Math 177: Additional chapters

Yakov Eliashberg

May 2018

Contents

Chapter 1

What is a differential equation?

1.1 Preliminaries

Differential equations and system of equations are equations or system of equations involving derivatives of unknown functions. If all the unknown functions are of the same one variable then the differential equations are called *ordinary*. In the case of functions of more than one variable one speaks of *partial* differential equations.

Thus any system of ordinary differential equations can be written as

$$F(t, u(t), u'(t), u''(t), \dots, u^{(k)}(t)) = 0, \tag{1.1.1}$$

$t \in [a, b]$, where $u : [a, b] \to \mathbb{R}^m$ is a vector-valued function, and F is a map of a domain U in the space $\mathbb{R}^N$, $N = km + 2$ to $\mathbb{R}^l$ for some integer l.

An important observation is that it is always possible to equivalently rewrite the system (1.1.1) to involve only the *first derivatives* of the unknown functions.

Indeed, the system

$$\begin{aligned}
&F(t, u(t), v_1(t), v_2(t), \ldots, v_{k-1}(t), v'_{k-1}(t)) = 0,\\
&u'(t) = v_1(t),\\
&v'_1(t) = v_2(t),\\
&\ldots\\
&v'_{k-2}(t) = v_{k-1}(t),
\end{aligned}$$

$t \in [a, b]$, $u, v_1, \ldots, v_{k-1} : [a, b] \to \mathbb{R}^m$, is equivalent to the system (1.1.1).

Let us stress the point that when dealing with concrete equations this transformation is not always the best way of action. However, in many cases it is, and also for theoretical purposes considering the systems of first order differential equations is sufficient and we will usually do that. In other words, we will be studying the systems

$$F(t, u(t), u'(t)) = 0, \tag{1.1.2}$$

$t \in [a, b]$, $u : [a, b] \to \mathbb{R}^m$, $F : U \to \mathbb{R}^l$, where U is a domain in $\mathbb{R}^{2k+1}$.

1.2 Vector fields

A *vector field* v on a domain $U \subset V$ is a function which associates to each point $x \in U$ a vector $v(x) \in V_x$, i.e. a vector originated at the point x.

Let v be a vector field on a domain $U \in V$. If we fix a basis in V and parallel transport this basis to all spaces $V_x, x \in V$, then for any point $x \in V$ the vector $v(x) \in V_x$ is described by its coordinates $(v_1(x), v_2(x), \ldots, v_n(x))$. Therefore, to define a vector field on U is the same as to define n functions $v_1, \ldots, v_n$ on U, i.e. to define a map $(v_1, \ldots, v_n) : U \to \mathbb{R}^n$. We call a vector field v C^k-smooth if the functions $v_1, \ldots, v_n$ are smooth on U.

Thus, if a basis of V is fixed, then the difference between the maps $U \to \mathbb{R}^n$ and vector fields on U is just a matter of geometric interpretation. When we speak about a vector field

v we view $v(x)$ as a vector in V_x, i.e. originated at the point $x \in U$. When we speak about a map $v : U \to \mathbb{R}^n$ we view $v(x)$ as a point of the space V, or as a vector with its origin at $\mathbf{0} \in V$.

Vector fields naturally arise in a context of Physics, Mechanics, Hydrodynamics, etc. as force, velocity and other physical fields.

There is another very important interpretation of vector fields as *first order differential operators*.

Let $C^\infty(U)$ denote the *vector space of infinitely differentiable functions* on a domain $U \subset V$. Let v be a C^∞-smooth vector field on V. We associate with v a linear operator

$$D_v : C^\infty(U) \to C^\infty(U),$$

given by the formula

$$D_v(f) = df(v),\ \ f \in C^\infty(U).$$

In other words, we compute at any point $x \in U$ the directional derivative of f in the direction of the vector $v(x)$. Clearly, the operator D_v is linear: $D_v(af + bg) = aD_v(f) + bD_v(g)$ for any functions $f, g \in C^\infty(U)$ and any real numbers $a, b \in \mathbb{R}$. It also satisfies the *Leibniz rule*:

$$D_v(fg) = D_v(f)g + fD_v(g).$$

In view of the above correspondence between vector fields and first order differential operators it is sometimes convenient just to view a vector field as a differential operator. Hence, when it will not be confusing we may drop the notation D_v and just directly apply the vector v to a function f (i.e. write $v(f)$ instead of $D_v(f)$).

