Hamiltonian vector field corresponding to the function H is denoted by X_H or sgrad H. Thus, we have

$$X_H = \operatorname{sgrad} H = \omega^{-1} dH.$$

Proposition 15.2. The flow of the Hamiltonian vector field $X_H = \omega^{-1} dH$ preserves both the function H and the symplectic form ω .

Proof. We have

$$L_{X_H}H = i_{X_H}dH = i_{X_H}i_{X_H}\omega = \omega(X_H, X_H) = 0$$

so H is indeed preserved. Further, by Cartan's magic formula, we get

$$L_{X_H}\omega = i_{X_H}d\omega + di_{X_H}\omega,$$

but ω is closed so this rewrites as

$$di_{X_H}\omega = d(dH) = 0,$$

which means that the symplectic form ω is preserved as well.

Corollary 15.3. The flow of a Hamiltonian vector field consists of symplectic diffeomorphisms: $\phi_t^* \omega = \omega$.

In fact, preservation of the symplectic form ω is almost equivalent to the Hamiltonian property. Namely, one has the following.

Proposition 15.4. Let v be a vector field on a symplectic manifold (M, ω) . Then $L_v \omega = 0$ if and only if v is locally Hamiltonian, which means that it can be written as X_H in a neighbourhood of any point in M. Furthermore, if $H^1(M, \mathbb{R}) = 0$, then $L_v \omega = 0$ if and only if v is Hamiltonian.

Proof. We have $L_v \omega = d(i_v \omega)$, so $L_v \omega = 0$ if and only if the 1-form $i_v \omega$ is closed. But closed forms can be equivalently characterized as locally exact forms, so $L_v \omega = 0$ if and only if $i_v \omega = dH$ locally, which means that $v = X_H$. Furthermore, if $H^1(M, \mathbb{R}) = 0$, then closed 1-forms and exact 1-forms are the same, so in that case $L_v \omega = 0$ if and only if v is Hamiltonian.

Vector fields v such that $L_v \omega = 0$ are known as *symplectic*. They can be equivalently characterized as vector fields whose flows consist of symplectic diffeomorphisms. The above proposition shows that a vector field is symplectic if and only if it is locally Hamiltonian. It is easy to construct an example of a vector field which is symplectic but not globally Hamiltonian: take a closed non-exact 1-form α and define $v = \omega^{-1} \alpha$. For example, one can take α to be the differential of the polar angle function in \mathbb{R}^2 without the origin.

From Proposition 15.2 we also get the following.

Corollary 15.5 (Liouville's theorem). Any Hamiltonian vector field preserves the volume form $\Omega = \wedge^n \omega$.

Proof. Since the flow ϕ_t of a Hamiltonian vector field preserves ω , it also preserves $\wedge^n \omega$. Alternatively, one can use the product rule for the Lie derivative:

$$L_{X_H}(\wedge^n \omega) = (L_{X_H}\omega) \wedge \omega \wedge \omega \wedge \cdots \wedge \omega + \omega \wedge (L_{X_H}\omega) \wedge \omega \wedge \cdots \wedge \omega + \cdots = 0.$$

We now discuss examples of Hamiltonian vector fields.

Example 15.6. Let (M, ω) be a 2-dimensional symplectic manifold (equivalently, a surface with an area form). Let also H be a function on M. Then the level sets of H are 1-dimensional. The Hamiltonian vector field X_H is tangent to those level sets (by Proposition 15.2), which means that in the 2D case the trajectories of X_H are almost the same as level sets of H. More precisely, we have the following:

Proposition 15.7. 1. Each connected component of each regular level set of H is a trajectory of X_H .

2. For an arbitrary level set $H^{-1}(c) \subset M$, each critical point $x \in H^{-1}(c)$ is a 1-point trajectory (an equilibrium point) of X_H , and each connected component of the set of non-singular points in $H^{-1}(c)$ is a trajectory of X_H .

Proof. The first statement is a particular case of the second one, so we only prove the second statement. By definition, we have $X_H = \omega^{-1} dH$, so critical points of H are indeed equilibrium points of X_H . Consider now the set N of non-critical points in the given level set of H. At such points we have $X_H \neq 0$, so the set N is a disjoint union of non-trivial (i.e. non-stationary) trajectories of X_H . Furthermore, since both N and the trajectories are 1-dimensional manifolds, each trajectory is open N. Also, each trajectory is by definition connected. So, N is a disjoint union of open connected subsets, which means that these subsets (the trajectories of X_H) are connected components of N, q.e.d.

In \mathbb{R}^2 (as well as in the sphere S^2), the class of Hamiltonian vector fields coincides with symplectic (=area-preserving) vector fields, by Proposition 15.4. If we endow \mathbb{R}^2 we the standard area form $dx \wedge dy$, then the area-preserving condition for a vector field $v = (v_x, v_y)$ is

$$L_v(dx \wedge dy) = 0.$$

Explicitly, we have

$$L_v(dx \wedge dy) = di_v(dx \wedge dy) = d(v_x dy - v_y dx) = \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}\right) dx \wedge dy$$

So, a vector field in \mathbb{R}^2 is Hamiltonian if and only if it has zero *divergence*

$$\operatorname{div} v = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}.$$

In particular, any divergence-free (=area-preserving) vector field in \mathbb{R}^2 automatically has a first integral, namely the Hamiltonian function. For more complicated surfaces this does not have to be so, as shown by the following example. Consider the torus T^2 with coordinates ϕ, ψ , defined modulo 2π . Then the vector field

$$\begin{cases} \dot{\phi} = 1, \\ \dot{\psi} = \alpha, \end{cases}$$

where α is constant, has zero divergence and thus preserves the area form $d\phi \wedge d\psi$. However, if α is irrational, then this vector field does not have any first integrals. Indeed, consider the trajectory of this vector field (i.e. a solution of the above system of ODEs) with initial condition $\phi = 0$, $\psi = 0$. It is explicitly given by

$$\begin{cases} \phi(t) = t \mod 2\pi, \\ \psi(t) = \alpha t \mod 2\pi. \end{cases}$$

Consider also the circle $\phi = c$. The above trajectory intersects this circle when $t = c + 2\pi k$. The corresponding values of ψ are $\alpha(c + 2\pi k) = \alpha c + 2\alpha \pi k$, which means that consecutive intersection points are obtained from each other by means of a rotation with the angle $2\alpha\pi$. This angle is not a rational multiple of π , so the intersection points are dense in the circle $\phi = c$ (see Lecture 3). But since our trajectory has a dense intersection with every circle $\phi = c$, and the union of such circles is the whole torus, it follows that the trajectory is dense in the torus. So, any continuos function which is constant along the trajectory must be constant on the whole torus. But that means that there are no non-constant first integrals, as desired.

The trajectories of the above vector field are known as irrational windings of the torus. One can similarly show that each trajectory is dense. This can be also derived from the fact that rotations (translations) of the torus take trajectories to trajectories, because the vector field defining the irrational winding commutes (has zero Lie bracket) with $\partial/\partial \phi$ and $\partial/\partial \psi$.

Example 15.8. Our next example is *natural mechanical systems*. A mechanical system is called natural if all exterior forces are potential. An example is a motion of an object under gravity.

The configuration space M of a mechanical system is the space of all its possible positions, endowed with the topology given by the natural notion of "closeness" of two positions. This topology can be rigorously defined as the initial topology with respect to the maps $\phi_i \colon M \to \mathbb{R}^3$ which compute the position of the *i*'th particle of the system for a given position of the whole system. In more down-to-earth terms, if a mechanical system consists of n particles, then to each configuration we can assign a point in \mathbb{R}^{3n} whose coordinates are positions of all the *n* particles. The configuration space is then defined as the subset of \mathbb{R}^{3n} defined by the constraints on the positions of particles. Furthermore, one can extend this definition to the case of infinitely many particles and describe the configuration space of a system as a subset of an infinite-dimensional space. However, this does not seem to be a good definition from a practical point of view. In practice, it is not necessary to consider positions of all particles of the system to describe the configuration space. It is sufficient to consider a subset of particles with the property that their positions uniquely determine the position of the whole system. As an example, consider the pendulum, i.e. a rigid rod fixed at one of its ends. Although it consists of infinitely many particles, its position is uniquely described by the position of its free endpoint. So, the configuration space is the space of positions of that point, subject to a constraint: the distance from the free end to the immovable end is fixed. This constraint describes either a circle (if the pendulum can only move in two dimensions) or a sphere (three-dimensional, or *spherical* pendulum). So, the configuration space of the pendulum is a circle, while the configuration space of the spherical pendulum is the sphere S^2 .

In good examples, the configuration space M is a manifold. Its dimension is called the number of *degrees of freedom* of the system. It can be thought of as a number of parameters needed to specify a position of the system. Any motion of a system is described by a curve in its configuration space. An infinitesimal motion is thus a tangent vector to M. The space al all tangent vectors, i.e. the tangent bundle TM of the configuration manifold M, is called the *phase space*. Each infinitesimal motion $v \in TM$ of the system determines the velocity of each individual particle. These velocities can be computed by applying the differentials of the above "position calculating" maps ϕ_i to v. Since the differential is a linear map, the velocity of each particle is a linear function of the velocity $v \in TM$ of the system.

The kinetic energy of the system is defined by

$$K = \frac{1}{2} \sum m_i v_i^2,$$

where the sum is taken over all particles, m_i is the mass of the *i*'th particle, and v_i is its velocity. In the case of infinitely many particles, this is replaced by the integral

$$K = \frac{1}{2} \int v^2 dm.$$

Since the velocity of each particle is a linear function of the velocity v of the system, K can be viewed as a quadratic form on tangent vectors to M. It is positive definite, so it defines an inner product on each tangent space. It also depends smoothly on the point in M. In other words, the *kinetic energy is a Riemannian metric on the configuration space*.

For convenience purposes, as a Riemannian metric on M we take *twice* the kinetic energy:

$$\langle v, v \rangle = 2K(v).$$

This metric defines a map from the tangent bundle TM to the cotangent bundle T^*M given by $v \mapsto \langle v, * \rangle$. This mapping is a particular case of a *Legendre transform* in mechanics. The image of the velocity v under the Legendre transform is denoted by p and is called the *momentum*. The space of momenta is the cotangent bundle T^*M . Since it is identified with TM by means of the Legendre transform, it is also called the phase space (sometimes only T^*M , but not TM, is called the phase space).

We also have the *potential energy* function U on M, defined by summing up (or integrating) the potential over all particles. The *total energy* of the system is H = K + U. It is a function on the tangent bundle TM. We can also transport this function to the cotangent bundle using the Legendre transform. The resulting function, also denoted H, is given by

$$H = \frac{1}{2} \langle p, p \rangle + U(q).$$

Here $p \in T_q^*M$. The inner product $\langle p, p \rangle$ on momenta is defined by inverting the inner product on velocities (i.e. inverting the corresponding map $TM \to T^*M$ and interpreting the resulting map $T^*M \to TM$ as a bilinear form on cotangent vectors). The physical meaning of this function is still total energy, but expressed in terms of momenta.

Theorem 15.9. The motion of a natural mechanical system is described by a Hamiltonian vector field on the phase T^*M . The corresponding Hamiltonian function is the total energy H, while the symplectic structure is the canonical symplectic structure on the cotangent bundle.

For the motion of a single particle in \mathbb{R}^n , we derived this result in Lecture 11.

Example 15.10. One more class of examples of Hamiltonian systems is provided by *geodesic flows*. Let M be a Riemannian manifold, i.e. a manifold endowed with an inner product on each tangent space which depends smoothly on the point. Then one can define the length of a curve $\gamma(t)$, $t_1 \leq t \leq t_2$ on M by

$$\int_{t_1}^{t_2} \sqrt{\langle \gamma'(t), \gamma'(t) \rangle} dt.$$

The length of a curve clearly does not depend on a parametrization.

Geodesics on M are defined as curves $\gamma(t)$ with the following property: for any real numbers t_1, t_2 sufficiently close to each other, the arc of $\gamma(t)$ corresponding to $t_1 \leq t \leq t_2$ is the shortest curve joining the points $\gamma(t_1), \gamma(t_2)$. For example, geodesics in \mathbb{R}^2 are straight lines, while geodesics on the sphere S^2 are great circles, i.e. sections of the sphere by 2-planes passing through its center. Note that this definition of geodesics does not assume any specific parametrization. A more precise definition of geodesics includes an additional requirement $\langle \gamma'(t), \gamma'(t) \rangle = \text{const}$, i.e. the velocity vector of the curve has constant magnitude. For surfaces in \mathbb{R}^n (with Riemannian metric induced by the inner product in the ambient space), geodesics can be alternatively defined as curves whose acceleration $\gamma''(t)$ is orthogonal to the surface.

Geodesics are described by a second order differential equation on M, i.e. a vector field on the tangent bundle TM. However, it turns out to be more convenient to write the geodesic equation on the cotangent bundle T^*M . To that end, one identifies the tangent and cotangent bundles using the Riemannian metric (which is the same as the Legendre transform in Example 15.8). This gives a vector field on T^*M , known as the geodesic flow (this term may also refer to the flow of that vector field).

Theorem 15.11. The geodesic flow is a Hamiltonian vector field on T^*M . The corresponding Hamiltonian function is the "kinetic energy"

$$K = \frac{1}{2} \langle p, p \rangle,$$

while the symplectic structure is the canonical symplectic structure on the cotangent bundle.

As in the previous example, the inner product $\langle p, p \rangle$ on momenta is defined by inverting the inner product on velocities.

By comparing this with the previous example, we see that the motion of any *free* mechanical system (which means no external forces) is given by geodesics on the configuration space. The corresponding metric is twice the kinetic energy of the system. For example, the spherical pendulum with no gravity moves along great circles (the kinetic energy of the pendulum coincides, up to a constant factor, with the standard metric on the sphere; this is because both of these metrics are rotationally invariant). Thus, free physical systems have a geometric interpretation. Conversely, geodesics on a surface embedded in the Euclidian space has a physical interpretation: they are trajectories of a particle that is constrained to a given surface. This in particular follows from the definition of geodesics as curves whose acceleration is orthogonal to the surface. Indeed, if a particle is moving freely on the surface, then the only force acting on it is the reaction force which is orthogonal to the surface. So, by Newton's second law, the acceleration is orthogonal to the surface as well, which means that trajectories of the particle are indeed geodesics (and the other way around).

Lecture 16: Poisson brackets

Definition 16.1. Let M be a symplectic manifold with symplectic form ω . The *Poisson bracket* of two smooth function f, g on M is a smooth function $\{f, g\}$ defined by⁶

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

Using that the 1-form $\omega(X_f,...)$ is, by definition of X_f , equal to df, we can rewrite the formula for the Poisson bracket as

$$\{f,g\} = -\omega(X_f, X_g) = \omega(X_g, X_f) = dg(X_f) = L_{X_f}g$$

Also, we have

$$dg(X_f) = dg(\omega^{-1}(df)) = \omega^{-1}(df, dg),$$

where in the latter formula ω^{-1} is interpreted as a form on cotangent vectors. This form is known as the *Poisson tensor*. We will denote it by π . To summarize, we have the following equivalent definitions of the Poisson bracket:

$$\{f,g\} = -\omega(X_f, X_g) = L_{X_f}g = \omega^{-1}(df, dg).$$

The formula $\{f, g\} = L_{X_f}g$ provides an alternative formulation of Hamiltonian dynamics: the evolution of any smooth function g along the flow of X_f is given by

$$\frac{dg}{dt} = \{f, g\}.$$

Here the derivative in the left hand side is the derivative along the trajectories of X_f , i.e. it is the time derivative of the function g(x(t)), where the point x(t) moves along the vector field X_f . In particular, we have the following.

Proposition 16.2. A function g is a first integral of the Hamiltonian system X_f (i.e. constant along trajectories of X_f) if and only if $\{f, g\} = 0$.

When functions f, g satisfy $\{f, g\} = 0$, one says that they are *in involution*.

We now describe algebraic properties of Poisson brackets.

Proposition 16.3. The Poisson bracket operation on a symplectic manifold has the following properties:

⁶Note that some authors define the Poisson bracket with an opposite sign

- 1. It is skew-symmetric, i.e. $\{g, f\} = -\{f, g\}$.
- 2. It is bilinear over real numbers. For example, linearity in the first argument means that $\{af+bg,h\} = a\{f,h\} + b\{g,h\}$, where f,g,h are functions, and a,b are real numbers.
- 3. It satisfies the product rule in both arguments, e.g. $\{fg,h\} = f\{g,h\} + g\{f,h\}$.
- 4. It satisfies the Jacobi identity

$$\{\{f,g\},h\} + \{\{g,h\},f\} + \{\{h,f\},g\} = 0.$$
(27)

Proof. The only property that does not follow right from the definition is the Jacobi identity. To prove the latter, we first rewrite it, using skew-symmetry, as

$$\{\{f,g\},h\} = \{f,\{g,h\}\} - \{g,\{f,h\}\},\$$

Using the property $\{f, g\} = L_{X_f}g$, this can be further rewritten as

$$L_{X_{\{f,g\}}}h = L_{X_f}L_{X_g}h - L_{X_g}L_{X_f}h.$$

But

$$L_{X_f}L_{X_g}h - L_{X_g}L_{X_f}h = L_{[X_f, X_g]}h,$$

so the Jacobi identity rewrites as

$$X_{\{f,g\}} = [X_f, X_g].$$
 (28)

To prove this identity, we will use the product rule for the Lie derivative: for any vector fields v, w and any k-form ξ we have

$$L_v(\xi(w)) = (L_v\xi)(w) + \xi(L_vw)$$

where $\xi(w) = i_w \alpha$. Using that $L_v w = [v, w]$, this can also be written as

$$L_v i_w = i_w L_v + i_{[v,w]},$$
(29)

where both sides should be understood as operations on differential forms. Expressing $i_{[v,w]}$, we also find that

$$i_{[v,w]} = L_v i_w - i_w L_v,$$

which is sometimes abbreviated as

$$i_{[v,w]} = [L_v, i_w].$$

This product rule follows from the product rule for the usual derivative, because the Lie derivative is defined as the time derivative of $\phi_t^*(\ldots)$.

Using (29), we find that

$$i_{[X_f,X_g]}\omega = L_{X_f}i_{X_g}\omega - i_{X_g}L_{X_f}\omega$$

But since Hamiltonian vector fields preserve the symplectic form, we have $L_{X_f}\omega = 0$, so

$$i_{[X_f,X_g]}\omega = L_{X_f}i_{X_g}\omega = L_{X_f}dg = di_{X_f}dg = d(L_{X_f}g) = d\{f,g\},$$

which is equivalent to (28). Thus, (28) and the Jacobi identity are proved.

Corollary 16.4. If g, h are first integrals of X_f , then so is their Poisson bracket $\{g, h\}$.

Proof. Since g and h are first integrals, we have $\{f, g\} = 0$ and $\{f, h\} = 0$, so the first and the last term on the left hand side of (27) vanish. Therefore, the second term also vanishes, i.e. $\{\{g, h\}, f\} = 0$, which is equivalent to saying that $\{g, h\}$ is a first integral of X_f .

Corollary 16.5. If g is a first integral of X_f , then X_g is a symmetry of X_f , i.e. $[X_f, X_g] = 0$.

Proof. Indeed, from (28) we have that $[X_f, X_q] = 0$ whenever $\{f, g\} = 0$.

Lecture 17: Liouville integrable systems on symplectic manifolds

Definition 17.1 (Continuous time Liouville integrable system). Let X_H be a Hamiltonian vector field on a 2*n*-dimensional symplectic manifold (M, ω) . Then X_H is called *completely integrable* (or *Liouville integrable*, or just *integrable*) if it has $n = \frac{1}{2} \dim M$ first integrals F_1, \ldots, F_n which are:

- 1. Functionally independent, which means that the differentials dF_1, \ldots, dF_n are linearly independent almost everywhere on M.
- 2. In involution, i.e. $\{F_i, F_j\} = 0$ for all i, j = 1, ..., n.

Let us comment on these conditions. First of all, since the Hamiltonian H itself is a first integral of X_H , one can always assume that $F_1 = H$. If this is the case, then the condition $\{F_1, F_i\} = 0$ is satisfied automatically. In particular, the $\{F_i, F_j\} = 0$ condition is vacuous for $2n \leq 4$. If $2n \geq 6$, then this condition becomes non-trivial and it is equivalent to requiring that each of the functions F_1, \ldots, F_n is a first integral for each of the Hamiltonian vector fields X_{F_1}, \ldots, X_{F_n} .

The condition that the differentials of F_1, \ldots, F_n are generically independent roughly speaking means that none of the F_i 's can be expressed as a function of the other ones. Indeed, if, say, $F_1 = G(F_2, \ldots, F_n)$, then

$$dF_1 = \sum_{i=2}^n \frac{\partial G}{\partial x_i}(F_2, \dots, F_n)dF_i,$$

so the differentials of F_1, \ldots, F_n are linearly dependent. Conversely, it is not hard to show that linear dependence of differentials implies that one of the F_i 's can be locally expressed as a function of the other ones.

One can also reformulate the functional independence condition in terms of the moment map, that is the map $F: M \to \mathbb{R}^n$ given by $F(x) = (F_1(x), \ldots, F_n(x))$. Clearly, the differentials of F_1, \ldots, F_n are linearly independent if and only if the differential of the moment map has maximal rank. So, dF_1, \ldots, dF_n are linearly independent almost everywhere on M if and only if the moment map is generically a submersion. In particular, generic level sets of the moment map of an integrable system on M are half-dimensional submanifolds of M.

Finally, notice that since dF_1, \ldots, dF_n are linearly independent almost everywhere, the same is true for the Hamiltonian vector fields X_{F_1}, \ldots, X_{F_n} . Furthermore, $\omega(X_{F_i}, X_{F_j}) = -\{F_i, F_j\} = 0$, so X_{F_1}, \ldots, X_{F_n} span an isotropic subspace in every tangent space of M. At generic points this subspace has dimension $n = \frac{1}{2} \dim M$ and is, therefore, a Lagrangian (=maximal isotropic) subspace. This in particular means that $n = \frac{1}{2} \dim M$ is the maximal number of independent first integrals in involution. Indeed, if we had more, then their Hamiltonian vector fields would span an isotropic subspace of dimension $> \frac{1}{2} \dim M$, which is not possible. So, completely integrable systems are, in a sense, "maximally integrable": they have a maximal possible number of first integrals in involution. There do exist systems which have more integrals, but in this case some of the integrals must have non-zero pairwise Poisson brackets. Such systems are known as *superintegrable*, or *degenerately integrable* (we will discuss a precise definition later on).