Let $v_1, \dots, v_n$ be a basis of V, and $x_1, \dots, x_n$ be the coordinate functions in this basis. We would like to introduce the notation for the vector field obtained from vectors $v_1, \dots, v_n$ by parallel transporting them to all points of the domain U. To motivate the notation which we are going to introduce, let us temporarily denote these vector fields by $\mathbf{v}_1, \dots, \mathbf{v_n}$. Observe that $D_{\mathbf{v}_i}(f) = \frac{\partial f}{\partial x_i},\ i = 1, \dots, n$. Thus the operator $D_{\mathbf{v}_i}$ is just the operator $\frac{\partial}{\partial x_i}$ of taking

i-th partial derivative. Hence, viewing the vector field $\mathbf{v_i}$ as a differential operator we will just use the notation $\frac{\partial}{\partial x_i}$ instead of $\mathbf{v}_i$. Given any vector field v with coordinate functions $a_1, a_2, \dots, a_n : U \to \mathbb{R}$ we have

$$D_v(f)(x) = \sum_{i=1}^{n} a_i(x) \frac{\partial f}{\partial x_i}(x), \text{ for any } f \in C^\infty(U),$$

and hence we can write $v = \sum\limits_{i=1}^{n} a_i \frac{\partial}{\partial x_i}$. Note that the coefficients a_i here are *functions* and not constants.

1.3 Differential equations as vector fields

If $m = l$, i.e. the number of equations is equal to the number of unknown functions the system is called *determined*. If $l > m$ it is called *over-dertermined* and if $l < m$ *under-determined*. We will be dealing in this class exclusively with determined systems.

More precisely, for determined system one usually imposes an additional condition, that the minor of the Jacobi matrix of the map $F : U \to \mathbb{R}^l$ corresponding to the last m coordinates does not vanish at every point $(t, u, y) \in U \subset \mathbb{R}^{2m+1} = \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m$ for which $F(t, u, y) = 0$. Then according to the implicit functions locally near each such point the system (1.1.2) can be solved with respect to the derivatives, i.e. written in the form

$$u'(t) = v(t, u(t)), \tag{1.3.1}$$

$t \in [a, b]$, $u : \mathbb{R}^m \to \mathbb{R}, v : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$.

Let us consider first the case when v is independent of t, i.e. the system has the form

$$u'(t) = v(u(t)), \tag{1.3.2}$$

$t \in [a, b]$, $u, v : \mathbb{R}^m \to \mathbb{R}$. A system of this type is called *autonomous*. It is useful to think about v as a vector field on $\mathbb{R}^m$, or on a domain $\Omega \subset \mathbb{R}^m$. In other words, if coordinates in $\mathbb{R}^m$ are denoted by $(u_1, \dots, u_m)$ and the coordinate functions of v are $(v_1, \dots v_m)$ then we can think of v as a vector field $v = \sum\limits_1^m v_i(u)\frac{\partial}{\partial u_i}$. Then the problem of solving the ODE (1.3.2) can be interpreted as finding a path

$$u : [a, b] \to \mathbb{R}^m \tag{1.3.3}$$

such that its velocity vector $u'(t)$ at each point $t \in [a, b]$ coincides with the vector field v at the point $u(t)$, i.e. with the vector $v(u(t))$. Usually one also impose an *initial condition* on the solution: $u(a) = A \in \mathbb{R}^m$.

The space $\mathbb{R}^m$ on which the vector field v lives is called the *phase space* of the system (1.3.2), and the solutions (1.3.3) are called *phase curves* or *integral curves* of the system (1.3.2). The dimension of the phase space is called the *order* of the system.