In the discrete setting, the definition of Liouville integrability is the same, with a map T instead of a vector field X_H . The Hamiltonian property of X_H is replaced by the assumption that the map T preserves the symplectic structure: $T^*\omega = \omega$. This is indeed a natural analog of the Hamiltonian property, because, as we know, Hamiltonian vector fields can be viewed as infinitesimal symplectic diffeomorphisms, and, conversely, any vector field whose flow preserves the symplectic structure is locally Hamiltonian.

Definition 17.2 (Discrete time Liouville integrable system). Let $T: M \to M$ be a symplectic diffeomorphism of a 2*n*-dimensional symplectic manifold (M, ω) (i.e. $T^*\omega = \omega$). Then T is called completely integrable (or Liouville integrable, or just integrable) if it has $n = \frac{1}{2} \dim M$ functionally independent first integrals F_1, \ldots, F_n in involution.

This definition in particular means that to every discrete time integrable system one can associate a continuous time system: as a Hamiltonian of such a system, one can take any of the integrals F_i . Conversely, one can obtain a discrete integrable system out of a continuous one by taking the time 1 shift along the trajectories of X_H .

Also note that while any continuous time Hamiltonian system on a 2D manifold is automatically integrable (because the Hamiltonian function is a first integral), there is no reason for the corresponding statement in the discrete setting to be true. In other words, an area-preserving map of a surface to itself does not have to be integrable.

The main structural result about integrable systems is the *Arnold-Liouville theorem*. We state it simultaneously in the continuous and discrete settings. The continuous version is by now considered a classical result, while the discrete version is relatively recent and belongs to A. Veselov [20].

Theorem 17.3 (Arnold-Liouville theorem). Consider a continuous or discrete time integrable system on a 2n-dimensional symplectic manifold (M, ω) . Let F_1, \ldots, F_n be its independent first integrals in involution, and let $F = (F_1, \ldots, F_n)$ be the corresponding moment map. Consider a level set $M_c = F^{-1}(c)$ of the moment map (i.e. a joint level set $\{x \in M \mid F_i(x) = c_i\}$), and assume that M_c is regular (i.e. dF has maximal rank on M_c), compact, and connected. Then:

- 1. The submanifold M_c is diffeomorphic to an n-dimensional torus.
- 2. The dynamics on M_c is quasi-periodic. In other words, there exist 2π -periodic coordinates ϕ_1, \ldots, ϕ_n on the torus M_c such that the continuous dynamics on M_c is given by

$$\dot{\phi}_i = k_i$$

while the discrete dynamics is given by

$$\phi_i \mapsto \phi_i + l_i.$$

3. The restriction of the symplectic structure ω to M_c vanishes (half-dimensional submanifolds with this property are called **Lagrangian submanifolds**; a submanifold is Lagrangian if and only if its tangent space at every point is Lagrangian). Furthermore, one can choose the angle coordinates ϕ_1, \ldots, ϕ_n on the torus M_c as well as on nearby level sets, and also choose new first integrals $s_i = s_i(F_1, \ldots, F_n)$ such that

$$\omega = \sum_{i=1}^{n} ds_i \wedge d\phi_i$$

The variables s_i are known as action variables, while ϕ_i 's are angle variables. Together they are called action-angle variables.

4. The dynamics on M_c (i.e. the integral trajectories in the continuous case and iterates of the map in the discrete case) can be explicitly described by quadratures.

The connectedness assumption of the theorem is not crucial. If the level set M_c is disconnected, then each of the connected components of M_c is a torus, while the rest of the theorem stays intact. In the discrete setting one needs an additional assumption that each connected component is invariant under the map T, otherwise one cannot restrict the map to M_c . In any case, this assumption is always satisfied for a sufficiently high power of T. The compactness assumption can also be weakened, as we will se further.

As a corollary of the Arnold-Liouville theorem, the phase space of an integrable system (i.e. the symplectic manifold where the system is defined) is almost everywhere foliated by Lagrangian tori, and the dynamics on these tori is quasi-periodic.

Exercises.

- 1. Let f, g be two functions in \mathbb{R}^2 such that df and dg are linearly dependent at every point. Show that there exists an open subset $U \subset \mathbb{R}^2$ in which either f is a function of g, or g is a function of f.
- 2. Let $\phi_i = k_i$ be a vector field on a torus. Prove that the trajectories of this vector field are periodic if and only if the vector $k = (k_1, \ldots, k_n)$ can be expressed as $k = \lambda z$, where $\lambda \in \mathbb{R}$, and $z \in \mathbb{Z}^n$ is a vector with integer components.
- 3. What is the condition on l_i 's under which the map $\phi_i \mapsto \phi_i + l_i$ of the torus to itself has periodic orbits?
- 4. Let s_i be one of the action variables. Prove that the trajectories of the vector field X_{s_i} are 2π -periodic (hint: compute the Poisson bracket $\{s_i, \phi_j\}$).
- 5. Assume that M_c is compact but possibly disconnected. Prove that in the discrete case there exists k > 0 such that T^k maps every connected component of M_c to itself.

Lecture 18: Proof of the first part of the Arnold-Liouville theorem: topology

In this lecture we will prove the topological part of the Arnold-Liouville theorem, namely that a compact connected regular joint level set M_c of n functions F_1, \ldots, F_n in involution on a 2n-dimensional symplectic manifold M is diffeomorphic to a n-dimensional torus. First of all, it is clear that M_c is indeed a smooth n-dimensional submanifold of M, because it is a regular joint level set of n smooth functions. What is non-trivial is that this submanifold is diffeomorphic to a torus. To prove that, observe the following:

- 1. Regularity of M_c means that the differentials dF_1, \ldots, dF_n are linearly independent at every point of M_c . Therefore, the vector fields $X_{F_i} = \omega^{-1} dF_i$ are also linearly independent at every point of M_c .
- 2. We have $L_{X_{F_i}}F_j = \{F_i, F_j\} = 0$, so every vector field X_{F_i} preserves the functions F_1, \ldots, F_n and thus restricts (i.e. is tangent) to the joint level set M_c of those functions.
- 3. $[X_{F_i}, X_{F_i}] = X_{\{F_i, F_i\}} = 0$, i.e. the vector fields X_{F_i} commute.

Thus, M_c is a compact connected *n*-dimensional manifold that admits *n* pairwise commuting vector fields X_{F_1}, \ldots, X_{F_n} that are linearly independent at every point. We will show that any manifold with this

property is a torus. To that end, denote by $\phi_{t_i}^{(i)}$ the time t_i shift along the trajectories of the vector field X_{F_i} . Further, for any $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$ let

$$\phi_t = \phi_{t_1}^{(1)} \circ \cdots \circ \phi_{t_n}^{(n)}.$$

Then, for every $t \in \mathbb{R}^n$, the mapping ϕ_t is a diffeomorphism of M_c to itself. Furthermore, the diffeomorphism corresponding to t + s is a composition of diffeomorphisms corresponding to s and t:

$$\phi_{s+t} = \phi_s \circ \phi_t. \tag{30}$$

In other words, ϕ_t defines an action of the group \mathbb{R}^n on the manifold M_c . To prove (30) one uses that the flows of the vector fields X_{F_i} commute with each other:

$$\phi_{s+t} = \phi_{s_1+t_1}^{(1)} \circ \dots \circ \phi_{s_n+t_n}^{(n)} = (\phi_{s_1}^{(1)} \circ \phi_{t_1}^{(1)}) \circ \dots \circ (\phi_{s_n}^{(n)} \circ \phi_{t_n}^{(n)}) = (\phi_{s_1}^{(1)} \circ \dots \circ \phi_{s_n}^{(n)}) \circ (\phi_{t_1}^{(1)} \circ \dots \circ \phi_{t_n}^{(n)}) = \phi_s \circ \phi_t.$$

Further, observe that the action ϕ_t has the following properties:

1. It is transitive, i.e. for any $x_1, x_2 \in M_c$ there exists $t \in \mathbb{R}^n$ such that $\phi_t(x_1) = x_2$. In other words, the action ϕ_t has only one orbit. To prove that, it suffices to show that every orbit is open. Since M_c is connected and orbits are disjoint, it then follows that there can be only one orbit. To prove that every orbit is open, pick $x_0 \in M_c$. Its orbit is the set $\{\phi_t(x_0) \mid t \in \mathbb{R}^n\}$. This set is the image of the map $\mathbb{R}^n \to M_c$ given by $t \mapsto \phi_t(x_0)$. Let us show that this map is open. Moreover, it is a local diffeomorphism. To prove that, we compute the differential of the map:

$$\frac{\partial}{\partial t_i}\phi_t(x_0) = \frac{\partial}{\partial t_i}(\phi_{t_1}^{(1)} \circ \cdots \circ \phi_{t_n}^{(n)})(x_0).$$

Using that the flows of X_{F_i} 's commute, the latter expression can be rewritten as

$$\frac{\partial}{\partial t_i}\phi_{t_i}^{(i)}((\phi_{t_1}^{(1)}\circ\ldots\phi_{t_{i-1}}^{(i-1)}\circ\phi_{t_{i+1}}^{(i+1)}\circ\cdots\circ\phi_{t_n}^{(n)})(x_0)),$$

which, by definition of the flow, is equal to

$$X_{F_i}(\phi_{t_i}^{(i)}((\phi_{t_1}^{(1)} \circ \dots \phi_{t_{i-1}}^{(i-1)} \circ \phi_{t_{i+1}}^{(i+1)} \circ \dots \circ \phi_{t_n}^{(n)})(x_0))) = X_{F_i}(\phi_t(x_0))$$

So, the differential of the map $t \mapsto \phi_t(x_0)$ takes the vector $\partial/\partial t_i$ to the vector $X_{F_i}(\phi_t(x_0))$, and using that X_{F_i} 's are linearly independent at every point we conclude that the differential of our map takes a basis in the tangent space to a basis in the tangent space and hence is an isomorphism. But this exactly means that our map is a local diffeomorphism and hence an open map, as desired.

2. The stabilizer of every point $x_0 \in M_c$ is a discrete subgroup $\Gamma \subset \mathbb{R}^n$ (note that since the action is transitive and \mathbb{R}^n is an Abelian group, the stabilizers of all points in M_c are actually the same). Indeed, we already saw that the map $\mathbb{R}^n \to M_c$ given by $t \mapsto \phi_t(x_0)$ is a local diffeomorphism. In particular, it maps a neighborhood of $0 \in \mathbb{R}^n$ bijectively to a neighborhood of $x_0 \in M_c$, which means that $0 \in \mathbb{R}^n$ is an isolated point in the stabilizer Γ of x_0 . But any subgroup $\Gamma \subset \mathbb{R}^n$ such that 0 is an isolated point of Γ must be discrete. Indeed, if Γ is not discrete, then there exists a Cauchy sequence $v_k \in \Gamma$. But then $v_{k+1} - v_k$ is a sequence of elements of Γ which tends to 0, which contradicts 0 being an isolated point.

Since \mathbb{R}^n acts on M_c transitively with stabilizer Γ , we can conclude that M_c is diffeomorphic to the quotient \mathbb{R}^n / Γ . This can be thought of a as a geometric version of the orbit-stabilizer theorem. We will not discuss the precise conditions of this theorem here (in the next lecture we will construct a diffeomorphism from M_c to the torus explicitly, so we can in fact avoid using this general theorem). It now remains to describe the discrete subgroup Γ . To that end, we will use the following classification of discrete subgroups of \mathbb{R}^n :

Theorem 18.1. Any discrete subgroup $\Gamma \subset \mathbb{R}^n$ is a **lattice**, i.e. a set of integral linear combinations of $m \leq n$ linearly independent vectors $v_1, \ldots, v_m \in \mathbb{R}^n$.

Proof. We will prove this theorem for n = 1, while the general case can be found e.g. in [2, Section 49]. Let Γ be a discrete subgroup of \mathbb{R} . If $\Gamma = \{0\}$, then the statement of the theorem holds trivially. So, we can assume that Γ contains a non-zero element. Therefore, Γ also contains a positive element (because it is a subgroup and thus always contains -x along with x). Let v be the minimal positive element in Γ (such an element exists since Γ is discrete and thus any it subset, in particular the subset of positive elements, is closed). Then any other element $w \in \Gamma$ must be of the form mv, where $m \in \mathbb{Z}$ is an integer. Indeed, take any $w \in \Gamma$. Dividing it by v, we can represent it as w = mv + r, where $m \in \mathbb{Z}$, and r is a real number such that $0 \leq r < v$. Since Γ is a subgroup, we have $r = w - mv \in \Gamma$. But r is less than v and thus cannot be positive, so r = 0, which proves that every element of Γ is indeed an integer multiple of v. On other hand, since $v \in \Gamma$, all integer multiples of v are also in Γ , so Γ is a lattice spanned by v, as desired.

So, the stabilizer of our \mathbb{R}^n action on M_c is a lattice Γ spanned (over \mathbb{Z}) by some linearly independent vectors $v_1, \ldots, v_m \in \mathbb{R}^n$. Performing a change of basis, we can assume that v_i is the *i*'th basis vector. Thus, Γ coincides with the subgroup $\mathbb{Z}^m \subset \mathbb{R}^n$ which consists of points whose first *m* coordinates are integers. To compute the quotient $\mathbb{R}^n / \mathbb{Z}^m$, notice that the *i*'th copy of \mathbb{Z} in \mathbb{Z}^m acts only on the *i*'th copy of \mathbb{R} in \mathbb{R}^n , so

$$\mathbb{R}^n / \mathbb{Z}^m \simeq (\mathbb{R} / \mathbb{Z})^m \times \mathbb{R}^{n-m} \simeq T^m \times \mathbb{R}^{n-m}.$$

Therefore, M_c is diffeomorphic to a product of an *m*-dimensional torus and an (n-m)-dimensional vector space. But since M_c is compact, it follows that m = n (i.e. Γ is a full rank lattice), and $M \simeq T^n$, as desired.

Remark 18.2. If M_c is regular and compact but not connected, then the same argument shows that each connected component of M_c is still diffeomorphic to an *n*-dimensional torus. Furthermore, if M_c is not compact, then the above argument shows that every connected component of M_c is diffeomorphic to a product of the form $T^m \times \mathbb{R}^{n-m}$, provided that all the vector fields X_{F_i} are complete (i.e. their integral trajectories are defined for all times). This completeness assumption is very important: if it is not satisfied, then M_c may have quite intricate topology. For this reason, the completeness assumption is sometimes included in the definition of an integrable system. There are, however, natural examples when there is no completeness but one can still understand the topology of M_c by using algebraic (=complex geometric) approach to integrability instead of the symplectic one.

Lecture 19: Proof of the second part of the Arnold-Liouville theorem: dynamics

In the previous lecture we proved that a compact connected regular joint level set M_c of n functions F_1, \ldots, F_n in involution on a 2n-dimensional symplectic manifold M is diffeomorphic to a n-dimensional torus. In this section, we will prove that *each* of the vector fields X_{F_i} , in particular the vector field X_H , where $H = F_1$, is linear on that torus, in appropriate coordinates. We will also prove that the discrete part of the Arnold-Liouville theorem: if there is a symplectic map T preserving F_1, \ldots, F_n , then the restriction of this map to the torus M_c is a shift. To prove those statements, we make the identification between M_c and the torus T^n constructed in the previous lecture more explicit.

Recall that we have a map $\mathbb{R}^n \to M_c$ given by

$$(t_1,\ldots,t_n)\mapsto\phi_{t_1}^{(1)}\circ\cdots\circ\phi_{t_n}^{(n)}(x_0),$$

where $\phi_{t_i}^{(i)}$ is the time t_i shift along the trajectories of the vector field X_{F_i} , and x_0 is an arbitrary point in M_c . Denote this map by ϕ . Recall that the map ϕ is a surjective local diffeomorphism. Furthermore, $\phi(t) = \phi(t')$ if and only if $t - t' \in \Gamma$, where Γ is a full rank lattice in \mathbb{R}^n , i.e. the set of integer linear combinations of some n linear independent vectors. Consider the linear map $\psi \colon \mathbb{R}^n \to \mathbb{R}^n$ which takes the vectors $2\pi e_i$, where e_1, \ldots, e_n is the standard basis in \mathbb{R}^n , to the basis vectors of the lattice Γ . Then the composite map $\phi \circ \psi \colon \mathbb{R}^n \to M_c$ is a surjective local diffeomorphism. Furthermore, $\phi \circ \psi(t) = \phi \circ \psi(t')$ if and only if the coordinates of the vector t - t' are integer multiples of 2π . Therefore, the map $\phi \circ \psi$ descends to a global diffeomorphism $\mathbb{R}^n/2\pi\mathbb{Z}^n \to M_c$. So, we once again see that M_c is diffeomorphic to the n-dimensional torus $\mathbb{R}^n/2\pi\mathbb{Z}^n = T^n$. We will now show that vector fields X_{F_i} are constant in standard coordinates ϕ_1, \ldots, ϕ_n on T^n , coming from Cartesian coordinates in \mathbb{R}^n . To prove that, it suffices to show that the pull-back of X_{F_i} to \mathbb{R}^n by means of the map $\phi \circ \psi$ is a constant vector field. For the map ϕ , we already saw that in the previous lecture: X_{F_i} is the ϕ -pushforward of the coordinate vector field $\partial/\partial t_i$. But then it follows that the pull-back of X_{F_i} by $\phi \circ \psi$ is constant as well, because ψ is a linear map and hence takes constant vector fields to constant vector fields. This ends the proof of the dynamical part of the Arnold-Liouville theorem in the continuos case.

In the discrete case, we have all of the above, plus a symplectic map T which preserves F_1, \ldots, F_n and thus restricts to M_c . Furthermore, since T preserves both the symplectic structure and the functions F_1, \ldots, F_n , it also preserves the symplectic gradients $X_{F_i} = \omega^{-1} dF_i$ of those functions. So, we have a map of the torus M_c to itself which preserves n linearly independent *constant* vector fields X_{F_i} . We claim that any map with this property is a translation. Indeed, the vector fields X_{F_i} are of the form

$$X_{F_i} = \sum_j k_{ij} \frac{\partial}{\partial \phi_j},$$

where k_{ij} 's are constant. Since X_{F_i} 's are linearly independent, the matrix k_{ij} is invertible, so there exists another matrix l_{ij} such that

$$\frac{\partial}{\partial \phi_i} = \sum_j l_{ij} X_{F_j}.$$

In other words, X_{F_i} 's and $\partial/\partial \phi_i$'s are just two different bases in the *n*-dimensional vector space of constant vector fields on the torus. And since the map T preserves the basis of X_{F_i} 's, it also preserves $\partial/\partial \phi_i$'s (as well as any other constant vector field). But this means that T commutes with the flow of $\partial/\partial \phi_i$, which is the translation in the direction ϕ_i . So, T commutes with all translations of the torus. Therefore,

$$T(\phi_1,\ldots,\phi_n)=T(0,\ldots,0)+(\phi_1,\ldots,\phi_n),$$

which means that T is a translation by T(0, ..., 0). Thus, the dynamical part of the Arnold-Liouville theorem is proved in the discrete case as well.

We skip the proof of the last two parts of the Arnold-Liouville theorem (action-angle coordinates and explicit solutions). Those proofs can be found e.g. in [2, Section 49].

Lecture 20: Noether's principle, first examples of integrable systems

As we saw in the previous lectures, Hamiltonian systems with many conserved quantities exhibit nice dynamical behavior. Now, we want to understand where such conserved quantities may come from. The most natural from the physical perspective source of conserved quantities is symmetries. Noether's principle is mechanics says that any symmetry of a physical system gives rise to a conserved quantity. The goal of this lecture is to make this statement precise, as well as to discuss some examples.

There are two types of symmetries of physical systems: discrete and continuous. Discrete symmetries do not lead to conserved quantities, but they are useful for understanding continuous ones, so we will discuss the discrete case first.

Consider a natural mechanical system. Such a system is described by a configuration manifold M, along with a Riemannian metric on it (which is twice the kinetic energy), and a smooth function (the potential). A *discrete symmetry* of such a system is a diffeomorphism $\phi: M \to M$ which preserves the metric (i.e. ϕ is an *isometry*) and the potential. Having such a symmetry ϕ , we define its *cotangent lift* $\hat{\phi}$ by the pushforward

$$\phi(p) = \phi_* p$$

for any $p \in T^*M$. It is easy to see that $\hat{\phi}$ is a diffeomorphism of the cotangent bundle T^*M onto itself.

Exercise 20.1. For any diffeomorphism $\phi: M \to M$ its cotangent lift $\hat{\phi}: T^*M \to T^*M$ is a symplectomorphism, i.e. it preserves the canonical structure symplectic structure on the cotangent bundle.

Exercise 20.2. Consider a diffeomorphism $\mathbb{R} \to \mathbb{R}$ given by a function f(q) with $f'(q) \neq 0$. Describe its cotangent lift and explicitly show that it preserves the symplectic form on the cotangent bundle.

Exercise 20.3. For any symmetry $\phi: M \to M$ of a natural mechanical system on M its cotangent lift $\hat{\phi}: T^*M \to T^*M$ preserves the associated Hamiltonian function on T^*M .

Thus, any symmetry of the configuration space preserving a natural mechanical system gives rise to a transformation of the phase space preserving both the symplectic form and the Hamiltonian.

We now turn our attention to continuous symmetries. Such a symmetry is given by a vector field v on the configuration manifold M such that the Lie derivative of both the metric and the potential with respect to v vanishes. This is equivalent to requiring that the flow ϕ_t of v consists of discrete symmetries (see Proposition 13.8). Thus, for every t, the cotangent lift $\hat{\phi}_t$ is a symplectomorhism preserving the Hamiltonian.

Exercise 20.4. Show that for any $t_1, t_2 \in \mathbb{R}$ we have $\hat{\phi}_{t_1+t_2} = \hat{\phi}_{t_1} \circ \hat{\phi}_{t_2}$. Deduce that $\hat{\phi}_t$ is itself a flow of the vector field \hat{v} on T^*M given by

$$\hat{v} = \left. \frac{d}{dt} \right|_{t=0} \hat{\phi}_t.$$

The vector field \hat{v} is called the *cotangent lift*, or *cotangent prolongation* of v.

Exercise 20.5. Show that in coordinates the cotangent prolongation of a vector field

$$v = \sum_i v_i \frac{\partial}{\partial q_i}$$

is given by

$$\hat{v} = \sum_{i} v_i \frac{\partial}{\partial q_i} - \sum_{i,j} \frac{\partial v_j}{\partial q_i} p_j \frac{\partial}{\partial p_i}.$$
(31)

Since the flow $\hat{\phi}_t$ of the vector field \hat{v} consists of symplectic diffeomorphisms, the vector field \hat{v} is symplectic, i.e. one has $L_{\hat{v}}\omega = 0$ for the canonical symplectic structure ω on the cotangent bundle. Furthermore, the vector field \hat{v} is Hamiltonian. The corresponding Hamiltonian function is the vector field v interpreted as a function on T^*M . We denote this function by f_v :

$$f_v(p) = p(v) \quad \forall p \in T^*M.$$

In local coordinates, we have

$$f_v = \sum_i v_i p_i.$$

Since the canonical symplectic structure is standard in (p,q) coordinates on T^*M , we can compute the associated Hamiltonian vector field as

$$\begin{aligned} q'_i &= \quad \frac{\partial f_v}{\partial p_i} = v_i, \\ p'_i &= \quad - \frac{\partial f_v}{\partial q_i} = - \sum_{i,j} \frac{\partial v_j}{\partial q_i} p_j \frac{\partial}{\partial p_i} \end{aligned}$$

This agrees with formulas (31) for the cotangent prolongation. Thus, the cotangent prolongation \hat{v} of any vector field v is indeed Hamiltonian, with the Hamiltonian function given by v itself (more precisely, the Hamiltonian function is $f_v(p) = p(v)$, which is nothing else but v regarded as function on the cotangent bundle). Assume further that v is a symmetry of a natural mechanical system on M. Then the cotangent prolongation \hat{v} preserves the associated Hamiltonian function H, so

$$\{f_v, H\} = L_{X_{f_v}} H = L_{\hat{v}} H = 0$$

Thus, f_v is a first integral for the Hamiltonian dynamics on T^*M given by the Hamiltonian function H. This is the Noether's principle: every symmetry v gives rise to a conserved quantity, namely f_v .

The fact that $\{f_v, H\} = 0$ for a symmetry v can also se seen as follows. Given any contravariant tensor T on M, i.e. a multilinear form on every cotangent space, we associate to it a function f_T on T^*M given by

$$f_T(p) = T(p, \dots, p)$$

This agrees with our earlier definition of f_v for a vector field v. Furthermore, if T is a tensor of type (0,0), i.e. a function on M, then f_T is simply the lift of T to the cotangent bundle.

Proposition 20.6. For any contravariant tensor T and any vector field v one has

$$\{f_v, f_T\} = f_{L_vT}.$$

Proof sketch. This identity is equivalent to

$$L_{\hat{v}}f_T = f_{L_vT}.$$

The latter is straightforward to verify when T is either a function or a vector field, using e.g. the coordinate form (31) of the cotangent prolongation \hat{v} . For an arbitrary tensor T, the necessary identity can be proved by writing T as

$$T = \sum T_{i_1 \dots i_k} \frac{\partial}{\partial q_{i_1}} \otimes \dots \otimes \frac{\partial}{\partial q_{i_k}}$$

and applying the product rule for the Lie derivative.

The Hamiltonian H of a natural system can be written as

$$H = \frac{1}{2}f_{g^{-1}} + f_U$$

where g^{-1} is the inverse of the metric, and U is the potential. Therefore, for a symmetry v we have

$$\{f_v, H\} = \frac{1}{2}\{f_v, f_{g^{-1}}\} + \{f_v, f_U\} = \frac{1}{2}f_{L_vg^{-1}} + f_{L_vU}$$

But since v is a symmetry, we have $L_v g^{-1} = 0$ and $L_v U = 0$, so $\{f_v, H\} = 0$, as desired.

Another corollary of Proposition 20.6 is the following:

Corollary 20.7. If v and w are commuting symmetries of a natural mechanical system, then the corresponding conserved quantities Poisson commute.

Proof. We have

$$\{f_v, f_w\} = f_{L_v w} = f_{[v,w]} = 0.$$

As an example of application of Noether's principle, consider the motion of a particle in \mathbb{R}^2 in a central force field. The corresponding configuration manifold is \mathbb{R}^2 with coordinates x, y, the metric is the standard Euclidian metric, and the potential is a function $U = U(x^2 + y^2)$. This system has a symmetry given by rotations around the origin. The corresponding symmetry vector field is

$$v = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$$

The associated function on the cotangent bundle is

$$f_v = xp_y - yp_x$$

It is known as the *angular momentum*. Since the system is invariant under rotations, the angular momentum Poisson-commutes with the Hamiltonian

$$H = \frac{1}{2}(p_x^2 + p_y^2) + U(x^2 + y^2).$$

The identity $\{f_v, H\} = 0$ is also easy to verify explicitly. Furthermore, it is easy to see that the Hamiltonian H and the angular momentum f_v are functionally independent. This can be shown by an explicit computation or using the following argument: the projections of the trajectories for the Hamiltonian vector field $X_{f_v} = \hat{v}$ to the configuration space are circles centered at the origin. Indeed, by construction those projections coincide with trajectories of v. If H and f_v were functionally dependent, then the same would be true for trajectories of X_H . However, we know that a point in the plane can move in any direction prescribed by initial velocity and does not, in general, move in circles around the origin.

Thus, the motion of a particle in \mathbb{R}^2 in a central force field is an example of an integrable system. The level sets of the energy function H are compact provided that $U(z) \to \infty$ as $z \to \infty$. Thus, joint level sets of H and f_v are also compact, and one can apply the Arnold-Liouville theorem: generic joint level sets are two-dimensional tori with quasi-periodic dynamics. For a generic potential and generic torus, the dynamics is genuinely quasi-periodic: the trajectory fills the torus densely. The projection of such a torus to the configuration space is a two-dimensional domain, filled densely by the trajectory. Thus, a generic trajectory of a particle in \mathbb{R}^2 in a central force field densely fills a two-dimensional domain.

There is, however, an exceptional situation, corresponding to the potential $U = \frac{1}{2}(x^2 + y^2)$ (or, more generally, any constant multiple of $x^2 + y^2$). The corresponding force field F = -grad U is linear: F = -(x, y). Such force is known as *Hooke's force*. To physically realize such a system one attaches the particle to the origin by a spring.

In the Hooke's case, the Hamiltonian is of the form

$$H = \frac{1}{2}(p_x^2 + p_y^2 + x^2 + y^2).$$

It can be separated into two parts, each of which depends either only on x or only on y:

$$H = \frac{1}{2}(p_x^2 + x^2) + \frac{1}{2}(p_y^2 + y^2).$$

Because of that, we have that H Poisson-commutes with its x part

$$H_x = p_x^2 + x^2,$$

so H_x is an additional conserved quantity (same is true for $H_y = p_y^2 + y^2$ but H_y can be expressed in terms of H and H_x). Thus, we have a Hamiltonian system on a 4-dimensional symplectic manifold with 3 > 2 independent conserved quantities (independence can be checked explicitly). Such systems where the number of first integrals exceeds $\frac{1}{2} \dim M$ are known as *superintegrable*. Of course, the Poisson brackets of H, f_v, H_x cannot be all zero, because there is at most two independent functions on a 4-manifold with vanishing Poisson brackets. Explicitly, one has

$$\{f_v, H_x\} = \{xp_y - yp_x, p_x^2 + x^2\} = 2(p_xp_y + xy) \neq 0.$$

Remark 20.8. Note that the latter function is also a first integral, as Poisson brackets of first integrals. It is however dependent with the ones we already have:

$$(p_x p_y + xy)^2 = H_x H_y - f_v^2.$$

It is in fact impossible (for non-zero vector field) to have four independent integrals on a 4-manifold, because in that case generic trajectories would be points.

A physical manifestation of superintegrability on 4-manifolds is closed trajectories. Indeed, joint level sets of three independent integrals must be one-dimensional manifolds, so by compactness they are circles. Therefore, the trajectories of the system are contained in circles and hence must be closed. In particular, all generic trajectories of a particle in \mathbb{R}^2 attached to a spring are closed. One can further show that those trajectories are ellipses centered at the origin (this should not be confused with gravitational potential where bounded trajectories are ellipses which have the origin as one of the foci).

Lecture 21: Further examples of integrable systems

spherical pendulum, geodesics on the sphere and in \mathbb{R}^2 , geodesics on ellipsoid, Chasles' theorem...

Lecture 22: Introduction to Poisson manifolds

In this lecture we introduce the notion of a Poisson manifold. Recall that on every symplectic manifold M we have a Poisson bracket operation on functions which has the following properties:

- 1. It is skew-symmetric, i.e. $\{g, f\} = -\{f, g\}$.
- 2. It is bilinear over real numbers. For example, linearity in the first argument means that $\{af+bg,h\} = a\{f,h\} + b\{g,h\}$, where f,g,h are functions, and a,b are real numbers.
- 3. It satisfies the product rule in both arguments, e.g. $\{fg,h\} = f\{g,h\} + g\{f,h\}$.
- 4. It satisfies the Jacobi identity

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

A Poisson manifold is a manifold endowed with an operation on functions satisfying exactly the same properties. Such a manifold in general does not have to be symplectic. For example, any manifold (in particular, an odd-dimensional manifold) endowed with a zero bracket is a Poisson manifold. But not any manifold has a symplectic structure (for example, odd-dimensional manifolds do not admit such structures).

We will now discuss what part of symplectic formalism can be developed for Poisson manifolds.

Proposition 22.1. For any smooth function f on a Poisson manifold, there exists a unique vector field X_f such that $\{f, g\} = L_{X_f}g$ for any smooth function g.

Proof. It follows from the properties of the Poisson bracket that for fixed f the operation $\{f, *\}$ is a differentiation. But any differentiation of functions is a differentiation along a certain (uniquely determined) vector field. Therefore, there exists a unique vector field X_f such that $\{f, g\} = dg(X_f)$, as desired. \Box

As in the symplectic case, the vector field X_f is called the *Hamiltonian vector field* with the Hamiltonian function f.

Proposition 22.2. For any Poisson bracket, there exists a bivector field π (i.e. a skew-symmetric bilinear form on cotangent vectors) such that $\{f, g\} = \pi(df, dg)$.

Proof. From the definition of the vector field X_f and properties of the Poisson bracket it follows that $X_{fg} = fX_g + gX_f$. Therefore, for any covector $\xi \in T_x^*M$, we have $\xi(X_{fg}) = f(x)\xi(X_g) + g(x)\xi(X_f)$. So, the correspondence $f \mapsto \xi(X_f)$ is, for fixed ξ , a differentiation at the point x. We know that any such differentiation is determined by a tangent vector. Denote this tangent vector by v_{ξ} . Then we have $\xi(X_f) = df(v_{\xi})$. It is also clear from this formula that v_{ξ} linearly depends on the covector ξ , i.e. we have a linear operator $\tilde{\pi} \colon T_x^*M \to T_xM$ such that $v_{\xi} = \tilde{\pi}(\xi)$. Thus, $\xi(X_f) = \tilde{\pi}(\xi, df)$, where in the latter formula $\tilde{\pi}$ is interpreted as a bilinear function on covectors. Further, taking $\xi = dg$, we get that $\{f,g\} = dg(X_f) = \tilde{\pi}(dg, df)$. To complete the proof, it now remains to denote $\pi = -\tilde{\pi}$. Then $\{f,g\} = \pi(df, dg)$. Skew-symmetry of π follows from skew-symmetry of the Poisson bracket, while smooth dependence on the point x can be seen e.g. from the formula $\pi_{ij} = \{x_i, x_j\}$ for the components.

The bivector field π is called the *Poisson tensor* (or *Poisson bivector*). We just saw that any Poisson bracket defines a Poisson bivector. The components of that bivector in coordinates are given by the formula $\pi_{ij} = \{x_i, x_j\}$. Conversely, given a bivector π , we can construct a bracket using the formula $\{f, g\} = \pi(df, dg)$. This bracket is automatically bilinear, skew-symmetric, and satisfies the product rule in both arguments. However, the Jacobi identity imposes additional restrictions on the tensor π .

Exercise 22.3. A Poisson bracket is uniquely determined by brackets of coordinate functions. A bracket satisfies the Jacobi identity if and only if it is satisfied by coordinate functions.

Recall now that for symplectic manifolds the Poisson tensor π is the inverse of the symplectic structure. Conversely, we can reconstruct the symplectic structure as the inverse of the Poisson structure. Thus, a Poisson manifold is symplectic if and only if the Poisson tensor is invertible. The closedness of the so-defined symplectic form can be deduced from the Jacobi identity for the Poisson structure.

Exercise 22.4. Show that the formulas

$$\{x, y\} = z, \quad \{y, z\} = x, \quad \{z, x\} = y.$$

define a Poisson structure in \mathbb{R}^3 . Show that it does not correspond to any symplectic structure. Show that the function $x^2 + y^2 + z^2$ has a zero Poisson bracket with any other function. Such functions are known as *Casimir functions*. Note that there are no such (non-zero) functions in the symplectic case.

Exercise 22.5. Let V be a vector space, and let π be a skew-symmetric bilinear form on V^* . Show that the formula $\{f, g\} = \pi(df, dg)$ defines a Poisson bracket on V. Describe the corresponding Poisson tensor and Casimir functions.

Exercise 22.6. A vector space \mathfrak{g} is called a *Lie algebra* if it is endowed with a bilinear skew-symmetric operation [,] satisfying the Jacobi identity. Show that the formula $\{f, g\}(x) = x([df(x), dg(x)])$ defines a Poisson bracket on the dual space \mathfrak{g}^* of any Lie algebra. This bracket is known as the *Lie-Poisson bracket*. Show that the Poisson bracket from Exercise 22.4 is the Lie-Poisson bracket corresponding to the Lie algebra \mathbb{R}^3 with cross-product operation.

Lecture 23: Symplectic leaves

In this lecture we will show that any Poisson manifold can be represented as a disjoint union of symplectic manifolds, called *symplectic leaves*.

Recall that the Hamiltonian vector field corresponding to a smooth function f on a Poisson manifold M is defined as the unique vector field X_f satisfying the identity

$$\{f,g\} = L_{X_f}g$$

for any smooth function g on M. This can be rewritten as

$$\{f,g\} = dg(X_f). \tag{32}$$

At the same time, we have

$$\{f, g\} = \pi(df, dg) = dg(\pi(df)), \tag{33}$$

where in the latter formula we interpret the Poisson tensor π as a map from the cotangent space to the tangent space. Comparing formulas (32) and (33), we see that the Hamiltonian vector field X_f is explicitly given by

$$X_f = \pi(df). \tag{34}$$

Recall that we have a similar definition of the Hamiltonian vector field X_f in the symplectic case: $X_f = \omega^{-1}(df)$. In the Poisson case, the Poisson tensor π plays the role of ω^{-1} .

Exercise 23.1. For the Poisson bracket from Example 22.4, compute the Hamiltonian vector field with Hamiltonian f = x.

From (34) it follows that for any function f the Hamiltonian vector field X_f belongs to the subspace Im π at every point $x \in M$: $X_f(x) \in \text{Im } \pi(x)$. In the symplectic case, this result is vacuous, because the Poisson tensor is invertible and hence surjective. However, if the Poisson tensor π is degenerate, then the result that we get is non-trivial: all Hamiltonian vector fields belong to a proper subspace of the tangent space at every point.

Assume now for simplicity that the Poisson tensor π has constant rank k, i.e. the rank of $\pi(x)$ is k for all $x \in M$. Then $\operatorname{Im} \pi(x)$ is a k-dimensional subspace of the tangent space $T_x M$. A choice of a k-dimensional subspace at every tangent space of a manifold is called a k-dimensional distribution (one also needs to assume that the dependence of the subspace on the point is smooth). We are already familiar with 1-dimensional distributions: those correspond to a choice of a direction in every tangent space, i.e. 1-dimensional distributions are the same as direction fields. As follows from the existence and uniqueness theorem for ODEs, for any direction field on a manifold M and any point $x \in M$ there exists an integral curve of the direction field through x, i.e. an unparametrized curve whose tangent line at every point coincides with the 1-dimensional subspace given by the direction field. Similarly, for k-dimensional distributions we would like to construct *integral surfaces*, i.e. surfaces whose tangent plane at every point coincides with the plane determined by the distribution. A distribution is called *integrable* if it admits an integral surface through every point. An integrable k-dimensional distribution gives rise to a k-dimensional foliation, i.e. a partition of the ambient manifold into a disjoint union of k-dimensional surfaces. Conversely, any foliation gives an integrable distribution: one just takes all tangent spaces of surfaces forming the foliation.

It turns out that for $k \ge 2$ a k-dimensional distribution does not have to be integrable. To see this, observe that if v, w are vector fields belonging to an integrable distribution, then their commutator [v, w] should also belong to the distribution. Indeed, since v, w belong to the distribution, they are tangent to surfaces forming the corresponding foliation, and hence the same is true for their commutator.

Exercise 23.2. Prove that there exists no surface in \mathbb{R}^3 which is orthogonal to the vector field v = (z, x, 0) at every point. Hint: find two vector fields orthogonal to v whose commutator is not orthogonal to v.

A distribution is called *involutive* if the commutator of any two vector fields belonging to the distribution also belongs to the distribution. We see that an integrable distribution must be involutive. Conversely, **any involutive distribution is integrable**: this is Frobenius's theorem. There is also a version of this result for *singular distributions*, i.e. distributions whose dimension vary with the point. This is usually referred to as the Stefan-Sussmann theorem, see e.g. [10].

Exercise 23.3. Explain why there may exist at most one foliation tangent to a given distribution.

A curious manifestation of the Frobenius theorem is our ability to navigate a motor vehicle with just two controls (steering and driving) in the four-dimensional configuration space (the four parameters are x and y coordinates of the car, its orientation, and orientation of the front wheels). This is possible due to the non-involutivity of the distribution spanned by the control vector fields, see [6, Section 2.6].

Exercise 23.4. Let v_1, \ldots, v_k be vector fields belonging to a certain distribution L. Furthermore, assume that foe very point x the vectors $v_1(x), \ldots, v_k(x)$ span L(x). Finally, assume that $[v_i, v_j] \in L$ for every i, j. Prove that the distribution L is involutive.

Thanks to this result, it is sufficient to verify the involutivity condition only for basis vector fields of the distribution. In our case, as such basis vector fields one can take Hamiltonan vector fields, as shown b the following exercise:

Exercise 23.5. Assume that v is a vector field belonging to the distribution Im π , where π is a Poisson tensor. Then v can be locally expressed as a linear combination of Hamiltonian vector fields (with non-constant coefficients).

Thus, to verify involutivity of the distribution $\text{Im }\pi$ it suffices to check that the commutator of Hamiltonian vector fields belongs to $\text{Im }\pi$. But this follows from the formula

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

which is established in the same way as in the symplectic case. So, by the Frobenius theorem the distribution $\text{Im }\pi$ is integrable. In the case when the rank of the Poisson tensor is non-constant, we can similarly apply the Stefan-Sussmann theorem (for that one needs to verify certain additional technical conditions), or, alternatively, deduce integrability from Weinstein's *splitting theorem* [21]. All in all, we get that every Poisson manifold is foliated by submanifolds tangent to the distribution $\text{Im }\pi$. These submanifolds are called *symplectic leaves*.

Proposition 23.6. Every symplectic leaf has a canonical structure of a symplectic manifold. In particular, every symplectic leaf (and thus the image of the Poisson tensor at every point) is even-dimensional.

Proof. Let L be a symplectic leaf. We will first show that L inherits a structure of a Poisson manifold. Take two arbitrary smooth functions f and g on L and extend them to smooth functions \tilde{f} and \tilde{g} defined in some open (in the ambient manifold M) neighborhood of L. Define the bracket of f and g by

$$\{f,g\} = \{f,\tilde{g}\}.$$

We claim that this bracket does not depend on the choice of extensions. Indeed, suppose we change one of the extensions, say, \tilde{f} to \hat{f} . Then

$$\{\tilde{f} - \hat{f}, \tilde{g}\} = \pi(d(\tilde{f} - \hat{f}), d\tilde{g}).$$

But the function $\tilde{f} - \hat{f}$ vanishes on L, so its differential at any point $x \in L$ vanishes on the tangent space $T_x L = \operatorname{Im} \pi(x)$. But since π is skew-symmetric, the annihilator of its image is its kernel. So, $d(\tilde{f} - \hat{f}) \in \operatorname{Ker} \pi$, and $\{\tilde{f} - \hat{f}, \tilde{g}\} = 0$, which proves that the bracket $\{f, g\}$ is well-defined. Furthermore, it satisfies all properties of a Poisson structure simply because it is defined using a Poisson structure on the ambient manifold. So, it remains to show that the Poisson structure on L is non-degenerate and hence gives rise to a symplectic structure. To that end, observe that

$$\{f,g\} = \{\tilde{f},\tilde{g}\} = d\tilde{g}(X_{\tilde{f}}).$$

But $X_{\tilde{f}}$ belongs to the tangent space of L, so $d\tilde{g}(X_{\tilde{f}}) = dg(X_{\tilde{f}})$, and the Hamiltonian vector field X_{f} (relative to the Poisson structure on L) coincides with the Hamiltonian vector field $X_{\tilde{f}}$ (relative to the Poisson structure on the ambient manifold). Therefore,

$$X_f = \pi d\tilde{f}$$

which means that the Poisson tensor π on M and the Poisson tensor π_L on L are related by

$$\pi = \pi_L \circ r$$

where $r: T_x^*M \to T_x^*L$ is the natural restriction map. So, since π is surjective (as a mapping on $T_xL = \text{Im }\pi(x)$), π_L is surjective as well, and hence an isomorphism, as desired.

Exercise 23.7. A mapping $\phi: M \to N$ between Poisson manifolds is called a *Poisson map* if $\{\phi^*f, \phi^*g\}_M = \phi^*\{f, g\}_N$ for any smooth functions f, g on N (here $\{,\}_M, \{,\}_N$ are Poisson brackets on M and N respectively). A submanifold $M \subset N$ of a Poisson manifold N is called a *Poisson submanifold* if the inclusion map $i: M \to N$ is Poisson. Prove that any symplectic leaf of a Poisson manifold is a Poisson submanifold. Prove that a submanifold $M \subset N$ is Poisson if and only if its tangent space contains the image of the Poisson tensor (equivalently, if any Hamiltonian vector field on N is tangent to M). Describe Poisson submanifolds of a symplectic manifold is called symplectic if the restriction of the symplectic structure to that submanifold is again symplectic).

Lecture 24: Symplectic leaves and Casimir functions

Recall that a function f on a Poisson manifold M is called a *Casimir function* if $\{f, g\} = 0$ for any other smooth function g on M.

Proposition 24.1. A function is a Casimir if and only if it is constant on every symplectic leaf.