If one thinks about the vector field v as a velocity vector field of a motion of some fluid then phase curves are trajectories of the individual particles. In the mechanical context, when we think about the parameter t as the time, it is customary to denote the derivative by the dot, i.e. to write $\dot{u}$ instead of u'.

Let us point out, however, that usually for problems arising from Mechanics the phase space is not the space in which the motion takes place. Indeed, consider, for instance, the so-called, 3-body problem when, three bodies (say, the Sun, the Earth and the Moon) move in the 3-space according to the law of gravity, The motion of this system can be described by Newton equations of the form

$$\ddot{u}_1 = f_1(u_1, u_2, u_3),$$
$$\ddot{u}_2 = f_2(u_1, u_2, u_3),$$
$$\ddot{u}_3 = f_3(u_1, u_2, u_3),$$

where $u_1, u_2, u_3 \in \mathbb{R}^3$ are positions of (the centers of mass) of the bodies. After transforming this into a system of first order equations we get a vector field in $\mathbb{R}^{18}$. *This is the phase space of our system.* Thus a motion of a the 3-body system corresponds to a phase trajectory of the corresponding point in its 18-dimensional phase space.

A *non-autonomous* system (1.3.1) can be viewed as a *time-dependent* vector field $v_t(u) = v(t,u)$. For instance, one encounters this situation when studying a non-steady flow of a fluid. Note that any non-autonomous system of order m can be viewed as an autonomous system of order $m+1$:

$$\begin{aligned}\dot{u} &= v(\tau(t), u(t)),\\ \dot{\tau} &= 1.\end{aligned}$$

The space $\mathbb{R}^{m+1} = \mathbb{R}^m \times \mathbb{R}$ of variables (u,τ) is called the *extended phase space* of the original non-autonomous system (1.3.1). In the extended phase space we can write the system as

$$\dot{\widehat{u}} = \widehat{v}(\widehat{u}(t)), \tag{1.3.4}$$

where $\widehat{u} = (u,\tau) \in \mathbb{R}^{m+1}$,

$$\widehat{v} = \sum v_i(\widehat{u})\frac{\partial}{\partial u_i} + \frac{\partial}{\partial \tau}.$$

1.4 Line (direction) fields and Pfaffian equations

Let us denote by λ the *line field* $\lambda := \mathrm{Span}(\widehat{v})$ generated by the vector field $\widehat{v}$. We note that the vector field $\widehat{v}$ can be uniquely reconstructed from λ, and hence the system (1.3.4) can be equivalently viewed as the line field λ. [1]

More generally, given any line field λ in a domain $U \subset \mathbb{R}^n$ one can consider the problem of its integration as finding integral curves for this line field, i.e. paths $u : [a,b] \to U$ such that $\dot{u}(t) \in \lambda_{u(t)}$ for any $t \in [a,b]$. Note that in this case while the direction of the velocity vector is prescribed at any point, its length is not. Hence, one can reparameterize γ by composing it with a diffeomorphism $\phi : [c,d] \to [a,b]$ and get a different integral path which corresponds to the same integral curve viewed as a submanifold of U.

Note that in our original example of the line field λ generated by the vector field when the line field λ has a non-singular projection to one of the coordinates lines (namely, τ). Hence,

[1] In Arnold's book is used the term *direction field* for the line field λ.

any integral curve is graphical with respect to this projection, and therefore we can choose τ as the parameter on them. In fact any line field, in a neighborhood of each point projects non-singularly to one of the coordinate axes, and hence the corresponding coordinate can be chosen as a parameter for integral curves near that point.

Consider now the case when $n = 2$, i.e. when λ is a line field on a domain $U \subset \mathbb{R}^2$. Then, if the line field λ is co-orientable it can be defined by a Pfaffian equation

$$\alpha = 0$$

for a 1-form $\alpha = Pdx + Qdy$ on U.