Proof. For any symplectic leaf L, we have

$$\{f,g\}|_L = \{f|_L, g|_L\}_L,$$

where $\{,\}_L$ is the Poisson bracket on L induced by the bracket on the ambient manifold M. So, $\{f,g\} = 0$ if and only if

$$\left\{ \left. f \right|_L, \left. g \right|_L \right\}_L = 0$$

for any symplectic leaf L. Furthermore, since g is arbitrary, $g|_L$ is also an arbitrary function on L. So, f is a Casimir if and only if its restriction to every symplectic leaf is a Casimir. But symplectic leaves are symplectic and hence have no Casimirs except for constant functions. The result follows.

Here is yet another characterization of Casimir functions:

Proposition 24.2. A smooth function f is a Casimir if and only if at every point we have $\pi df = 0$, where π is the Poisson tensor (interpreted as a map from the cotangent space to the tangent space).

Proof. We have $\{f, g\} = \pi(df, dg)$, so f is a Casimir if and only if $\pi(df, dg) = 0$ for any smooth function g. But any covector is a differential of a suitable smooth function, so f is a Casimir if and only if $\pi(df, \xi) = 0$ for an arbitrary covector ξ , which is equivalent to saying that $\pi df = 0$.

Example 24.3. Let us find Casimir functions of the bracket from Example 22.4. The corresponding Poisson tensor reads

$$\pi = \left(\begin{array}{ccc} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{array} \right).$$

The kernel of π at generic points is spanned by the covector xdx + ydy + zdz. Thus f is a Casimir if and only if

$$df = \lambda(xdx + ydy + zdz),$$

where λ is a scalar function. In particular, $f = x^2 + y^2 + z^2$ is a Casimir, corresponding to $\lambda = 2$. We can now use this information to describe the symplectic leaves. The level sets of the function $f = x^2 + y^2 + z^2$ are concentric spheres $x^2 + y^2 + z^2 = a$, and also the origin, which is a single-point level set. Since Casimir functions are constant on symplectic leaves, symplectic leaves must be subsets of level sets of Casimir functions. In particular, each of the spheres $x^2 + y^2 + z^2 = a$, a > 0 is a disjoint union of symplectic leaves. But since the rank of the Poisson tensor is equal to two away from the origin, it follows that symplectic leaves contained in the sphere $x^2 + y^2 + z^2 = a$ are two-dimensional and thus coincide with the sphere itself (a connected manifold cannot be represented as a disjoint union of several manifolds of the same dimension). So, each of the spheres $x^2 + y^2 + z^2 = a$, a > 0 is a two-dimensional symplectic leaf. The union of such spheres is the whole space minus the origin, so the origin must be a zero-dimensional leaf. This agrees with the fact that the rank of the Poisson tensor at the origin is zero.

From this description of symplectic leaves it also follows that any other Casimir is a function of $f = x^2 + y^2 + z^2$. Indeed, any function constant on spheres $x^2 + y^2 + z^2 = a$ must be a smooth function of $x^2 + y^2 + z^2$.

Exercise 24.4. Show that the symplectic form on the symplectic leaves $x^2 + y^2 + z^2 = a$ is proportional to the standard area form on the sphere.

Example 24.5. Consider the following modification of the bracket from Example 22.4:

$$\{x, y\} = -z, \quad \{y, z\} = x, \quad \{z, x\} = y.$$
(35)

Similarly to Example 24.3, we find that the function $f = x^2 + y^2 - z^2$ is a Casimir. Its level sets $x^2 + y^2 - z^2 = a$ are

- One-sheeted hyperboloids for a > 0.
- Two-sheeted hyperboloids for a < 0.
- A cone for a = 0.

Using the same argument as in Example 24.3, we conclude that one-sheeted hyperboloids, as well as each of the two sheets of two-sheeted hyperboloids are symplectic leaves. As for the cone, it cannot be a single leaf in particular because it is not a manifold. To represent it as a union of leaves, observe that the Poisson bracket vanishes at the origin, so the origin is a zero-dimensional leaf. Furthermore, removing the origin from the cone, we are left with a two-dimensional manifold, so the same argument as in Example 24.3 shows that each of the open halves of the cone, i.e. each of the sets $x^2 + y^2 - z^2 = 0, z > 0$ and $x^2 + y^2 - z^2 = 0, z < 0$ is a symplectic leaf. So, all in all, the cone is a union of three symplectic leaves.

As in Example 24.3, it is not hard to show that any other Casimir of the bracket (35) is a function of $x^2 + y^2 - z^2$, at least locally. So, in both of the above examples there is essentially only one Casimir, in the sense that any other Casimir is functionally dependent with the one that we found. More generally, we have the following result:

Proposition 24.6. Assume that dim Ker $\pi = k$ almost everywhere on M (such a number k exists in all reasonable example of Poisson brackets; in the above two examples, we have k = 1). Then π has at most k functionally independent Casimirs.

Proof. Let f_1, \ldots, f_l be functionally independent Casimirs. Then their differentials df_1, \ldots, df_l are linearly independent almost everywhere in M. In particular, there is a point $x \in M$ such that dim Ker $\pi(x) = k$, and $df_1(x), \ldots, df_l(x)$ are linearly independent. So, since $df_i(x) \in \text{Ker } \pi(x)$, it follows that $l \leq k$, as desired.

It turns out that locally this bound is exact: if dim Ker $\pi = k$, then one can always find k functionally independent local Casimirs:

Proposition 24.7. Assume that dim Ker $\pi = k$ in the neighborhood of a given point $x \in M$. Then there exist k functionally independent local Casimirs near x.

Proof sketch. Since dim Ker $\pi = k$ near x, all symplectic leaves close to x are submanifolds of the same codimension k. Changing coordinates, we can assume that these submanifolds are planes, all parallel to each other. Moreover, we can in fact arrange that these submanifolds are given by the equations $x_i = c_i$, where $i = 1, \ldots, k$, and c_1, \ldots, c_k are constants (whose values depend on the choice of the leaf). Then x_1, \ldots, x_k are the sought local Casimirs.

Globally, this result does not need to hold, as shown by the following example:

Exercise 24.8. Consider the Poisson bracket in \mathbb{R}^3 given by $\{x, y\} = 0, \{z, x\} = x, \{z, y\} = y$. Show that the corresponding Poisson tensor has a one-dimensional kernel at generic points, but there are no globally defined Casimirs.

Lecture 25: Lie groups and Lie algebras

In this lecture we give a brief introduction to the theory of Lie groups and algebras. Details can be found in any textbook on Lie theory, see e.g. [7, Chapter 1].

Recall that a group G is called a *Lie group* if G is smooth manifold, and the group operations in G, i.e. multiplication and inversion, are smooth maps.

Example 25.1. The group GL_n of invertible $n \times n$ matrices over real numbers is a Lie group of dimension n^2 . Indeed, GL_n is a smooth manifold of dimension n^2 because it is an open subset in the n^2 -dimensional vector space of all $n \times n$ real matrices. Furthermore, the group operations are smooth because their expression in terms of matrix entries is given by smooth functions, namely polynomial (quadratic) functions for multiplication and rational functions for inversion.

Example 25.2. The group O_n of $n \times n$ orthogonal matrices is a Lie group of dimension $\frac{1}{2}n(n-1)$. To prove this, it suffices to show that it is a submanifold of GL_n of the indicated dimension (then the group operations are automatically smooth as restrictions of smooth group operations in GL_n). To prove the latter, consider the map Φ from all $n \times n$ matrices to symmetric $n \times n$ matrices given by $X \mapsto XX^t$. Then O_n can be described as the level set $\Phi^{-1}(\mathrm{Id})$. To prove the smoothness of that level set it suffices to show that Φ is a submersion at points $X \in \Phi^{-1}(\mathrm{Id})$, i.e. that the differential of Φ is surjective. The differential of Φ at X is given by

$$d_X \Phi(Y) = Y X^t + X Y^t.$$

It is a linear map from all $n \times n$ matrices to symmetric $n \times n$ matrices. We need to show that this map is surjective, i.e. that any symmetric matrix A can be written as $YX^t + XY^t$, where X is a fixed orthogonal matrix, and Y is an arbitrary matrix. Such a representation of A can be found by solving the equation

$$YX^t = \frac{1}{2}A.$$

Taking Y solving this equation we get that

$$YX^{t} + XY^{t} = YX^{t} + (YX^{t})^{t} = \frac{1}{2}A + \frac{1}{2}A^{t} = A,$$

as desired.

Example 25.3. The group SO_n of $n \times n$ orthogonal matrices with determinant 1 is also a Lie group of dimension $\frac{1}{2}n(n-1)$. Indeed, any orthogonal matrix has determinant ± 1 , so SO_n can be described as the subset of O_n which consists of orthogonal matrices with positive determinant. That is an open subset and hence a submanifold of O_n whose dimension is the same as the dimension of O_n .

Exercise 25.4. Show that O_n consists of two connected components, one of which is SO_n . Hint: use the canonical form theorem for orthogonal matrices.

For a Lie group G, we will denote by \mathfrak{g} it tangent space at the identity element $id \in G$.

Example 25.5. The space \mathfrak{gl}_n is the tangent space at the identity of GL_n . Since GL_n is an open subset in the vector space of all $n \times n$ matrices, \mathfrak{gl}_n is just the space of all $n \times n$ matrices.

Example 25.6. Since SO_n is an open subset of O_n , they have the same tangent spaces at the identity: $\mathfrak{so}_n = \mathfrak{o}_n$ (the notation \mathfrak{o}_n is not very common). Furthermore, since O_n can be described as the level set of the map $\Phi: X \mapsto XX^t$ (see Example 25.2), its tangent space \mathfrak{so}_n at the identity coincides with the kernel of the differential

$$d_{\mathrm{Id}}\Phi(Y) = Y + Y^t$$

Thus, \mathfrak{so}_n is the space of skew-symmetric matrices.

Exercise 25.7. Since $T_{\text{Id}}SO_n = \mathfrak{so}_n$, it follows that for any skew-symmetric matrix $X \in \mathfrak{so}_n$ there is a smooth curve $Y(t) \in SO_n$ of orthogonal matrices such that Y(0) = Id and Y'(0) = X. Give an explicit construction of the curve Y(t) for given X. Hint: one approach is to once again use the canonical form of X; alternatively, one can define Y(t) by the matrix exponential

$$Y(t) = \exp(tX) = \sum_{k=0}^{\infty} \frac{1}{k!} (tX)^k;$$

what needs to be shown is that the exponential of a skew-symmetric matrix is orthogonal.

We now show that the tangent space \mathfrak{g} at the identity of any Lie group G has a natural structure of a Lie algebra, i.e. it equipped with a binary operation [,] which is bilinear, skew-symmetric, and satisfies the Jacobi identity. So, to every group one can associate a Lie algebra. Conversely, one can show that every finite-dimensional Lie algebra corresponds to a Lie group (Lie's third theorem). Moreover, such a group G is unique if we require that G is connected and simply connected.

To construct the Lie algebra \mathfrak{g} of a group G, for a fixed $g \in G$ consider the mapping $\Phi_g: G \to G$ given by $\Phi_g(h) = ghg^{-1}$. This is a group action since $\Phi_{g_1} \circ \Phi_{g_2} = \Phi_{g_1g_2}$. Observe also that for any $g \in G$ we have $\Phi_g(\mathrm{id}) = \mathrm{id}$. Therefore, the differential $d_{\mathrm{id}}\Phi_g$ is a mapping of the tangent space \mathfrak{g} at the identity to itself. Denote $d_{\mathrm{id}}\Phi_g$ by Ad_g . It is also a group action, i.e. $\mathrm{Ad}_{g_1g_2} = \mathrm{Ad}_{g_1} \circ \mathrm{Ad}_{g_2}$. This action is known as the *adjoint action*, or *adjoint representation*. It is indeed a group representation, because Ad_g is a linear operator for any g.

Proposition 25.8. For a matrix Lie group G, i.e. for GL_n or any its Lie subgroup such as SO_n , we have $Ad_A(B) = ABA^{-1}$.

Proof. By definition of the differential, we have

$$\operatorname{Ad}_{A}(B) = \left. \frac{d}{dt} \right|_{t=0} \left(A\hat{B}(t)A^{-1} \right),$$

where $\hat{B}(t)$ is a curve in G such that $\hat{B}(0) = \text{Id}$ and $\hat{B}'(0) = B$. But

$$\frac{d}{dt}\Big|_{t=0} \left(A\hat{B}(t)A^{-1}\right) = A\left(\left.\frac{d}{dt}\right|_{t=0}\hat{B}(t)\right)A^{-1} = ABA^{-1}. \quad \Box$$

Example 25.9. The adjoint representation of SO_n on \mathfrak{so}_n is given by the conjugation action of orthogonal matrices on skew-symmetric matrices. This action is well-defined because for orthogonal A and skew-symmetric B we have

$$(ABA^{-1})^t = (ABA^t)^t = AB^tA^t = -ABA^{-1}.$$

The adjoint representation $g \mapsto \operatorname{Ad}_g$ can be viewed as a homomorphism $G \to GL(\mathfrak{g})$, where $GL(\mathfrak{g})$ stands for invertible linear transformations of \mathfrak{g} . Taking the differential of this homomorphism at the identity, we get a linear map between tangent spaces to G and $GL(\mathfrak{g})$ at the identity, i.e. a map $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$. The image of an arbitrary element $\xi \in \mathfrak{g}$ under this map is denoted by $\mathfrak{ad}_{\mathcal{E}}$.

Proposition 25.10. For a matrix Lie group we have $\mathfrak{ad}_A(B) = [A, B] = AB - BA$.

Proof. By definition, we have

$$\mathfrak{ad}_A(B) = \left. \frac{d}{dt} \right|_{t=0} \operatorname{Ad}_{\hat{A}(t)} B,$$

where $\hat{A}(t)$ is a curve in the group such that $\hat{A}(0) = \text{Id}$, and $\hat{A}'(0) = A$. Therefore,

$$\mathfrak{ad}_A(B) = \left. \frac{d}{dt} \right|_{t=0} \left(\hat{A}(t) B \hat{A}(t)^{-1} \right) = \left(\left. \frac{d}{dt} \right|_{t=0} \hat{A}(t) \right) B - B \left(\left. \frac{d}{dt} \right|_{t=0} \hat{A}(t) \right) = AB - BA,$$

where we used that

$$\frac{d}{dt}\hat{A}(t)^{-1} = -\hat{A}(t)^{-1} \left(\frac{d}{dt}\hat{A}(t)\right)\hat{A}(t)^{-1}$$

for any smooth curve $\hat{A}(t)$ in GL_n .

For an arbitrary Lie group G the formula $\mathfrak{ad}_{\xi}\eta = [\xi, \eta]$ is taken as the definition of the bracket $[\xi, \eta]$:

$$[\xi,\eta] = \mathfrak{ad}_{\xi}\eta \quad \forall \, \xi,\eta \in \mathfrak{g}.$$

One can show that this bracket is bilinear, skew-symmetric, and satisfies the Jacobi identity, so it turns the tangent space \mathfrak{g} of the Lie group G at the identity into a Lie algebra. The mapping $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ given by $\xi \mapsto \mathfrak{ao}_{\xi}$ is called the *adjoint representation of the Lie algebra* \mathfrak{g} .

Exercise 25.11. Using the Jacobi identity, show that $\xi \mapsto \mathfrak{ad}_{\xi}$ is indeed a Lie algebra representation in the sense that $\mathfrak{ad}_{[\xi_1,\xi_2]} = [\mathfrak{ad}_{\xi_1},\mathfrak{ad}_{\xi_2}]$, where the latter bracket is the usual commutator of operators. In other words, $\xi \mapsto \mathfrak{ad}_{\xi}$ is a homomorphism of Lie algebras $\mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$.

Exercise 25.12. Show that \mathfrak{gl}_n endowed with the matrix commutator is indeed a Lie algebra (one needs to verify the Jacobi identity).

Exercise 25.13. For a Lie group G let $\lambda_g: G \to G$ be the *left translation by* g, i.e. the mapping given by $\lambda_g(h) = gh$. A tensor field T on G is called *left-invariant* if $(\lambda_g)_*T = T$ for any $g \in G$ (observe that λ_g is a diffeomorphism so the push-forward $(\lambda_g)_*T$ is well-defined). Prove that for any $\xi \in \mathfrak{g}$ there exists a unique left-invariant vector field v_{ξ} on G such that $v_{\xi}(\mathrm{id}) = \xi$.

Exercise 25.14. Prove that the commutator of left-invariant vector fields is left-invariant. A more difficult part: prove that the mapping $\xi \mapsto v_{\xi}$ is a Lie algebra isomorphism between \mathfrak{g} and left-invariant vector fields on G. Thus, one can define the bracket on \mathfrak{g} via the Lie bracket of the corresponding left-invariant vector fields. Hint: let g(t) be the integral curve of v_{ξ} with $g(0) = \mathrm{id}$. Then, by left-invariance, hg(t) is also an integral curve for any $h \in G$. Thus, the flow of v_{ξ} is given by $h \mapsto hg(t)$, i.e. the flow of \mathfrak{a} left-invariant vector field is given by right translations. Denoting the right translation by g as ρ_g , we have that the time t shift along the trajectories of v_{ξ} is given by $\rho_{g(t)}$. This allows us to compute the bracket $[v_{\xi}, v_{\eta}]$ as the t-derivative of $(\rho_{g(t)})^* v_{\eta}$ at t = 0. Further, observe, that right translations commute with left translations, so the pull-back of a left-invariant vector field v_{η} by a right translation is again a left-invariant vector field. Express this left-invariant vector field in the form v_* and use this expression to compute the Lie bracket $[v_{\xi}, v_{\eta}]$.

Exercise 25.15. Show that the space SL_n of real $n \times n$ matrices with determinant 1 is a Lie group. Show that its Lie algebra is the space \mathfrak{sl}_n of matrices with trace 0.

Exercise 25.16. Show that invertible upper-triangular matrices form a Lie group. Describe its Lie algebra.

Lecture 26: Lie-Poisson brackets and their symplectic leaves

Let \mathfrak{g} be a finite-dimensional Lie algebra. Then, on the dual space \mathfrak{g}^* of \mathfrak{g} one has the following Poisson bracket, called the *Lie-Poisson bracket*:

$${f,g}(x) = x([df(x), dg(x)]).$$

Here f and g are arbitrary smooth functions on \mathfrak{g}^* , and $x \in \mathfrak{g}^*$. Note that the differentials df(x), dg(x) belong to the cotangent space $T_x^*\mathfrak{g}^*$, but since \mathfrak{g}^* is a vector space, that cotangent space can be canonically identified with $(\mathfrak{g}^*)^* = \mathfrak{g}$. Therefore, the Lie bracket [df(x), dg(x)] is well-defined. It is an element of the Lie algebra \mathfrak{g} (depending on x), so x([df(x), dg(x)]) is a well-defined function of x.

Note also that as coordinate functions on \mathfrak{g}^* one can take linear functions. Furthermore, linear functions on \mathfrak{g}^* can be identified with elements of \mathfrak{g} . For two such functions $\xi, \eta \in \mathfrak{g}$, we have $d\xi(x) = \xi$ and $d\eta(x) = \eta$, so their Lie-Poisson bracket is given by

$$\{\xi,\eta\}(x) = x([\xi,\eta]),$$

which means that the bracket of two linear functions ξ, η is again a linear function, corresponding to the Lie algebra element $[\xi, \eta]$. In other words, for linear functions we have

$$\{\xi,\eta\} = [\xi,\eta]. \tag{36}$$

Since a Poisson structure is uniquely determined by pairwise brackets of coordinate function, formula (36) can be taken as an alternative definition of the Lie-Poisson bracket. In other words, the Poisson-Lie bracket is the unique Poisson bracket that coincides with the Lie bracket on linear function. This definition also implies the Jacobi identity for the Lie-Poisson bracket. Indeed, by (36) the Jacobi identity is satisfied for coordinate functions, and hence for all functions (see Exercise 22.3).

Exercise 26.1. Let V be a finite-dimensional vector space endowed with a Poisson structure such that the bracket of linear functions is a linear function (such Poisson structures are known as linear). Show that V^* has a canonical Lie algebra structure. Thus, there is a one-to-one correspondence between finite-dimensional Lie algebras and finite-dimensional vector spaces with linear Poisson structures.

Exercise 26.2. Show that the bracket from Example 22.4 is the Lie-Poisson bracket on the dual of the algebra \mathfrak{so}_3 of 3×3 skew-symmetric matrices.

Exercise 26.3. Show that the bracket (35) is the Lie-Poisson bracket on the dual of the algebra \mathfrak{sl}_2 of 2×2 traceless matrices.

We will now describe symplectic leaves of Lie-Poisson brackets. To that end, we will need the notion of the coadjoint representation for a Lie group and Lie algebra. Recall that the adjoint representation $g \mapsto \operatorname{Ad}_g$ of a Lie group G is a representation of G on its Lie algebra \mathfrak{g} , i.e. Ad_g is an (invertible) operator $\mathfrak{g} \to \mathfrak{g}$ for every $g \in G$. The coadjoint representation $g \mapsto \operatorname{Ad}_g^*$ is defined as the dual representation on \mathfrak{g}^* . Namely, the operator Ad_g^* is, by definition, the dual of the operator $\operatorname{Ad}_{g^{-1}} = (\operatorname{Ad}_g)^{-1}$. In other words, for every $x \in \mathfrak{g}^*, \xi \in \mathfrak{g}$ we have

$$(\mathrm{Ad}_a^* x)(\xi) = x(\mathrm{Ad}_a^{-1}\xi).$$

This is a group representation. Indeed, denote by A^{\dagger} the dual of A. Then

$$\mathrm{Ad}_{g_1g_2}^* = ((\mathrm{Ad}_{g_1g_2})^{-1})^{\dagger} = ((\mathrm{Ad}_{g_1}\mathrm{Ad}_{g_2})^{-1})^{\dagger} = ((\mathrm{Ad}_{g_1})^{-1})^{\dagger} ((\mathrm{Ad}_{g_2})^{-1})^{\dagger} = \mathrm{Ad}_{g_1}^*\mathrm{Ad}_{g_2}^*.$$

Example 26.4. For the Lie group GL_n , we can identify its Lie algebra \mathfrak{gl}_n and its dual \mathfrak{gl}_n^* by means of the non-degenerate symmetric bilinear form

$$\operatorname{Tr} XY = \sum_{i,j=1}^{n} X_{ij} Y_{ji}.$$

Thus, elements of \mathfrak{gl}_n^* can be also represented by matrices. In this model, the coadjoint operator Ad_Z^* is defined as the adjoint (with respect to the Tr XY inner product) of Ad_Z^{-1} . But Ad_Z is an orthogonal operator for every $Z \in GL_n$, because

$$\operatorname{Tr} \operatorname{Ad}_Z X \operatorname{Ad}_Z Y = \operatorname{Tr} Z X Z^{-1} Z Y Z^{-1} = \operatorname{Tr} Z X Y Z^{-1} = \operatorname{Tr} X Y.$$

So, the adjoint of Ad_Z^{-1} coincides with Ad_Z , i.e. $\operatorname{Ad}_Z^* = \operatorname{Ad}_Z$. In other words, the coadjoint representation of GL_n coincides with the adjoint representation, provided that we identify \mathfrak{gl}_n and its dual \mathfrak{gl}_n^* by means of the form $\operatorname{Tr} XY$. More generally, coadjoint and adjoint representations of a Lie group G coincide if its Lie algebra \mathfrak{g} admits an *invariant inner product*, i.e. a non-degenerate (but not necessarily positive-definite) symmetric bilinear form such that

$$\langle \mathrm{Ad}_g \xi, \mathrm{Ad}_g \eta \rangle = \langle \xi, \eta \rangle \quad \forall g \in G, \quad \xi, \eta \in \mathfrak{g}.$$

Besides \mathfrak{gl}_n , another example of a Lie algebra admitting an invariant inner product is \mathfrak{so}_n . The inner product on \mathfrak{so}_n is obtained by restricting the Tr XY product from \mathfrak{gl}_n . The so-obtained inner product is still non-degenerate. Moreover, it is negative definite, since for a skew-symmetric matrix X we have

$$\operatorname{Tr} X^{2} = \sum_{i,j=1}^{n} X_{ij} X_{ji} = -\sum_{i,j=1}^{n} X_{ij}^{2},$$

which is negative as long as $X \neq 0$. Thus, the coadjoint representation for \mathfrak{so}_n also coincides with the adjoint representation. The algebra \mathfrak{so}_n is an example of a simple Lie algebra (i.e. a Lie algebra with no non-trivial ideals). All simple Lie algebras admit an invariant inner product (given by the so called *Killing form*), and their coadjoint and adjoint representations coincide. More generally, this is true for all *reductive* Lie algebras, such as \mathfrak{gl}_n . There are also examples of non-reductive Lie algebras with this property, see e.g. [3].

Exercise 26.5. Another example of a simple Lie algebra is the algebra \mathfrak{sl}_n of traceless matrices. Show that as an invariant inner product on \mathfrak{sl}_n one can again take the form form $\operatorname{Tr} XY$ (a non-trivial part is to explain why this form is non-degenerate). Hint: what is the orthogonal complement of the identity matrix in \mathfrak{gl}_n with respect to the $\operatorname{Tr} XY$ form?

Exercise 26.6. Show that for the group of invertible 2×2 upper-triangular matrices the adjoint and coadjoint representations are different.

The coadjoint representation $\xi \mapsto \mathfrak{ad}_{\xi}^*$ of a Lie algebra is defined by the corresponding group representation in the same way as in the adjoint case (more generally, any Lie group representation defines a representation of the corresponding Lie algebra). Namely the mapping $\xi \mapsto \mathfrak{ad}_{\xi}^*$ from \mathfrak{g} to $\mathfrak{gl}(\mathfrak{g}^*)$ is the differential of the mapping $g \mapsto \operatorname{Ad}_q$ from G to $GL_n(\mathfrak{g}^*)$.

Proposition 26.7. The operator $\mathfrak{ad}^{*}\xi$ is dual to $-\mathfrak{ad}_{\xi} = \mathfrak{ad}_{-\xi}$.

Proof. Let g(t) be a curve in the corresponding group G such that g(0) = id and $g'(0) = \xi$. Then, by definition of the coadjoint representation, we have

$$(\mathrm{Ad}_{g(t)}^* x)(\eta) = x(\mathrm{Ad}_{g(t)}^{-1} \eta).$$

for every $x \in \mathfrak{g}^*$, $\eta \in \mathfrak{g}$. Differentiating this with respect to t at t = 0 and using that

$$\mathfrak{ad}_{\xi} = \left. \frac{d}{dt} \right|_{t=0} \mathrm{Ad}_{g(t)}, \quad \mathfrak{ad}_{\xi}^* = \left. \frac{d}{dt} \right|_{t=0} \mathrm{Ad}_{g(t)}^*$$

we get

$$(\mathfrak{ad}_{\xi}^*x)(\eta) = -x(\mathfrak{ad}_{\xi}\eta)$$

as desired.

We can now rewrite the Lie-Poisson bracket in terms of the coadjoint representation:

$$\{f,g\}(x) = x([df(x), dg(x)]) = x(\mathfrak{ad}_{df(x)}dg(x)) = -(\mathfrak{ad}_{df(x)}^*x)(dg(x)).$$

The latter expression for the Poisson bracket means that the corresponding Poisson tensor at a point $x \in \mathfrak{g}^*$, understood as an operator from the cotangent space to tangent space, i.e. as an operator $\mathfrak{g} \to \mathfrak{g}^*$, is given by

$$\xi \mapsto -\mathfrak{a}\mathfrak{d}_{\xi}^* x. \tag{37}$$

In particular, the image of the Poisson tensor at x consists of all possible images of x under the coadjoint action of \mathfrak{g} .

Corollary 26.8. Symplectic leaves of the Lie-Poisson bracket on the dual of a Lie algebra \mathfrak{g} are orbits of the coadjoint representation of the corresponding Lie groups G.

Remark 26.9. If the group G is disconnected, then its coadjoint orbits may be disconnected as well. In this case, symplectic leaves of the Lie-Poisson bracket coincide with connected components of coadjoint orbits.

Proof of Corollary 26.8. It suffices to show that the tangent space to such an orbit at every point coincides with the image of the Poisson tensor. In other words, the tangent space to the orbit of coadjoint representation of the group coincides with the orbit of the coadjoint representation of the corresponding algebra. This can be easily proved using the following well-known fact: for any smooth Lie group action, the mapping $g \mapsto g \circ x$ from the group G to the orbit of x is a submersion. Applying this for the coadjoint action, we see that the tangent space to the orbit of x consists of vectors of the form

$$\left. \frac{d}{dt} \right|_{t=0} \left(\operatorname{Ad}_g^*(t) x \right),$$

where g(t) are all possible smooth curves in the group such that g(0) = id. But this can be rewritten as $\mathfrak{ad}^*_{a'(0)}x$, hence the result.

Corollary 26.10. Coadjoint orbits of any Lie group G carry a canonical symplectic structure.

This structure is known as the Kirillov-Kostant-Souriau symplectic form.

Example 26.11. Let \mathfrak{g} be a Lie algebra which admits an invariant inner product. Then coadjoint and adjoint orbits coincide, so symplectic leaves of the corresponding Lie-Poisson bracket can be identified with adjoint orbits. In particular, for matrix Lie algebras, such as \mathfrak{gl}_n , \mathfrak{sl}_n , and \mathfrak{so}_n , the adjoint action is given by $B \mapsto ABA^{-1}$, so symplectic leaves coincide with conjugacy classes of matrices. For example, two matrices $B, C \in \mathfrak{gl}_n$ belong to the same leaf if and only if there exists $A \in GL_n$ such that $C = ABA^{-1}$.

Remark 26.12. The conjugacy class of a given matrix is, generally speaking, disconnected. So, a more precise characterization of symplectic leaves for matrix Lie algebras is that they are connected components of conjugacy classes. This issue does not arise for the groups SL_n and SO_n , because those groups are connected.

Lecture 27: Lie-Poisson brackets and Casimirs

Recall that a function f on a Poisson manifold is called a Casimir function if $\{f, g\} = 0$ for any other smooth function g. Equivalently, Casimirs can be characterized as functions constant on every symplectic leaf. For Lie-Poisson brackets, symplectic leaves coincide with orbits of the coadjoint action of the connected Lie group G on the dual \mathfrak{g}^* of the Lie algebra \mathfrak{g} of G. Therefore, Casimirs are functions constant on coadjoint orbits. In other words, a smooth function f on \mathfrak{g}^* is a Casimir if an only if it is invariant under the coadjoint action, i.e.

$$(\mathrm{Ad}_a^*)^* f = f \quad \forall \, g \in G,$$

or, which is the same,

$$f(\operatorname{Ad}_q^* x) = f(x) \quad \forall g \in G, \quad x \in \mathfrak{g}^*.$$

Example 27.1. For matrix Lie groups (such as GL_n , SL_n , and SO_n) coadjoint action coincides with the adjoint one, i.e. the conjugation action. Thus, for such groups Casimirs of the Lie-Poisson bracket are precisely conjugation invariant functions (more precisely, in the GL_n case one needs to consider functions that are invariant under conjugation by matrices with positive determinant; this is because GL_n is disconnected, so in that case symplectic leaves are coadjoint orbits for the connected component of the identity, which is precisely the subgroup of matrices with positive determinant). In the GL_n case as conjugation invariant functions one can take coefficients $f_i(A)$ of the characteristic polynomial

$$\det(A - \lambda \mathrm{Id}) = \sum_{i=0}^{n} f_i(A)\lambda^i.$$

This gives n conjugation-invariant polynomial functions f_1, \ldots, f_{n-1} (we exclude f_n because it is constant, namely $f_n = (-1)^n$). For example,

$$f_0(A) = \det A, \quad f_{n-1}(A) = (-1)^{n-1} \operatorname{Tr} A.$$

One can show that any other conjugation-invariant polynomial can be expressed as a polynomial of f_0, \ldots, f_{n-1} . In more algebraic terms, f_1, \ldots, f_{n-1} generate the associative algebra of conjugationinvariant functions (moreover, it can be shown that they generate this algebra freely, i.e. there are no polynomial relations between f_1, \ldots, f_{n-1}). As another set of generators, one can take functions $h_i(A) = \operatorname{Tr} A^i$, where $i = 1, \ldots, n$. Functions h_1, \ldots, h_n can be expressed in terms of f_0, \ldots, f_{n-1} , and vice versa. The corresponding formulas are known as *Newton's identities*. For example, for 2×2 matrices we have $f_0 = \det A$, $f_1 = -\operatorname{Tr} A$, so

$$h_1 = \operatorname{Tr} A = -f_1, \quad h_2 = \operatorname{Tr} A^2 = f_1^2 - 2f_0.$$

The latter formula follows from the following expressions in terms of eigenvalues λ_1, λ_2 :

$$f_1 = -(\lambda_1 + \lambda_2), \quad f_0 = \lambda_1 \lambda_2, \quad h_2 = \lambda_1^2 + \lambda_2^2$$

Thus, as Casimirs of the Lie-Poisson bracket on \mathfrak{gl}_n^* one can take either coefficients of the characteristic polynomial, or traces of powers (up to $\operatorname{Tr} A^n$). Furthermore, the same functions can be taken as Casimirs for other matrix Lie algebras, although some of them become trivial. In particular, in the SL_n case we have $\operatorname{Tr} A = 0$, so as Casimirs it suffices to take $\operatorname{Tr} A^2, \ldots, \operatorname{Tr} A^n$, or, alternatively, all coefficients of the characteristic polynomial up to the coefficient of λ^{n-2} . In the SO_n case, even more Casimirs become trivial. Indeed, if $A \in \mathfrak{so}_n$ is skew-symmetric, then so is any odd power of A, therefore $\operatorname{Tr} A = \operatorname{Tr} A^3 = \cdots = 0$, and as Casimirs we can take traces of even powers of A.

Exercise 27.2. Describe Casimirs of the Lie-Poisson bracket on \mathfrak{so}_n^* in terms of coefficients of the characteristic polynomial. Which of those coefficients vanish identically?

Remark 27.3 (On Lax form of Hamiltonian equations). To say that traces of powers are Casimirs in the dual \mathfrak{g}^* of a martix Lie algebra \mathfrak{g} is the same as to say that those functions are first integrals for *any* Hamiltonian evolution in \mathfrak{g}^* . The latter can be also seen from the *Lax form* of Hamiltonian equations. From the form (37) of the Lie-Poisson tensor on \mathfrak{g}^* we get that the Hamiltonian equation with Hamiltonian function H reads

$$\dot{x} = -\mathfrak{a}\mathfrak{d}^*_{dH(x)}x.$$

If \mathfrak{g} admits an invariant inner product (in particular if \mathfrak{g} is a matrix Lie algebra such as \mathfrak{gl}_n , \mathfrak{sl}_n , or \mathfrak{so}_n), then identifying \mathfrak{g}^* with \mathfrak{g} by means of that inner product we identify \mathfrak{ad}^* with \mathfrak{ad} . Therefore in that case the above Hamiltonian equation rewrites as

$$\dot{x} = -\mathfrak{a}\mathfrak{d}_{dH(x)}x = -[dH(x), x] = [x, dH(x)].$$
(38)

In the case of a matrix () Lie algebra the latter bracket is the usual matrix commutator, so this Hamiltonian equation is an example of a *Lax equation*. One says that a matrix (or, more generally, an operator in an infinite-dimensional vector space) L depending on time t satisfies a Lax equation if

$$\frac{dL}{dt} = [L, A]$$

for some other (generally speaking, t-dependent) matrix (operator) A. The matrices L and A entering this equation are known as a *Lax pair*. An equation that can be written in a Lax form is said to admit a *Lax representation*. An immediate corollary of the Lax form is preservation of traces of powers:

Exercise 27.4. Show that if L satisfies a Lax equation, then L^k also satisfies a Lax equation, with the same A. Conclude that the traces of powers of L do not depend on time.

We thus once again see that Hamiltonian equations on the dual of matrix Lie algebras preserve traces of powers, which means that those traces are Casimirs.

Observe also that preservation of traces of powers of L in Lax evolution is equivalent to preservation of the spectrum (eigenvalues) of L. Thus, Lax equations describe *isospectral deformations*. In fact, even a stronger statement is true: Lax evolution preserves Jordan normal form. In other words, the matrix L(t) is conjugate to L(0) for any t:

Exercise 27.5. A matrix L(t) satisfies a Lax equation if and only if there exists another t-dependent matrix U(t) such that $L(t) = U(t)^{-1}L(0)U(t)$. In other words, a Lax equation describes an evolution of a matrix in its conjugacy class (we already know this for Hamiltonian equations on duals of matrix Lie algebras, because conjugacy classes are precisely symplectic leafs, which are preserved by Hamiltonian vector fields).

Example 27.6. Let us discuss in more detail the Lie-Poisson bracket on \mathfrak{so}_3^* , the corresponding Casimirs, and symplectic leaves. We actually already discussed that bracket in Example 22.4, but now we will do the same from the Lie theory perspective. Recall that \mathfrak{so}_3^* can be identified with \mathfrak{so}_3 by means of the form Tr XY and thus consists of skew-symmetric matrices

$$A = \begin{pmatrix} 0 & x & y \\ -x & 0 & z \\ -y & -z & 0 \end{pmatrix}.$$

The only "non-trivial" Casimir in this case is $\operatorname{Tr} A^2 = -2(x^2 + y^2 + z^2)$ (cf. Example 22.4). Its level sets are spheres centered at the origin, as well as the origin itself. From this it can be deduced, as in Example 22.4, that symplectic leaves in this example coincide with level sets of the Casimir. This

can also be seen directly, arguing as follows. Symplectic leaves in $\mathfrak{so}_3^* = \mathfrak{so}_3$ coincide with orbits of the conjugation action of SO_3 on \mathfrak{so}_3 . Every skew-symmetric operator on \mathbb{R}^3 can be brought, by an orthogonal transformation, to the form

$$A = \begin{pmatrix} 0 & \omega & 0 \\ -\omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$
 (39)

Moreover, such a normal form is unique, up to replacing ω with $-\omega$, because the eigenvalues of the latter matrix are $\pm \omega \sqrt{-1}$, so matrices with different $|\omega|$ are not conjugate to each other. Thus, conjugacy classes of 3×3 skew-symmetric matrices are characterized by $|\omega| = \sqrt{-\frac{1}{2} \operatorname{Tr} A^2}$.

Exercise 27.7. In this argument we actually used that every skew-symmetric operator in \mathbb{R}^3 can be brought, by a *special* orthogonal transformation (i.e. a transformation in SO_n), to the form (39). Explain why in this case one does not need to distinguish between orthogonal and special orthogonal transformations.

Example 27.8. We will now discuss in detail the case of \mathfrak{sl}_2^* (cf. Example 24.5). The space $\mathfrak{sl}_2^* = \mathfrak{sl}_2$ consists of matrices of the form

$$A = \left(\begin{array}{cc} a & b \\ c & -a \end{array}\right).$$

As a Casimir function we can take det $A = -a^2 - bc$. By performing change of variables a = x, b = y + z, c = y - z, we can rewrite this function in the form det $A = -x^2 - y^2 + z^2$. Therefore, the level set det A = c is a two-sheeted hyperboloid for c > 0, cone for c = 0, and one-sheeted hyperboloid for c > 0. From this one obtains the following description of symplectic leaves (cf. Example 24.5): each one-sheeted hyperboloid is a leaf, each sheet of each two-sheeted hyperboloid is a leaf, while the cone consists of three leaves: the origin, and two connected components of the complement to the origin. At the same time, we know that symplectic leaves in \mathfrak{sl}_2^* are given by Jordan normal forms of traceless 2×2 matrices. To describe such forms, observe that from $\mathrm{Tr} = 0$ condition it follows that the eigenvalues form a pair $\pm \lambda$. Therefore, the following cases are possible:

• Two real non-zero eigenvalues $\pm \lambda$. The corresponding normal form is

$$\left(\begin{array}{cc}\lambda & 0\\ 0 & -\lambda\end{array}\right).$$

In this case, the determinant is $-\lambda^2 < 0$, so the corresponding symplectic leaf is a one-sheeted hyperboloid.

• Two imaginary non-zero eigenvalues $\pm \omega \sqrt{-1}$. The corresponding normal form is

$$\left(\begin{array}{cc} \omega\sqrt{-1} & 0\\ 0 & -\omega\sqrt{-1} \end{array}\right),$$
$$\left(\begin{array}{cc} 0 & \omega\\ -\omega & 0 \end{array}\right).$$

or

In this case, we have det = $\omega^2 > 0$, which means that the level set of the Casimir consists of two symplectic leaves (two sheets of a two-sheeted hyperboloid). The point is that even though all 2×2 matrices with eigenvalues $\pm \omega \sqrt{-1}$ are conjugate, they may not be conjugate in SL_2 (see Exercise 27.9 below).

• The matrix is non-zero, but both eigenvalues are zero. This corresponds to matrices with Jordan normal form

$$\left(\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right).$$

The corresponding leaf is one of the two open halves of the complement of the origin in the cone. There are again two leafs because it may be not possible to bring to a Jordan normal form using a matrix from SL_2 .

Exercise 27.9. Show that the matrices

$$\left(\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right), \quad \left(\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right)$$

have the same eigenvalues and Jordan normal forms but are not conjugate in $SL_2(\mathbb{R})$. Are they conjugate in $SL_2(\mathbb{C})$?

Lecture 28: Lie-Poisson brackets from Poisson reduction

In this lecture we will give a construction of Lie-Poisson structures starting from the canonical Poisson structure on the cotangent bundle. This explains the relevance of Lie-Poisson structures in physics. The construction is based on the notion of *Poisson reduction*, which we will now discuss.

Assume that M is a manifold, and G is a group acting on M. This means that for every $g \in G$ we have a mapping $\rho_g \colon M \to M$, and that $\rho_{gh} = \rho_g \circ \rho_h$. We will say that the *G*-action is good (this is not a standard term) if

- 1. The quotient space M/G is a smooth manifold.
- 2. The natural projection mapping $pr: M \to M/G$ is smooth. As a corollary, any smooth function f on M/G lifts to a smooth G-invariant function $f \circ pr$ on M.
- 3. Conversely, any smooth G-invariant function on M descends to a smooth function on M/G.

Remark 28.1. There is no reason why these assumptions should be satisfied for general group actions on manifolds. They are, however, satisfies for certain classes of actions. For example, by the quotient manifold theorem [15] the above assumptions hold if G is a Lie group acting on M smoothly (i.e. the action mapping $(g, x) \mapsto \rho_g(x)$ is smooth), freely, and properly (i.e. the mapping $G \times M \to M \times M$ given by $(g, x) \mapsto (\rho_g(x), x)$ is proper). In particular, our assumptions are satisfied for free actions of compact (and, in particular, finite) Lie groups.

Example 28.2. Consider an \mathbb{R} action on \mathbb{R}^2 given by $(x, y) \mapsto (x + t, y)$. The orbits of this action are horizontal lines, so the quotient $\mathbb{R}^2 / \mathbb{R}$ can be identified with \mathbb{R} . The projection mapping $pr \colon \mathbb{R}^2 \to \mathbb{R}^2 / \mathbb{R}$ is the map $(x, y) \mapsto y$. It is easy to see that all the above assumptions are satisifed: the quotient is smooth, and the projection mapping is smooth as well. Finally, an \mathbb{R} -invariant smooth function on \mathbb{R}^2 must be of the form f(y), so it descends to a smooth function on the quotient. Thus, this action is a good action.

Example 28.3. Consider an action of S^1 on \mathbb{R}^2 by rotations about the origin. For this action, the orbits are circles centered at the origin, plus the origin itself. The quotient is homeomorphic to a closed ray, which is not a manifold. Thus, this action is not good. It does, however, become good if remove the origin from \mathbb{R}^2 (as follows e.g. from the quotient manifold theorem, see Remark 28.1). This situation is quite typical: there are many actions which are not good but become good upon removal of a certain zero measure singular set from the manifold M.

For good actions, there is natural isomorphism between the space $C^{\infty}(M/G)$ of smooth functions on the quotient and the space $C^{\infty}(M)^G$ of smooth *G*-invariant functions on *M*. Explicitly, the mapping $C^{\infty}(M/G) \to C^{\infty}(M)^G$ is given by the pull-back pr^* .

Perhaps the easiest way to show that a given action is good is to construct a *section*. We say that a smooth submanifold $S \subset M$ is a section of a *G*-action if it intersects each *G*-orbit at exactly one point. Furthermore, we will say that *S* is a *good section* (again, non-standard terminology) if the mapping $M \to S$ which takes *x* to the intersection point of its *G*-orbit with *S* is smooth.

Example 28.4. Show that if a G-action of M admits a good section S, then M/G is homeomorphic to S and therefore carries a smooth manifold structure. Hence show that an action admitting a good section is good.