A solution of this equation, or which is the same, an integral curve of the line field $\lambda = \{\alpha = 0\}$. Hence, if it is given parametrically by $x = x(t), y = y(t)$, $t \in [a, b]$, then we get

$$(P(x(t), y(t))\dot{x}(t) + Q(x(t), y(t))\dot{y}(t))\, dt = 0$$

or

$$P(x(t), y(t))\dot{x}(t) + Q(x(t), y(t))\dot{y}(t) = 0.$$

Near a point where $(x_0, y_0) \in U$ where $Q(x_0, y_0) \neq 0$ (i.e. near a point where the projection of the line field λ to the x-axis is non-singular, we can equivalently write the equation $Pdx + Qdy = 0$ as $dy = -\frac{P}{Q}dx$, and hence look for solutions $y = f(x)$ of the equation

$$f'(x) = -\frac{P(x, f(x))}{Q(x, f(x)},$$

and similarly if $P(x_0, y_0) \neq 0$ we can write the equation in the form $dx = -\frac{Q}{P}dy$ and look for solutions $x = g(y)$ of the equation

$$g'(y) = -\frac{Q(g(y), y)}{P(g(y), y)}.$$

Example 1.1. Vector field on the line. Consider a vector field $v(x) = f(x)\frac{\partial}{\partial x}$ on $\mathbb{R}$ where $f(x) \neq 0$ for all $x \in \mathbb{R}$. Consider the corresponding differential equation

$$\dot{x} = v(x).$$

Passing to the extended phase space $\mathbb{R}^2$ with coordinates (x, t) this equivalent to a Pafaffian equation

$$dx = f(x)dt,$$

which in turn can be rewritten as

$$dt = \frac{dx}{f(x)},$$

because by our assumption $f(x) \neq 0$. Suppose we are looking for an integral curve passing through a point (t_0, x_0). Then integrating this equation we get

$$t - t_0 = \int\limits_{x_0}^{x} \frac{dx}{f(x)}.$$

Chapter 2

Phase flow

In this chapter we denote by U, V domains in $\mathbb{R}^n$. However, everything can be generalized to the case when U and V are any two n-dimensional manifolds.

2.1 Action of a diffeomorphism on a vector field

Let $f : U \to V$ be a diffeomorphism. Let us denote by $\mathrm{Vect}(U)$ and $\mathrm{Vect}(V)$ the spaces of vector fields on U and V, respectively.

Given a diffeomorphism $f : U \to V$ one can define the *push-forward map* $f_* : \mathrm{Vect}(U) \to \mathrm{Vect}(V)$ as follows. Let $X \in \mathrm{Vect}(U)$ be a vector field on U. Then we define the vector field $Y = f_*X$ by the formula

$$Y(v) = d_x f(X(u)), \text{ where } u = f^{-1}(v).$$

Let us point out that unlike the pull-back operator f^* on differential forms which defined for any smooth maps and not, necessarily for diffeomorphisms, the push-forward operator f_* on vector fields is defined *only for diffeomorphisms* (why?).

We can similarly define the push-forward operator on *line* fields. If X is a vector field and $\lambda = \mathrm{Span}(X)$ the line field which it generates then $f_*\lambda = \mathrm{Span}(f_*v)$.

Exercise 2.1. 1. Suppose $n = 2$ and a line field λ on U is defined by a Pfaffian equation $\alpha = 0$, where α is a 1-form on U. Show that given a diffeomorphism $f : U \to V$ the line field $f_*\lambda$ on V can be defined by a Pfaffian equation $\beta = 0$, where

$$\beta := \left(f^{-1}\right)^* \alpha = (f^*)^{-1} \alpha.$$

2. Let $P : U \to V$ be the map introducing polar coordinates. In other words, $U = \{0 < r < \infty,\ 0 < \phi < 2\pi\}$ is a domain in $\mathbb{R}^2$ with Cartesian coordinates (r, ϕ), $V = \mathbb{R}^2 \setminus \{y = 0, x \geq 0\}$ in $\mathbb{R}^2$ with Cartesian coordinates (x, y) and P is defined by the formula