Exercise 28.5. Show that x = 0 is a good section of the action from Exercise 28.2. Show that $y = x^3$ is also a section, but not good.

If an action admits a good section S, then we can identify the quotient M/G with S. We therefore have an isomorphism between the spaces $C^{\infty}(M)^G$ and $C^{\infty}(S)$. Explicitly, this isomorphism is given by restriction.

We will now assume that M is a Poisson manifold, and that the *G*-action on M is Poisson, which means that $\rho_g: M \to M$ is a Poisson diffeomorphism for every g. In other words,

$$\rho_a^* \{ f_1, f_2 \} = \{ \rho_a^* f_1, \rho_a^* f_2 \} \quad \forall g \in G, \quad f_1, f_2 \in C^\infty(M) \}$$

Exercise 28.6. Show that a diffeomorphism $\rho: M \to M$ is Poisson if and only if $\rho_*\pi = \pi$, where π is the Poisson tensor.

Proposition 28.7. Assume that we are given a Poisson action of G on M. Then the Poisson bracket of G-invariant functions is a G-invariant function.

Proof. Let $f_1, f_2 \in C^{\infty}(M)^G$ be G-invariant functions. Then, since the G-action is Poisson, we have

$$\rho_q^* \{ f_1, f_2 \} = \{ \rho_q^* f_1, \rho_q^* f_2 \}.$$

Further, by G-invariance, we get

$$\{\rho_g^* f_1, \rho_g^* f_2\} = \{f_1, f_2\},\$$

 \mathbf{SO}

$$\rho_g^* \{ f_1, f_2 \} = \{ f_1, f_2 \},\$$

which exactly means that $\{f_1, f_2\}$ is a *G*-invariant function.

Thus, for Poisson G-actions, the space $C^{\infty}(M)^G$ of G-invariant functions is closed under the Poisson bracket. For good actions, that space can be canonically identified with the space of $C^{\infty}(M/G)$ of smooth functions on the quotient. Thus, the algebra of smooth functions on the quotient by a good Poisson action carries a Poisson structure. In other words, if we have a good Poisson action of a group G on a Poisson manifold M, then M/G also carries a structure of a Poisson manifold. This procedure allowing one to construct a Poisson structure on the quotient is known as *Poisson reduction*, and the Poisson structure on M/G is called the *reduced Poisson structure*. Reduction of Poisson manifolds was introduced in [13] as a generalization of an earlier notion of symplectic reduction [12].

Our main example here is left action of a Lie group G on T^*G . The group G acts on itself via left translations $\lambda_g(h) = gh$. The cotangent lift $\hat{\lambda}_g$ of that action preserves the canonical symplectic structure on T^*G (see Exercise 20.1), and hence the Poisson structure. So, the left action of G on T^*G is Poisson.

Exercise 28.8. Show that $\mathfrak{g}^* = T^*_{id}G$ is a good section of the left *G*-action on T^*G . Therefore, that action is a good action.

Thus, the quotient T^*G/G has a Poisson manifold structure. Furthermore, since \mathfrak{g}^* is a good section, we can identify T^*G/G with \mathfrak{g}^* . To compute the reduced Poisson structure on $T^*G/G = \mathfrak{g}^*$, we need to extend two given functions on \mathfrak{g}^* to *G*-invariant functions on T^*G , compute the bracket of obtained functions, and then again restrict to \mathfrak{g}^* . Since a Poisson structure is determined by brackets of coordinate functions, it suffices to do this for linear functions on \mathfrak{g}^* .

Exercise 28.9. Let $\xi \in \mathfrak{g}$ be a linear function on \mathfrak{g}^* . Then its extension to a left-invariant function on T^*G is given by $f_{v_{\xi}}$, where v_{ξ} is a left-invariant vector field on G with $v_{\xi}(\mathrm{id}) = \xi$ (see Exercise 25.14), and $f_{v_{\xi}}$ is the corresponding function on T^*G (see Section 20).

Exercise 28.10. Using the above exercise, Exercise 25.14, and Proposition 20.6 show that the reduced Poisson structure on $T^*G/G = \mathfrak{g}^*$ coincides with the Lie-Poisson structure.

Thus, the Lie-Poisson structure on \mathfrak{g}^* arises as a reduction of the canonical symplectic structure on T^*G by the left *G*-action. One can also show that the right action gives rise to the same Poisson structure, but with an opposite sign.

Lecture 29: Reduction of Hamiltonian dynamics and rigid body motion

In the previous lecture we considered the situation when a group G acts in a Poisson way on a Poisson manifold M. We saw that if the action is what we called *good*, then the quotient M/G carries a canonical Poisson structure, called the *reduced Poisson structure*. We will now show that this reduction procedure can also be applied to Hamiltonian dynamics, provided that the Hamiltonian function is G-invariant.

Assume first that we have an arbitrary G-invariant vector field v on M. Then, assuming that M/G is smooth, and so is the projection $pr: M \to M/G$, we can push the vector field v forward to M/G. This pushforward is defined by the formula

$$(pr_*v)(y) = d(pr)(v(x)),$$

where $y \in M/G$ is arbitrary and $x \in M$ is such that pr(x) = y. Note that the expression d(pr)(v(x))does not depend on the choice of $x \in pr^{-1}(y)$, because if x, x' both belong to $pr^{-1}(y)$, then they belong to the same *G*-orbit, i.e. there is $g \in G$ such that $\rho_q(x) = x'$. But then by *G*-invariance of v we have,

$$d\rho_q(v(x)) = v(x'),$$

 \mathbf{SO}

$$d(pr)(v(x')) = d(pr \circ \rho_g)(v(x)) = d(pr)(v(x))$$

where we used that $pr \circ \rho_g = pr$. Thus, a *G*-invariant vector field on *M* descends to a vector field pr_*v on the quotient M/G.

Exercise 29.1. Show that if the *G*-action on *M* is good, and *v* is a *G*-invariant vector field, then the push-forward pr_*v is smooth. Hint: show that for an arbitrary smooth function *f* on M/G we have

$$(L_{pr_*v}f)(y) = (L_v(pr^*f))(x)$$

for any $x \in M$ and y = pr(x). Show that the function $L_v(pr^*f)$ is G-invariant. Deduce that $L_{pr*v}f$ is a smooth function.

Now assume that $H \in C^{\infty}(M)^G$ is a smooth *G*-invariant function on *M*. Then, if the action of *G* on *M* is Poisson, the Hamiltonian vector field X_H is also *G*-invariant. So, as explained above, we have a well-defined push-forward pr_*X_H on M/G.

Proposition 29.2. The push-forward vector field pr_*X_H is Hamiltonian with respect to the reduced Poisson structure on M/G. The corresponding Hamiltonian functions is the push-forward of the Hamiltonian pr_*H :

$$pr_*X_H = X_{pr_*H}.$$

Proof. Let $f \in C^{\infty}(M/G)$ be arbitrary. Then

$$L_{pr_*X_H}f = df(pr_*X_H).$$

By definition of pull-back as the dual of the push-forward, this can be further rewritten as

$$(pr^*df)(X_H) = d(pr^*f)(X_H) = L_{X_H}pr^*f = \{H, pr^*f\}_M.$$

The latter is the bracket of two G-invariant functions on M, so be definition of the reduced Poisson structure on G/M that expression is equal to $\{pr_*H, f\}_{M/G}$. Thus, we have

$$L_{pr_*X_H}f = \{pr_*H, f\}_{M/G},$$

which means that the vector field pr_*X_H is Hamiltonian with respect to the reduced Poisson structure, with Hamiltonian function given by pr_*H .

Thus, Hamiltonian dynamics invariant under a Poisson action descends to Hamiltonian dynamics on the quotient by the action. As an example, consider the motion of a rigid body fixed at one of its points moving under the action of no external forces (to emphasize the absence of external forces, one usually calls this a *free* rigid body). To describe a position of a body, consider an orthonormal frame which is *frozen* in the body, i.e. rotating together with the body. It is then clear that the position (orientation) of the body determines the frozen frame and vice versa. Furthermore, the frozen frame is determined by an operator taking the vectors of a fixed orthonormal frame in the space (a *space frame*) to the vectors of the frozen frame (*body frame*). Assuming that the space frame consists of the standard basis vectors in \mathbb{R}^3 , the matrix of such an operator has the vectors of the frozen frame as its columns. Since the frozen frame is orthonormal, that matrix A is orthogonal. Furthermore, if the frozen frame is chosen positively oriented, then one has det A = 1, and thus $A \in SO_3$. Therefore, one can identify a position of a rigid body with a fixed point with a matrix $A \in SO_3$. In other words, SO_3 is the configuration space of such a rigid body. The dynamics of the body is therefore governed by a Hamiltonian system on T^*SO_3 , with the Hamiltonian given by the inverse of the kinetic energy (there are no external forces and hence no potential energy).

Let us now show that the kinetic energy of the rigid body is given by a *left-invariant metric* on SO_3 , i.e. a metric that is invariant under the left action of SO_3 on itself. The origin of that invariance is what physicists call "isotropy of space", which means that Newtonian law of motions are the same in every space direction (provided there is no gravity field which breaks the isotropy). Explicitly, the kinetic energy is given by the integral over the whole body of the quantity $\frac{1}{2}\langle \dot{R}, \dot{R} \rangle$, where R is the radius-vector (in the space frame) of the given point in the body. Assuming that both the body and the space frame are centered at the fixed point, we have R = Ar, where r is the radius-vector of the same point in the body frame. Since the body frame is frozen in the body, we have r = const: the radius-vector of any point in the body has fixed coordinates in the body frame. Therefore, we have

$$\langle \dot{R}, \dot{R} \rangle = \langle \dot{A}r, \dot{A}r \rangle$$
But the latter expression is invariant under transformations of the form $A \mapsto BA$, where $B \in SO_3$ is timeindependent. Therefore, the kinetic energy is indeed invariant under the left action of G on itself, while the associated Hamiltonian is invariant under the left action of G on T^*G . Thus, the rigid body dynamics descends to a Hamiltonian system on $T^*SO_3/SO_3 = \mathfrak{so}_3^*$. The corresponding Hamiltonian function can be obtained by restricting the inverse of the kinetic energy from T^*SO_3 to \mathfrak{so}_3^* and is, therefore, a quadratic form. Thus, we obtain the following:

Corollary 29.3. The dynamics of a rigid body fixed at one of its points moving under the action of no external forces is described by a Hamiltonian system. The corresponding Poisson structure is the Lie-Poisson structure on \mathfrak{so}_3^* , while the Hamiltonian function is quadratic.

What is the physical meaning of this reduced system? Taking the quotient by SO_3 means identifying configurations (or, more precisely, phase space points) related to each other by rotation. We thus completely disregard the position of the body and from now on only measure its velocity. This is somewhat easier to understand at the level of the tangent bundle. The dynamics of the body can be thought of as a flow on TSO_3 . The corresponding vector field is invariant under the left SO_3 -action and thus descends to the quotient TSO_3/SO_3 . The latter can be identified with the Lie algebra \mathfrak{so}_3 , in the same way as we identified T^*SO_3/SO_3 with \mathfrak{so}_3^* . The projection $TSO_3 \to T^*SO_3/SO_3 = \mathfrak{so}_3$ maps the tangent vector $\dot{A} \in T_ASO_3$ to the vector $\Omega = A^{-1}\dot{A}$. The latter expression can be interpreted as the velocity of rotation of the body frame, expressed in terms of the body frame itself. It is known as the *angular velocity*. The reduction of the dynamics on TSO_3 to the dynamics on \mathfrak{so}_3 means rewriting the dynamical equations in terms of the angular velocity. The invariance of the initial equations under the SO_3 -action guarantees that the so obtained equations form a closed system, i.e. that the change of the angular velocity can be expressed as the function of the angular velocity itself and does not depend on the position.

Remark 29.4. The angular velocity has the following interpretation. First recall that the space \mathfrak{so}_3 of 3×3 skew-symmetric matrices can be identified with \mathbb{R}^3 . This identification if performed by mapping a vector $x \in \mathbb{R}^3$ to the operator $A_x(y) = x \times y$, where the latter product is the cross-product of vectors. Using this identification, one can identify the angular velocity matrix with a vector. That vector is known as the *angular velocity vector*. Its direction has a meaning of the instantaneous rotation axis, while its magnitude is the speed of rotation about that axis.

The reduction of dynamics to \mathfrak{so}_3^* can be interpreted in a similar way. Since the flows of T^*SO_3 and TSO_3 are related by the Legendre transform, the equations on \mathfrak{so}_3 and \mathfrak{so}_3^* are related by a linear operator $I: \mathfrak{so}_3 \to \mathfrak{so}_3^*$, determined by the kinetic energy at the identity element of the group. That operator I is determined by the shape of the body and is known as the *the inertia tensor*. The image of the angular velocity Ω under the inertia operator is denoted by M and is called the *angular momentum*. The Hamiltonian equation on \mathfrak{so}_3^* is thus the evolution of the angular momentum. Since it is related to the angular velocity by means of a constant operator, we can again interpret this equation as the velocity evolution separated from the position evolution.

Since the free rigid body dynamics is a Hamiltonian system on \mathfrak{so}_3^* , it is an integrable system. Indeed, the symplectic leaves in \mathfrak{so}_3^* are two-dimensional, so any Hamiltonian system on \mathfrak{so}_3^* is integrable. The corresponding first integrals are the Casimir function $\operatorname{Tr} M^2$ and the energy function.

Remark 29.5. One can show that this system is also integrable on the initial phase space T^*SO_3 . This, however, does not directly follow from Noether's principle. The problem is that although the system is SO_3 -invariant, the group SO_3 is non-commutative, so in addition to the Hamiltonian we get three first integrals f_x , f_y , f_z which do not Poisson commute with each other. Thus what we get is four non-Poissoncommuting integrals instead of three Poisson-commuting ones. This is once again a superintegrability (or *non-commutative integrability*) situation. One can in fact show that f_x and $f_x^2 + f_y^2 + f_z^2$ Poisson commute (for that one shows that the functions f_x , f_y , f_z are given by *right-invariant vector fields*, and their pairwise Poisson brackets are equal to negative Lie brackets of the corresponding Lie algebra elements), so in this case non-commutative integrability implies commutative integrability.

Exercise 29.6. Using this remark, show that generic trajectories of the free rigid body system on T^*SO_3 evolve on two-dimensional tori.

Lecture 30: Free rigid body dynamics

In the previous lecture we saw that the motion of a free rigid body with a fixed point can be regarded as a dynamical system on \mathfrak{so}_3^* which is Hamiltonian with respect to the Lie-Poison structure, with quadratic Hamiltonian function. The exact form of that quadratic function is revealed in the following exercises:

Exercise 30.1. Let $A(t) \in SO_3$ be the matrix encoding the position of (the frozen frame in) a rigid body. Define the angular velocity $\Omega \in \mathfrak{so}_3$ by $\Omega = A^{-1}\dot{A}$. Show that the kinetic energy of the body is given by

$$K = \frac{1}{2} \langle J\Omega + \Omega J, \Omega \rangle, \tag{40}$$

where the inner product on \mathfrak{so}_3 is defined by⁷

$$\langle X, Y \rangle = -\frac{1}{2} \operatorname{Tr} XY = \frac{1}{2} \operatorname{Tr} XY^{t},$$

and J is the mass tensor relative to the fixed point, which is defined as the integral over the whole body of the matrix rr^t where r is the (column) radius-vector relative to the fixed point (expressed in the body frame). The operator $\mathfrak{so}_3 \to \mathfrak{so}_3$ given by $\Omega \mapsto J\Omega + \Omega J$ is known as the *inertia tensor* (we will reveal the connection of this definition with the physical one in the next exercise).

Exercise 30.2. Consider a straight line l passing through the fixed point of the body. The moment of inertia of the body with respect to l is defined as the integral over the whole body of the squared distance to l. Show that this moment can be computed as

$$I_l = I(e, e),$$

where e is a unit vector in the direction of l, and I is a quadratic form. Show that upon identification of \mathbb{R}^3 with \mathfrak{so}_3 (see Remark 29.4) the form I coincides with the inertia tensor.

Exercise 30.3. Show that the frozen frame can be always chosen in such a way that the mass tensor J is given by a diagonal matrix. The vectors of such a frozen frame are called *principal axes of inertia*. By definition, principal axes of inertia are eigenvectors of the mass tensor. Show that upon identification of \mathbb{R}^3 and \mathfrak{so}_3 principal axes of inertia are also eigenvectors of the inertia tensor. The corresponding eigenvalues of the inertia tensor are known as *principal moments of inertia*. How are they related to the eigenvalues of the mass tensor?

Exercise 30.4. Find principal axis of inertia and principal moments for a solid bounded by an ellipsoid.

Exercise 30.5. Show that the Hamiltonian function of the free rigid body on \mathfrak{so}_3^* has the form

$$H = \frac{1}{2} \langle I^{-1}(M), M \rangle,$$

⁷The motivation for this choice of inner product is that it coincides with the standard inner product upon identification with \mathbb{R}^3 .

where M is an element of \mathfrak{so}_3^* identified with a skew-symmetric matrix by means of the form $-\frac{1}{2} \operatorname{Tr} XY$, and I is the inertia tensor. Furthermore, assuming that the frozen frame consists of principal axes of inertia, we have

$$H = rac{1}{2} \left(rac{x^2}{I_x} + rac{y^2}{I_y} + rac{z^2}{I_z}
ight),$$

where x, y, z are matrix entries of $M \in \mathfrak{so}_3^* = \mathfrak{so}_3$, and I_x , I_y , I_z are principal moments of inertia.

Using the above formula for the Hamiltonian along with (38) we get the following explicit form of the rigid body equation:

$$\dot{M} = [M, I^{-1}(\Omega)].$$

Here the angular momentum $M \in \mathfrak{so}_3^* = \mathfrak{so}_3$ is obtained from the angular velocity Ω by Legendre transform, which, in view of the form (40) of the kinetic energy, is given by $M = I(\Omega)$. Thus, the equation can be rewritten as

$$\dot{M} = [M, \Omega]$$

Identifying \mathfrak{so}_3 with \mathbb{R}^3 , we recover the classical form of the free rigid body equation:

$$\dot{m} = m \times \omega.$$

Here $m \in \mathbb{R}^3$ is the angular momentum vector, expressed in the frozen frame. By definition, this vector is obtained from the angular velocity ω by applying the inertia tensor.

Exercise 30.6. Show that the angular momentum coincides with the integral over the whole body of the quantity $r \times v$, where r is the radius-vector (relative to the fixed point) and v is the velocity. Compare this with the definition of the angular momentum in the lecture on Noether's principle.

Exercise 30.7. Using the above evolution equation of M and the definition of the angular velocity $\Omega = A^{-1}\dot{A}$ show that the matrix AMA^{-1} is constant. What's the physical meaning of the corresponding vector in \mathbb{R}^3 ? How is this related to Noether's principle?

Apart from the Hamiltonian, the free rigid body equation has another conserved quantity, namely the Casimir

$$-\frac{1}{2}\operatorname{Tr} M^2 = x^2 + y^2 + z^2.$$

Therefore, the trajectories can be described as intersections of level sets of this function (which are spheres) and level sets of the Hamiltonian (ellipsoids). We begin with describing single-point trajectories, i.e. fixed points. This can be done e.g. by computing critical points of the Hamiltonian restricted to symplectic leaves. Such points occur when the differentials of the Hamiltonian and the Casimir are collinear. A straightforward computation shows that for distinct I_x , I_y , I_z this critical set is the union of coordinate axes. By construction, this corresponds to rotations about principal axes. Another way to see that is to notice that $m \times \omega = 0$ if and only if m and ω are collinear, which again means that ω is an eigenvector of the inertia tensor I. We therefore conclude with the following:

Proposition 30.8. Rotations of a free rigid body around its principal axes of inertia are steady motions, *i.e.* for such motions the direction and speed of rotation does not change in time. Conversely, any steady rotation is a rotation about one of the principal axes.

Our further goal is to study more general, non-steady, rotations. Of particular interest are those that are close to steady ones. We would in particular like to know whether steady rotations are stable. This can be in principle done using standard ODE techniques, but it is easier to do geometrically, by intersecting spheres with ellipsoids. See e.g. wikipedia article on the tennis racket theorem (https:

//en.wikipedia.org/wiki/Tennis_racket_theorem) for visualization of such intersections. One sees that trajectories close to rotations about the axes corresponding to minimal and maximal moments of inertia are curves which stay close to the corresponding fixed point. Thus, those steady rotations are stable. On the other hand, trajectories which start close to rotations about the axis with intermediate moment of inertia go far away from the corresponding equilibrium, making that equilibrium unstable. In fact, one can see that trajectories starting close to (0, c, 0) (where we assume that $I_x > I_y > I_z$) after a certain period of time come close to the point (0, -c, 0), which means that the direction of rotation is reversed. This can be clearly seen in the video on the "Dzhanibekov effect" at the same wikipedia page (note that in this video the body is not fixed, but this is equivalent to fixing the body at the center of mass). One can also observe instability of rotations about the middle axis experimentally by trying to spin a tennis racket or a book about the corresponding axis (the gravity force present in this experiment does not play a role, because the gravity will only affect translational but not rotational motion).

Lecture 31: Integrable cases of heavy rigid body motion

In this lecture we consider the dynamics of a rigid body with a fixed point under the action of the gravity force. Here we are talking about gravity near the surface of the earth, so the gravity force is always pointing downward and its magnitude is equal to the mass times a universal constant g (acceleration due to gravity). The configuration space of this "heavy" rigid body is the same as for the free one, namely the group SO_3 . The expression for the kinetic energy is also the same, in particular it is a left-invariant function on T^*SO_3 . A new ingredient in the heavy case is the presence of a non-trivial potential.

Exercise 31.1. Show that the potential energy of a rigid body is equal to mgh, where m is the total mass, g is acceleration due to gravity, and h is the height of the center of mass relative to the fixed point (or relative to any other point which occupies a fixed position in space). In particular, if the fixed point coincides with the center of mass, then the potential energy is a constant function.

From this formula for the potential it is clear that it is invariant under rotations about the z-axis. Such rotations correspond to a subgroup

$$\begin{pmatrix} \cos(\phi) & \sin(\phi) & 0\\ -\sin(\phi) & \cos(\phi) & 0\\ 0 & 0 & 1 \end{pmatrix} \subset SO_3$$

acting on SO_3 (and hence T^*SO_3) from the left. This subgroup is isomorphic to a circle S^1 and we will denote it by S^1 .

Exercise 31.2. Derive an explicit formula for the potential as a function on SO_3 and show that it is invariant under the left action of the above subgroup S^1 .

So while the kinetic energy of the heavy rigid body still has SO_3 symmetry, the potential is only invariant under the subgroup $S^1 \subset SO_3$. For that reason in this case we cannot perform reduction by the SO_3 action, but can still reduce by the S^1 action (note that the action of S^1 on T^*SO_3 is symplectic since it is a lift of an action of S^1 on SO_3 , cf. the free rigid body case). Since T^*SO_3 has dimension six, and S^1 is one-dimensional, the reduced space T^*SO_3/S^1 is five-dimensional. It carries a reduced Poisson structure, and the dynamics of the heavy rigid body is Hamiltonian with respect to that structure. One can show that generic symplectic leaves of the Poisson structure T^*SO_3/S^1 are four-dimensional, and there is one globally defined Casimir function. **Exercise 31.3.** Use Noether's principle to show that the left action of S^1 on T^*SO_3 is given by a Hamiltonian vector field. Show that the Hamiltonian function of that vector field descends to the quotient T^*SO_3/S^1 , where it becomes a Casimir function. Does that Hamiltonian function descend to the quotient T^*SO_3/SO_3 ?

Remark 31.4. One can construct coordinates on the quotient T^*SO_3/S^1 as follows. First, for any $\xi \in T^*_ASO_3$ one defines the angular momentum $M \in \mathfrak{so}^*_3$ as the pull-back of ξ by the left translation λ_A . By construction, entries of M are left-invariant. This gives three functions on the five-dimensional space T^*SO_3/S^1 , so we need more functions to get a coordinate system (as follows from our discussion of the free rigid body case, these three functions define a coordinate system on the quotient T^*SO_3/SO_3). As additional functions one can take entries of the vector Γ which is the unit vector (0,0,1) expressed in the frozen frame. The entries of Γ are S^1 -invariant and are subject to a relation $\langle \Gamma, \Gamma \rangle = 1$. Thus, the quotient T^*SO_3/S^1 can be identified with the direct product of \mathbb{R}^3 (parametrized by the entries m_x , m_y , m_z of the angular momentum matrix) and sphere S^2 (the set of vectors $\Gamma = (\gamma_x, \gamma_y, \gamma_z) \in \mathbb{R}^3$ with $\langle \Gamma, \Gamma \rangle = 1$). It can be shown that the Poisson bracket on T^*SO_3/S^1 reads

$$\{m_x, m_y\} = m_z, \quad \{m_y, m_z\} = m_x, \quad \{m_z, m_x\} = m_y \\ \{m_x, \gamma_y\} = \gamma_z, \quad \{m_y, \gamma_z\} = \gamma_x, \quad \{m_z, \gamma_x\} = \gamma_y,$$

while other brackets of coordinate functions vanish. It follows from our construction that this bracket satisfies the Jacobi identity on the subset $\langle \Gamma, \Gamma \rangle = 1$ (because that subset is the image of the reduction map), but in fact it can be checked that the Jacobi identity holds everywhere. So, the reduced bracket naturally extends from $T^*SO_3/S^1 = \mathbb{R}^3 \times S^2$ to $\mathbb{R}^3 \times \mathbb{R}^3 = \mathbb{R}^6$. It follows from our construction that T^*SO_3/S^1 is a Poisson submanifold of \mathbb{R}^6 with respect to that bracket. Also notice that the so-obtained bracket on \mathbb{R}^6 is linear, so it can be regarded as Lie-Poisson bracket for an appropriate Lie algebra. That Lie algebra is denoted by \mathfrak{e}_3 and is isomorphic to the Lie algebra of Lie group of rigid motions of \mathbb{R}^3 (including translations). The appearance of that Lie algebra in our problem is apparently accidental: since one of the points of the body is fixed, the problem does not have translational symmetries.

The Hamiltonian of the rigid body also naturally extends to $\mathbb{R}^6 = \mathfrak{e}_3^*$, so we can regard heavy rigid body equations as equations on \mathfrak{e}_3^* (but only motions on the submanifold $\langle \Gamma, \Gamma \rangle = 1$ have actual physical meaning). This does not change much from the integrability perspective: the symplectic leaves in \mathfrak{e}_3^* are still four-dimensional, and one needs a single additional first integral to ensure complete integrability.

Since the symplectic leaves in T^*SO_3/S^1 (as well in the extended space \mathfrak{e}_3^*) are four-dimensional, we need to find one first integral in addition to the Hamiltonian function to establish complete integrability. In general, such an additional integral does not exist, see [22] and references therein. It does, however, exist in the following three special cases:

I. The Euler case. This is the case when the fixed point coincides with the center of mass. In this case the potential is a constant function (see Exercise 31.1), so it is equivalent to free rigid body motion. We already saw that system is integrable on \mathfrak{so}_3^* . Furthermore, it is integrable on T^*SO_3/S^1 as well: an additional integral can be constructed by pulling back the Casimir function on \mathfrak{so}_3^* using the natural map $T^*SO_3/S^1 \to T^*SO_3/SO_3$. The so obtained function on T^*SO_3/S^1 coincides with $m_x^2 + m_y^2 + m_z^2$, the squared magnitude of the angular momentum.

II. The Lagrange case. This is the case when two principal moments of inertia (say, I_x and I_y) coincide, and the center of mass is located on the principal axis corresponding to the third principal moment (i.e. the z axis). An example is a standard spinning top, i.e. a uniform body which is rotationally symmetric with respect to a line l through a fixed point (there is no actually fixed point in a spinning top, but the point where the top touches the table can be assumed to be fixed if the friction is large).

An additional integral in this case can be constructed, using Noether's principle, from the rotational symmetry (these rotations about the symmetry axis of the top should not be confused with rotations about the z axis in space; the latter have been already factored out).

Exercise 31.5. Show that the rotational symmetry of the Lagrange top is given by the **right** action of S^1 on SO_3 . Use that to show that this rotation and the corresponding Hamiltonian function descends to the quotient T^*SO_3/S^1 .

III. The Kovalevskaya case. This is the most "mysterious" case, with no apparent physical symmetry. In this case the ratio of moments of inertia is 2:2:1, and the center of mass is located in the plane of two principal axes with equal moments. An additional first integral in this case is a certain degree four polynomial which does not come from any physical considerations. There is plenty of literature on this beautiful dynamical system, although it seems that no real-life model was ever built. Computer simulations can be found in the film "Kovalevskaya Top", see https://av.tib.eu/media/10361. The film is accompanied by a paper found at the bottom of the page.

Lecture 32: The Toda lattice

In this lecture we will discuss the *Toda lattice*, a multidimensional integrable system introduced by Japanese physicist M. Toda in 1967 [19]. The Toda lattice consists of n unit mass particles on the real line which interact via a repelling force of magnitude e^{-r} where r is the distance between the particles. It is also assumed that every particle only interacts with its nearest neighbours, i.e. the *i*'th particle only interacts with the (i - 1)'st and (i + 1)'st. The configuration space of this system is \mathbb{R}^n coordinatized by positions q_1, \ldots, q_n of the particles. The phase space is therefore $T^*\mathbb{R}^n = \mathbb{R}^{2n}$, with coordinates $p_1, \ldots, p_n, q_1, \ldots, q_n$, where p_i 's are momenta or, which in this case is the same, velocities of the particles. The Poisson structure on the phase space is canonical, i.e. $\{p_i, q_i\} = 1$ while other brackets of coordinate functions vanish. The Hamiltonian function is

$$H = \frac{1}{2} \sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n-1} e^{q_i - q_{i+1}},$$

and the corresponding canonical Hamilton's equations are

$$\dot{q}_i = \{H, q_i\} = \frac{\partial H}{\partial p_i} = p_i, \quad \dot{p}_i = \{H, p_i\} = -\frac{\partial H}{\partial q_i} = -e^{q_i - q_{i+1}} + e^{q_{i-1} - q_i}.$$

Note that the latter formula is technically not valid for i = 1 and i = n. However, we can make it work for any i = 1, ..., n by formally setting $q_0 = -\infty$, $q_{n+1} = +\infty$.

The above Hamilton's equations can be also rewritten in the Newtonian form, as

$$\ddot{q}_i = -e^{q_i - q_{i+1}} + e^{q_{i-1} - q_i}.$$

Remark 32.1. This version of the Toda lattice is known as finite, non-periodic Toda. M. Toda himself initially considered a version with infinitely many particles indexed by $i \in \mathbb{Z}$, as well as the periodic version with Hamiltonian

$$H = \frac{1}{2} \sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} e^{q_i - q_{i+1}},$$

where q_{n+1} is, by definition, equal to q_1 . Also note that Toda's original interaction potential was $e^{-r}+r-1$. However, in the periodic and infinite cases the additional terms cancel out and do not affect the equations of motion. We will establish the integrability of the Toda lattice through a Lax representation. A Lax form for the Toda lattice was found by H. Flaschka [4], but the one that we will use is due to B. Kostant [8] (the relation between these two Lax representations is explained in Exercise 32.7 below). Define the variables

$$a_i = e^{q_i - q_{i+1}}, \quad b_i = -p_i.$$

Then b_1, \ldots, b_n determine the momenta p_1, \ldots, p_n , while a_1, \ldots, a_{n-1} determine coordinates q_1, \ldots, q_n up to simultaneous translation $q_i \mapsto q_i + c$. Accordingly, $a_1, \ldots, a_{n-1}, b_1, \ldots, b_n$ can be regarded as coordinates on the quotient of the phase space $T^*\mathbb{R}^n$ by the \mathbb{R} -action $q_i \mapsto q_i + c$. Note also that the \mathbb{R} -action preserves the Hamiltonian, so the Toda lattice can be rewritten as a Hamiltonian system on the quotient $T^*\mathbb{R}^n/\mathbb{R}$. Explicitly, we have

$$\dot{a}_i = \frac{d}{dt}e^{q_i - q_{i+1}} = a_i \frac{d}{dt}(q_i - q_{i+1}) = a_i(p_i - p_{i+1}) = a_i(b_{i+1} - b_i)$$

and

$$\dot{b}_i = -\dot{p}_i = a_i - a_{i-1}$$

So, on the quotient by the translation action, the Toda system rewrites as

$$\begin{cases} \dot{a}_i = a_i(b_{i+1} - b_i), \\ \dot{b}_i = a_i - a_{i-1}, \end{cases}$$
(41)

where, by definition, we set $a_0 = a_n = 0$, in agreement with the convention $q_0 = -\infty$, $q_{n+1} = +\infty$.

Proposition 32.2. The Toda system (41) is equivalent to the Lax equation L = [L, A], where

$$L = \begin{pmatrix} b_1 & 1 & & & \\ a_1 & b_2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-2} & b_{n-1} & 1 \\ & & & & a_{n-1} & b_n \end{pmatrix},$$
(42)

and A is the strictly lower-triangular part of L:

$$A = \begin{pmatrix} 0 & & & & \\ a_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & a_{n-2} & 0 & \\ & & & & a_{n-1} & 0 \end{pmatrix}$$

Proof. Write L as A + B, where

$$B = \begin{pmatrix} b_1 & 1 & & & \\ & b_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & b_{n-1} & 1 \\ & & & & & b_n \end{pmatrix}$$

Then

$$[L,A] = [B,A] = \begin{pmatrix} a_1 & & & \\ a_1(b_2 - b_1) & a_2 - a_1 & & \\ & \ddots & \ddots & \\ & & a_{n-2}(b_{n-1} - b_{n-2}) & a_{n-1} - a_{n-2} \\ & & & a_{n-1}(b_n - b_{n-1}) & -a_{n-1} \end{pmatrix},$$

so the equation $\dot{L} = [L, A]$ coincides with (41).

Corollary 32.3. The functions $\operatorname{Tr} L, \ldots, \operatorname{Tr} L^n$ are first integrals of the Toda system.

Exercise 32.4. Show that these functions are functionally independent. Hint: establish functional independence on the submanifold $a_1 = \cdots = a_{n-1} = 0$.

Thus, the Toda system has n independent first integrals, which is exactly the number needed for complete integrability, provided that these first integrals Poisson commute. We will obtain Poisson commutativity of those functions as a corollary of the *Adler-Kostant-Symes theorem*, which will be discussed in the next lecture.

Remark 32.5. Note that the Toda system can be considered as a system either on the symplectic manifold $T^*\mathbb{R}^n$, or on its quotient by the \mathbb{R} -action. In both cases complete integrability requires n integrals. For $T^*\mathbb{R}^n/\mathbb{R}$ this is so because its symplectic leaves are of dimension 2n-2, so we need n-1 first integrals on the leaves, plus one Casimir function to distinguish between the leaves.

Exercise 32.6. Show that $\operatorname{Tr} L$, viewed as a function on $T^*\mathbb{R}^n$, is Noether's integral corresponding to the \mathbb{R} -action, while on the quotient $T^*\mathbb{R}^n/\mathbb{R}$ it becomes a Casimir function. Also show that $\frac{1}{2}\operatorname{Tr} L^2$ is the Hamiltonian.

Exercise 32.7. Show that if a time-dependent matrix L satisfies a Lax equation L' = [L, A], and D is another time-dependent matrix, then the matrix $\tilde{L} = DLD^{-1}$ also satisfies a Lax equation: $\tilde{L}' = [\tilde{L}, \tilde{A}]$. Show that the Lax matrix (42) of the Toda lattice is conjugate to the matrix

$$\tilde{L} = \begin{pmatrix} b_1 & \sqrt{a_1} & & & \\ \sqrt{a_1} & b_2 & \sqrt{a_2} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{a_{n-2}} & b_{n-1} & \sqrt{a_{n-1}} \\ & & & & \sqrt{a_{n-1}} & b_n \end{pmatrix}$$

Hence show that \tilde{L} satisfies a Lax equation $\tilde{L}' = [\tilde{L}, \tilde{A}]$ and find the corresponding matrix \tilde{A} . This is the Lax representation found by Flaschka.

Lecture 33: The Adler-Kostant-Symes scheme

In this lecture we will describe a general approach which allows one to construct Hamiltonian systems which possess the following properties:

- They can be written in the Lax form $\dot{L} = [L, A]$. As a corollary, functions of the form $\text{Tr} L^k$ (equivalently, coefficients of the characteristic polynomial of L) are preserved by the flow.
- Moreover, the functions $\operatorname{Tr} L^k$ pairwise Poisson commute.

If, in addition, the functions $\operatorname{Tr} L^k$ are independent and there are sufficiently many of them, then this construction gives a completely integrable system. This construction is known as the Adler-Kostant-Symes (or AKS) scheme. It was developed in the works of M. Adler [1], B. Kostant [8], and W. Symes [17]. We will present an r-matrix version of the AKS scheme, due to M. Semenov-Tyan-Shanskii [16]. After discussing this general scheme we will show that the Toda lattice is a particular example and hence deduce its complete integrability.

Definition 33.1. Let \mathfrak{g} be a Lie algebra. A linear operator $r: \mathfrak{g} \to \mathfrak{g}$ is a called a *classical r-matrix* if the bracket

$$[x,y]_r = [rx,y] + [x,ry]$$

is a Lie bracket, i.e. if it satisfies the Jacobi identity (this bracket is automatically bilinear and skewsymmetric, so it is a Lie bracket if and only if it satisfies the Jacobi identity).

Example 33.2. For any Lie algebra \mathfrak{g} , the identity operator is a classical *r*-matrix. Indeed, for r = id we have

$$[x,y]_r = 2[x,y]$$

and the Jacobi identity for this bracket follows from the Jacobi identity for the [,] bracket.

Example 33.3 (Main example). Let \mathfrak{g} be a Lie algebra, and let \mathfrak{g}_{\pm} be its Lie subalgebras such that, as a vector space, \mathfrak{g} can be decomposed as $\mathfrak{g} = \mathfrak{g}_{+} \oplus \mathfrak{g}_{-}$. Let p_{\pm} be the projectors to \mathfrak{g}_{\pm} respectively, and let

$$r = \frac{1}{2}(p_+ - p_-)$$

Then r is a classical r-matrix. Indeed, denote $x_{\pm} = p_{\pm}(x)$. Then

$$[x,y]_r = [rx,y] + [x,ry] = \frac{1}{2} \left([x_+ - x_-, y_+ + y_-] + [x_+ + x_-, y_+ - y_-] \right) = [x_+, y_+] - [x_-, y_-].$$

The Jacobi identity for this bracket follows from the exercise below:

Exercise 33.4. Denote by \mathfrak{g}_r the algebra \mathfrak{g} equipped with the *r*-bracket from this example. Show that the mapping $\mathfrak{g}_r \to \mathfrak{g}_+ \oplus \mathfrak{g}_-$ given by $x \mapsto (x_+, -x_-)$ is an isomorphism of algebras (not just vector spaces). As a corollary, the *r*-bracket satisfies the Jacobi identity.

Example 33.5. Let $\mathfrak{g} = \mathfrak{gl}_n$ be the Lie algebra of $n \times n$ matrices. Then we have a vector space decomposition $\mathfrak{g} = \mathfrak{g}_+ \oplus \mathfrak{g}_-$ where \mathfrak{g}_+ consists of strictly lower triangular matrices, and \mathfrak{g}_- consists of upper triangular matrices. As we will see below, the corresponding *r*-matrix is related to the Toda lattice.

Example 33.6. For $\mathfrak{g} = \mathfrak{gl}_n$ we also have another decomposition $\mathfrak{g} = \mathfrak{g}_+ \oplus \mathfrak{g}_-$ where $\mathfrak{g}_+ = \mathfrak{so}_n$ consists of skew-symmetric matrices, and \mathfrak{g}_- consists of upper triangular matrices. This decomposition also arises in the study of the Toda lattice.

Given an *r*-matrix on a Lie algebra \mathfrak{g} , we have two different Lie brackets in that algebra: the standard bracket [,] and the *r*-bracket $[,]_r$. Accordingly, on the dual space \mathfrak{g}^* we obtain two Lie-Poisson brackets, which we denote by $\{,\}$ and $\{,\}_r$. We will also use the notation π and π_r for the corresponding Poisson tensors, and X_H, X_H^r for Hamiltonian vector fields with Hamiltonian H with respect to the two brackets. Let also G and G_r be two Lie groups associated with \mathfrak{g} : one corresponding to the [,] bracket, and another one corresponding to the $[,]_r$ bracket.

Theorem 33.7 (AKS theorem). Assume that \mathfrak{g} is a Lie algebra endowed with an r-matrix r. Then the following statements hold.

1. Assume that $H \in C^{\infty}(\mathfrak{g}^*)$ is an invariant of the coadjoint action of G on \mathfrak{g}^* , i.e. a Casimir function of the $\{,\}$ bracket. Then the Hamiltonian vector field X_H^r of H with respect to the $\{,\}_r$ bracket is given by

$$\dot{x} = -\mathfrak{ad}^*_{rdH(x)}x,\tag{43}$$

where $\mathfrak{a}\mathfrak{d}^*$ is the coadjoint action corresponding to the [,] bracket. If, in addition, \mathfrak{g} has an invariant (with respect to the adjoint action of G) inner product, then, upon identification of \mathfrak{g}^* and \mathfrak{g} by means of that product the above equation takes the Lax form

$$\dot{x} = [x, rdH(x)]$$

2. All invariants of the coadjoint action of G, i.e. Casimirs of the $\{,\}$ bracket, Poisson-commute with respect to the $\{,\}_r$ bracket.

Remark 33.8. From the second statement it follows that the restrictions of invariants of the coadjoint action to any symplectic leaf of the $\{,\}_r$ bracket Poisson commute with respect to that bracket. If, in addition, the number of independent invariants is half the dimension of the leaf, then we obtain an integrable system. As a Hamiltonian function, one can take any of the Poisson-commuting invariants. The tori of that integrable system can be described as joint level sets of Casimirs of *both brackets*. In other words, they are intersections of symplectic leaves of the $\{,\}$ bracket and symplectic leaves of the $\{,\}_r$ bracket. Here we use that symplectic leaves coincide with joint level sets of Casimirs. This is usually the case for generic leaves.

Proof of Theorem 33.7. We begin with the first statement. The Poisson tensor of the bracket $\{,\}_r$ at $x \in \mathfrak{g}^*$ is given by $\xi \mapsto -{}^r \mathfrak{ad}_{\xi}^* x$, where ${}^r \mathfrak{ad}_{\xi}^*$ is the coadjoint operator corresponding to the $[,]_r$ bracket, i.e. the negative dual of the operator ${}^r \mathfrak{ad}_{\xi}$ given by ${}^r \mathfrak{ad}_{\xi} \eta = [\xi, \eta]_r$. Therefore, the Hamiltonian vector field X_H^r is given by

$$\dot{x} = -{}^{r}\mathfrak{a}\mathfrak{d}^*_{dH(x)}x. \tag{44}$$

Further, we have

$$\mathfrak{fad}_{\xi}\eta = [\xi,\eta]_r = [r\xi,\eta] + [\xi,r\eta] = \mathfrak{ad}_{r\xi}\eta + (\mathfrak{ad}_{\xi} \circ r)(\eta), \tag{45}$$

where \mathfrak{ad} is the adjoint operator for the [,] bracket, i.e. $\mathfrak{ad}_{\xi}(\eta) = [\xi, \eta]$. Now, rewriting (45) as

$${}^{r}\mathfrak{ad}_{\xi} = \mathfrak{ad}_{r\xi} + \mathfrak{ad}_{\xi} \circ r$$

and taking the negative dual of both sides, we get

$${}^{r}\mathfrak{ad}_{\xi}^{*} = \mathfrak{ad}_{r\xi}^{*} + r^{*} \circ \mathfrak{ad}_{\xi}^{*}.$$

Plugging this into (44) we obtain

$$\dot{x} = -\mathfrak{a}\mathfrak{d}^*_{rdH(x)}x - r^*(\mathfrak{a}\mathfrak{d}^*_{dH(x)}x).$$

But $\mathfrak{a0}_{dH(x)}^* x$ is the negative of the Hamiltonian vector field X_H , and since H is a Casimir, that vector field vanishes. This gives us the desired expression for the Hamiltonian vector field X_H^r of H with respect to the $\{,\}_r$ bracket. Further, in case when we have an invariant inner product, we can identify \mathfrak{g}^* with \mathfrak{g} and $\mathfrak{a0}^*$ with $\mathfrak{a0}$, so we get

$$\dot{x} = -\mathfrak{a}\mathfrak{d}_{rdH(x)}x = -[rdH(x), x] = [x, rdH(x)]$$

as desired.

To prove the second statement notice that by (43) the Hamiltonian vector field X_H^r belongs to the image of the Poisson tensor π . Therefore, the trajectories of that vector field stay within the symplectic leaves of the $\{,\}$ bracket. So, any Casimir function \tilde{H} of the $\{,\}$ bracket is constant along trajectories of the vector field X_H^r , which means that $\{H, \tilde{H}\}_r = 0$ for arbitrary Casimirs H and \tilde{H} of the $\{,\}$ bracket, q.e.d.

Exercise 33.9. Deduce the second part of the AKS theorem directly from the definition of the Lie-Poisson bracket.

We now apply the AKS scheme to the Toda lattice. To that end, we consider an *r*-matrix on \mathfrak{gl}_n from Example 33.5. As a Casimir of the Lie-Poisson bracket, we take the function

$$H = \frac{1}{2} \operatorname{Tr} L^2.$$

Exercise 33.10. The differential of the function $\operatorname{Tr} L^k$ on \mathfrak{gl}_n is given by $d\operatorname{Tr} L^k = kL^{k-1}$. Here one identifies \mathfrak{gl}_n and \mathfrak{gl}_n^* by means of the form $\operatorname{Tr} XY$.

Thus, by the AKS theorem the Hamiltonian flow of H with respect to the r-bracket is

1

$$\dot{L} = [L, rdH(L)] = \left[L, \frac{1}{2}(L_{+} - L_{-})\right],$$

where L_{\pm} are, respectively, strictly lower triangular and upper triangular parts of L. Furthermore, this can be rewritten as

$$\left[L, \frac{1}{2}(L_{+} - L_{-})\right] = \left[L, L_{+} - \frac{1}{2}(L_{+} + L_{-})\right] = \left[L, L_{+} - \frac{1}{2}L\right] = [L, L_{+}].$$
equation

We thus get an equation

$$\dot{L} = [L, L_+] \tag{46}$$

which is Hamiltonian with respect to the *r*-bracket with Hamiltonian function $H = \frac{1}{2} \text{Tr} L^2$. Furthermore, by the second part of the AKS theorem traces of powers of *L* Poisson-commute (with respect to the *r*-bracket). This is almost what we need: equation (46) coincides with the Toda lattice equation when restricted to matrices *L* of the form (42). What is however not clear is why the restrictions of traces of powers of *L* to such special matrices still Poisson-commute. In fact, so far we have no relation between the bracket on arbitrary matrices (the *r*-bracket) and the bracket on matrices of the form (42) (which is obtained from the cotangent bundle bracket by means of reduction by translations). This relation is explained in the following exercises:

Exercise 33.11. Let \mathfrak{g} be a Lie algebra. Prove that a vector subspace $\mathfrak{h} \subset \mathfrak{g}^*$ is a Poisson submanifold (see Exercise 23.7) if if and only if its annihilator \mathfrak{h}^{\perp} is an ideal in \mathfrak{g} (i.e. $[x, y] \in \mathfrak{h}^{\perp}$ for any $x \in \mathfrak{g}$ and any $y \in \mathfrak{h}^{\perp}$). Furthermore, prove that an affine subspace $a + \mathfrak{h} \subset \mathfrak{g}^*$ (where $a \in \mathfrak{g}^*$ is fixed, and $\mathfrak{h} \subset \mathfrak{g}^*$ is a vector subspace) is a Poisson submanifold if and only if the annihilator \mathfrak{h}^{\perp} is an ideal in \mathfrak{g} and, in addition, the linear function $a \in \mathfrak{g}^*$ vanishes on the subspace $[\mathfrak{g}, \mathfrak{h}^{\perp}]$ (defined as the span of all commutators of the form [x, y] where $x \in \mathfrak{g}$ and $y \in \mathfrak{h}^{\perp}$).

Exercise 33.12. Deduce that the affine subspace of matrices of the form (42) is a Poisson submanifold of \mathfrak{gl}_n under the *r*-bracket (with *r*-matrix from Example 33.5). Further show that the restriction of the *r*-bracket to that submanifold coincides with the bracket obtained from reduction $T^*\mathbb{R}^n/\mathbb{R}$.

From the latter exercise it follows that the traces of powers are commuting first integrals of the Toda lattice (with respect to the natural, "mechanical" bracket). As a corollary, the Toda lattice is a completely integrable system.

Exercise 33.13. Deduce integrability of the Toda lattice using the Lax representation obtained in Exercise 32.7 and the *r*-matrix from Example 33.6.

Lecture 34: Factorization theorem

In this lecture we will provide a general method, called *factorization*, to solve integrable systems constructed by the AKS scheme. We will consider the case of the Lie algebra \mathfrak{gl}_n with the *r*-matrix given by the difference of two projections, but in fact a similar approach works in the general setting too.

Consider any vector space decomposition $\mathfrak{gl}_n = \mathfrak{g}_+ \oplus \mathfrak{g}_-$, where \mathfrak{g}_{\pm} are subalgebras. Then

$$r = \frac{1}{2}(p_+ - p_-),$$

where p_{\pm} are projectors onto \mathfrak{g}_{\pm} , is an *r*-matrix on \mathfrak{gl}_n . Therefore, by the AKS theorem any conjugationinvariant function H on \mathfrak{gl}_n gives rise to a Lax flow

$$\dot{L} = [L, rdH(L)] = [L, (dH(L))_+].$$
(47)

All such flows commute with each other. We want to obtain a solution of (47) with initial condition $L(0) = L_0$. Consider first the case of a more general Lax equation

$$\dot{L} = [L, A]. \tag{48}$$

We look for a solution with initial condition $L(0) = L_0$ in the form $L(t) = U(t)^{-1}L_0U(t)$, where U(0) = Id(cf. Exercise 27.5). Plugging this into the Lax equation, we see that U(t) must satisfy

$$U^{-1}U' = A \quad \Leftrightarrow \quad U' = UA. \tag{49}$$

Assuming that the dependence of A on time is already known, this is a system of linear differential equations with time-dependent coefficients. Therefore, it has a solution with initial condition U(0) = Id, which means that solutions of the Lax equation (48) can indeed be represented in the form

$$L(t) = U(t)^{-1} L_0 U(t)$$

with U(t) being a solution to (49). We now apply this to solve (47). We need to find a time-dependent matrix U(t) such that

$$U^{-1}U' = (dH(L))_+.$$
(50)

To that end, consider the subgroups $G_{\pm} \subset GL_n$ corresponding to the subalgebras \mathfrak{g}_{\pm} . It can be shown that since $\mathfrak{gl}_n = \mathfrak{g}_+ \oplus \mathfrak{g}_-$, any element $X \in GL_n$ sufficiently close to the identity matrix has a unique factorization of the form $X = X_+X_-$, where $X_{\pm} \in G_{\pm}$ (for uniqueness we need to assume that G_{\pm} are connected). Moreover, the set of all elements of GL_n admitting such a representation is dense in the connected component of the identity (i.e. in the set of matrices with positive determinant).

Example 34.1. Let \mathfrak{g}_+ be the subalgebra of strictly lower-triangular matrices, and \mathfrak{g}_- be the subalgebra of upper-triangular matrices. Then the corresponding subgroup G_+ consists of lower-triangular unidiagonal matrices, while G_- consists of invertible upper-triangular matrices. The claim in this case is that almost any matrix X (regardless of the sign of the determinanant) admits a decomposition $X = X_+X_-$, where X_+ is lower-triangular unidiagonal, and X_- is upper triangular. This decomposition is closely related to *Gauss decomposition* and can be constructed by Gauss elimination. Indeed, any matrix X can be brought to an upper triangular form by elementary row operations, and for almost all matrices this can be done without swapping rows. Such elementary operations without swapping rows correspond to multiplying X by a lower-triangular unidiagonal matrix L from the left, so we have that LX = U is upper-triangular, and $X = L^{-1}U$ is the desired factorization.

Example 34.2. For \mathfrak{g}_+ consisting of skew-symmetric matrices and \mathfrak{g}_- consisting of upper-triangular matrices we get the QR decomposition X = QR, where Q is orthogonal and R is upper-triangular. Such a decomposition exists for any invertible matrix X and can be constructed by applying the Gram-Schmidt orthogonalization process to columns of X.

Exercise 34.3. Let $Q \subset GL_n$ be a subgroup, and let $\mathfrak{q} \subset \mathfrak{gl}_n$ be its Lie algebra, i.e. the tangent space at the identity. Show that the tangent space T_XQ , where $X \in K$ consists of matrices of the form XY, where $Y \in \mathfrak{q}$. Hence show that the solution U(t) to the equation

$$U^{-1}U' = X$$

with $X(t) \in \mathfrak{q}$ and $U(0) = \mathrm{Id}$, belongs to the subgroup Q.

From this exercise it follows that the solution U to (50) belongs to the subgroup G_+ . To find the solution, we rewrite (50) as

$$U^{-1}U' = dH(L) + R_{-} \tag{51}$$

where R_{-} is an element of \mathfrak{g}_{-} whose precise nature is of no interest to us. Now recall that H is a conjugation-invariant function.

Exercise 34.4. Prove that for any conjugation-invariant function H on \mathfrak{gl}_n one has

$$dH(U^{-1}XU) = U^{-1}dH(X)U.$$

Note that this result is immediate for $H(X) = \operatorname{Tr} X^k$, because in that case we have

$$dH(U^{-1}XU) = k \operatorname{Tr} (U^{-1}XU)^{k-1} = U^{-1}(k \operatorname{Tr} X^{k-1})U = U^{-1}dH(X)U.$$

More generally, show that if H is an invariant of the coadjoint representation of a group G on the dual \mathfrak{g}^* of its Lie algebra, then $dH(\operatorname{Ad}_q^* x) = \operatorname{Ad}_g dH(x)$ for any $x \in \mathfrak{g}$ and $g \in G$.

Using this exercise, and denoting the constant matrix $dH(L_0)$ by X we rewrite (51) as

$$U^{-1}U' = U^{-1}XU + R_- \quad \Leftrightarrow \quad U' - XU = UR_-.$$

Now multiply the latter equation from the left by an integrating factor $\exp(-tX)$ to make the left hand side a complete derivative. This gives $V' = VR_-$, where $V = \exp(-tX)U$. Now notice that since $R_- \in \mathfrak{g}_$ and $V(0) = \mathrm{Id}$, according to Exercise 34.3 we have $V(t) \in G_-$. So,

$$\exp(tX) = UV^{-1}$$

where $U \in G_+$ and $V \in G_-$. This gives us a way to find U and thus the solution of the initial Lax equation:

Theorem 34.5 (Factorization theorem). Consider the Lax equation

$$\dot{L} = [L, (dH(L))_+],$$

where H is a conjugation invariant function on \mathfrak{gl}_n , and $dH(L)_+$ is the projection of its differential onto a subalgebra \mathfrak{g}_+ along a complementary subalgebra \mathfrak{g}_- . Then the solution to this equation with initial condition $L(0) = L_0$ is given by

$$L(t) = U(t)^{-1}L_0U(t)$$
(52)

where U(t) is found from the factorization

$$\exp(t \cdot dH(L_0)) = U(t)V(t)^{-1}$$

with $U(t) \in G_+$, $V(t) \in G_-$. Here G_{\pm} are the subgroups of GL_n corresponding to the subalgebras \mathfrak{g}_{\pm} .

Remark 34.6. It follows from Exercise 34.4 that L_0 commutes with $dH(L_0)$ which allows one to rewrite the above solution formula as

$$L(t) = V(t)^{-1} \exp(-tdH(L_0))L_0 \exp(tdH(L_0))V(t) = V(t)^{-1}L_0V(t).$$

Also note that while we were assuming in the derivation of the factorization formula that the Lie algebra \mathfrak{g} is \mathfrak{gl}_n , a simple modification of the above argument works for any Lie algebra with an invariant inner product, with conjugation in (52) replaced by the adjoint action. There is also a more general version of the factorization theorem that applies for *r*-matrices not expressible as difference of two projectors [16].

Lecture 35: Asymptotic behavior of the Toda lattice

In this lecture we will use the factorization method to solve the Toda lattice equations for two particles. As a result, we will see that the Toda lattice has very special asymptotic behavior, known as *scattering*.

For the sake of simplicity, we consider a special initial condition, namely we assume that both particles are initially at rest: $p_1(0) = p_2(0) = 0$, on distance α from each other. Due to translation invariance of the system we can further assume that the initial positions of particles are $q_2(0) = \alpha/2$ and $q_1(0) = -\alpha/2$. Note also that the total momentum $p_1 + p_2$ is conserved (cf. Exercise 32.6), so

$$\frac{d}{dt}(q_1(t) + q_2(t)) = p_1(t) + p_2(t) = p_1(0) + p_2(0) = 0,$$

which means that

$$q_1(t) + q_2(t) = q_1(0) + q_2(0) = 0,$$

i.e. for our special solution the position of the center of mass is also conserved (but it is not a conserved quantity for the general solution!).

We now apply the factorization method to find the solution of Toda equations with given initial conditions. Denote $\lambda = e^{-\alpha/2}$. Then the Lax matrix (42) at time t = 0 is given by

$$L_0 = \left(\begin{array}{cc} 0 & 1\\ \lambda^2 & 0 \end{array}\right).$$

According to the factorization theorem, we need to compute the matrix $\exp tL_0$ (note that for the Toda Hamiltonian we have dH(L) = L). Since tL_0 is a 2 × 2 matrix, we have

$$\exp tL_0 = at \cdot L_0 + b \cdot \mathrm{Id}$$

for certain constants $a, b \in \mathbb{R}$. The latter equation will still hold if we replace L_0 with its eigenvalues $\pm \lambda$, so

$$\begin{cases} a\lambda t + b = e^{\lambda t}, \\ -a\lambda t + b = e^{-\lambda t} \end{cases}$$

from where we find that

$$a = \frac{\sinh \lambda t}{\lambda t}, \quad b = \cosh \lambda t$$

and

$$\exp tL_0 = \frac{\sinh \lambda t}{\lambda} \cdot L_0 + \cosh \lambda t \cdot \mathrm{Id} = \begin{pmatrix} \cosh \lambda t & \lambda^{-1} \sinh \lambda t \\ \lambda \sinh \lambda t & \cosh \lambda t \end{pmatrix}.$$

To bring this matrix to an upper-triangular form we need to add to the second row the first one multiplied by $-\lambda \tanh \lambda t$. Therefore, in the factorization $\exp(tL_0) = UV^{-1}$, where U is lower-triangular unidiagonal and L is upper-triangular, we have

$$U(t) = \left(\begin{array}{cc} 1 & 0\\ -\lambda \tanh \lambda t & 1 \end{array}\right).$$

We now find the solution of the Toda system using formula (52):

$$L(t) = U(t)^{-1}L_0U(t) = \begin{pmatrix} \lambda \tanh \lambda t & 1\\ \lambda^2 \operatorname{sech}^2 \lambda t & -\lambda \tanh \lambda t \end{pmatrix}.$$

From this we find that the dependence of momenta of particles on time is given by

$$p_1(t) = -\lambda \tanh \lambda t, \quad p_2(t) = \lambda \tanh \lambda t.$$

As for coordinates, we have

$$e^{q_1-q_2} = \lambda^2 \operatorname{sech}^2 \lambda t$$

which, in view of the identity $q_1 + q_2 = 0$, gives

$$q_1(t) = -\frac{\alpha}{2} - \ln(\cosh \lambda t), \quad q_2(t) = \frac{\alpha}{2} + \ln(\cosh \lambda t).$$

These expressions can of course be also found from the formulas for $p_1(t), p_2(t)$ using integration and initial conditions. A particular simple solution of the Toda lattice corresponding to $\alpha = 0$ is given by

$$q_1(t) = -\ln(\cosh \lambda t), \quad q_2(t) = \ln(\cosh \lambda t).$$
(53)

Exercise 35.1. Derive this solution from the Toda lattice equations in a, b variables. Hint: eliminate two out of three variables using conservation of energy and momentum.

We would now like to study the asymptotic behavior of the Toda lattice. First observe that the velocities $p_1(t) = -\lambda \tanh \lambda t$, $p_2(t) = \lambda \tanh \lambda t$ have finite limits as $t \to \pm \infty$. In other words, for large absolute value of t the particles in the Toda lattice move as if they do not interact. This is not too surprising because

$$\lim_{t \to +\infty} (q_2(t) - q_1(t)) = +\infty,$$

so for large absolute value of t the interaction force $e^{q_1-q_2}$ is indeed weak. This type of asymptotically free motion is known as *scattering*.

Another thing that we observe is that for $t \to +\infty$ we have

$$\lim_{t \to +\infty} p_1(t) = -\lambda, \quad \lim_{t \to +\infty} p_2(t) = +\lambda,$$

while for $t \to -\infty$ we have

$$\lim_{t \to -\infty} p_1(t) = +\lambda, \quad \lim_{t \to -\infty} p_2(t) = -\lambda.$$

This means two things. First, this observation provides a geometric meaning for the spectrum of the Lax matrix L (and hence for the integrals of the Toda lattice which are symmetric functions of the spectrum): the eigenvalues of L are asymptotic momenta (velocities) of the particles⁸. Second, we see

⁸In general the eigenvalues of L are equal to negative asymptotic momenta. In our case this does not matter because the spectrum is symmetric with respect to the origin.



Figure 22: Collision of elastic balls and the Toda lattice.

that the particles in the Toda lattice exchange momenta as they scatter. This means that the interaction of particles in the Toda lattice for n = 2 is similar to a collision of two elastic balls moving along a straight line: the balls also exchange momenta when they collide. Figure 22 shows two graphs. The graph on the left shows position functions for two elastic balls of equal mass moving towards each other with unit speeds and colliding at the origin. The graph on the right shows position functions for the particular solution (53) of the Toda lattice with two particles. As one can see, the graphs look alike. There is, however, one remarkable difference in the asymptotic behavior of the Toda lattice as compared to elastic balls. Namely, in collision of elastic ball the trajectory of the second ball after the collision exactly continues the trajectory of the first ball before the collision. In other words, the position function of the first ball when $t \to +\infty$ is the same as the position function of the second one for $t \to -\infty$. For the Toda lattice this is not the case. Indeed, the $t \to +\infty$ asymptote of the first particle trajectory is parallel to the $t \to -\infty$ asymptote of the second particle trajectory, but these two lines are not the same (Figure 23). This phenomenon is known as *phase shift* and is common for integrable systems with scattering behavior.

Exercise 35.2. Compute the value of the phase shift for solution (53).



Figure 23: Asymptotes of particle trajectories in the Toda lattice.

For general solutions of the Toda lattice with n particles one has the following result due to J. Moser [14]:

Theorem 35.3. The motion in the finite non-periodic Toda lattice is asymptotically free. Furthermore, if $\lambda_1 < \cdots < \lambda_n$ are the eigenvalues of the Lax matrix, then the asymptotic velocities of particles are given by $p_1 = -\lambda_1, \ldots, p_n = -\lambda_n$ when $t \to -\infty$ and $p_1 = -\lambda_n, \ldots, p_n = -\lambda_1$ when $t \to +\infty$.

In other words, in process of scattering the *i*'th particle eventually exchanges its momentum with particle n + 1 - i. Again, this is similar to what happens in a collision of *n* elastic balls (we assume that collisions happen one by one and not simultaneously). However, in contrast to balls, the particles in the Toda lattice exhibit a phase shift.

Exercise 35.4. Prove Moser's theorem in the case of two particles. Hint: either write down the general solution explicitly, or study the level sets of the energy function as the function of a_1, b_1, b_2 on the plane of the constant momentum $b_1 + b_2 = \text{const.}$

References

- [1] M. Adler. On a trace functional for formal pseudo-differential operators and the symplectic structure of the Korteweg-de Vries type equations. *Invent. Math.*, 50(3):219–248, 1978.
- [2] V.I. Arnold. Mathematical methods of classical mechanics, volume 60. Springer Science & Business Media, 2013.
- [3] S. Benayadi and A. Elduque. Classification of quadratic Lie algebras of low dimension. Journal of Mathematical Physics, 55(8):081703, 2014.
- [4] H. Flaschka. The Toda lattice. II. Existence of integrals. *Physical Review B*, 9(4):1924, 1974.
- [5] P. Griffiths and J. Harris. A Poncelet theorem in space. Commentarii Mathematici Helvetici, 52(2):145–160, 1977.
- [6] B. Khesin and S. Tabachnikov. Fun problems in geometry and beyond. arXiv:1912.05740, 2019.
- B. Khesin and R. Wendt. The geometry of infinite-dimensional groups, volume 51. Springer Science & Business Media, 2008.
- [8] B. Kostant. The solution to a generalized Toda lattice and representation theory. Advances in Mathematics, 34(3):195–338, 1979.
- [9] V. Kovachev. Smoothness of the billiard ball map for strictly convex domains near the boundary. Proceedings of the American Mathematical Society, 103(3):856–860, 1988.
- [10] S. Lavau. A short guide through integration theorems of generalized distributions. Differential Geometry and its Applications, 61:42–58, 2018.
- [11] M. Levi and S. Tabachnikov. The Poncelet grid and billiards in ellipses. The American Mathematical Monthly, 114(10):895–908, 2007.
- [12] J. Marsden and A. Weinstein. Reduction of symplectic manifolds with symmetry. Reports on mathematical physics, 5(1):121–130, 1974.

- [13] J.E. Marsden and T. Ratiu. Reduction of Poisson manifolds. Letters in Mathematical Physics, 11(2):161–169, 1986.
- [14] J. Moser. Finitely many mass points on the line under the influence of an exponential potential an integrable system. In Dynamical systems, theory and applications, pages 467–497. Springer, 1975.
- [15] S. Schmitt. The quotient manifold theorem, lecture notes, https://www.mathi.uniheidelberg.de/~lee/stephanss16.pdf.
- [16] M.A. Semenov-Tyan-Shanskii. What is a classical r-matrix? Functional analysis and its applications, 17(4):259–272, 1983.
- [17] WW Symes. Systems of Toda type, inverse spectral problems, and representation theory. Inventiones mathematicae, 59(1):13–51, 1980.
- [18] S. Tabachnikov. Geometry and billiards. American Mathematical Society, 2005.
- [19] M. Toda. Vibration of a chain with nonlinear interaction. In Selected Papers of Morikazu Toda, pages 97–102. World Scientific, 1993.
- [20] A.P. Veselov. Integrable maps. Russian Mathematical Surveys, 46(5):1, 1991.
- [21] A. Weinstein. The local structure of Poisson manifolds. Journal of differential geometry, 18(3):523– 557, 1983.
- [22] S.L. Ziglin. The absence of an additional real-analytic first integral in some problems of dynamics. Functional Analysis and Its Applications, 31(1):3–9, 1997.