$$P(r, \phi) = (r \cos\phi, r \sin\phi).$$

Let $X = a\frac{\partial}{\partial r} + b\frac{\partial}{\partial \phi}$ be a vector field on U. Find $Y := P_*X = A\frac{\partial}{\partial x} + B\frac{p}{\partial y}$. This can also be equivalently formulated as relating the expressions of a given vector field Y on $\mathbb{R}^2$ in two different bases, the basis $\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$ and $\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \phi}\right)$.

2.2 Isotopy and diffeotopy

Let us denote by $\Delta \subset \mathbb{R}$ an interval in $\mathbb{R}$. This interval can be closed, open, semi-open, and even concides with the whole $\mathbb{R}$ or the rays (a, ∞) or $(-\infty, a)$.

Let us recall that a homotopy $f_t : U \to V$, $t \in \Delta$, is just a continuous family of continuous maps $U \to V$, which depends continuously on the parameter Δ. Equivalently, one can think of a homotopy as a continuous map $F : U \times \Delta \to V$. The relation to the first definition is given by the formula

$$F(x, t) = f_t(x), \quad \text{for } \ x \in U, \ \ t \in \Delta.$$

In this course we will always assume *all homotopies to be smooth*, i.e. $F : U \times \Delta \to V$ is at least a C^1-smooth map.

We will also need two special cases of a homotopy, called an *isotopy* and a *diffeotopy.*

A homotopy $f_t : U \to V$, $t \in \Delta$, is called a *diffeotopy* if $f_t : U \to V$ is a diffeomorphism for each $t \in U$. A homotopy $f_t : U \to V$, $t \in \Delta$, is called an *isotopy* if for each $t \in U$ the map

$f_t : U \to V$ is an *embedding*, i.e. a diffeomorphism onto its image $f_t(U)$. Thus, an embedding need not to be *onto*, and the image $f_t(U)$ can move during an isotopy. Of course, a diffeotopy is a special case of an isotopy.

Let $f_t : U \to U$ (note that the source and the target are the same!) be a diffeotopy. Then we can define a family of vector fields X_t on U by the formula

$$X_t(x) = \frac{df_t}{dt}(f_t^{-1}(x)), \quad x \in U, t \in \Delta. \tag{2.2.1}$$

Equivalently, one can write

$$X_t(f_t(x)) = \frac{df_t}{dt}(x), \quad x \in U, t \in \Delta,$$

which means that *for every $x_0 \in U$ the path $t \mapsto f_t(x_0)$, $t \in \Delta$, is a solution of the equation*

$$\dot{x} = X_t(x) \tag{2.2.2}$$

For any $t_0 \in \Delta$ this solution satisfies the initial condition $x(t_0) = f_t(x_0)$.

2.3 Rectification theorems

Theorem 2.2. *Let X be a C^1-smooth vector field in a domain $\Omega \subset \mathbb{R}^n$. Then for any point $x_0 \in \Omega$ there exists $\epsilon > 0$ and a neighborhood $U \ni x_0, U \subset \Omega$, such that there exists an isotopy $f_t : U \to \Omega$, $t \in (-\epsilon, \epsilon)$ such that*

$$f_0(x) = x \quad \textit{for all } x \in U; \tag{2.3.1}$$

$$\frac{df_t(x)}{dt} = X(f_t(x)). \tag{2.3.2}$$

We will prove this theorem later on in Section **??**.

The isotopy f_t is called the *local phase flow* of the vector field X. If f_t defined globally, i.e. it is a diffeotopy $U \to U$, even defined for small interval of time $(-\epsilon, \epsilon)$ then it is automatically defined for all $t \in \mathbb{R}$, see the next section.

Theorem 2.2 have several corollaries, most of which are essentially equivalent to the theorem itself.

First, we note that by the standard trick of reducing the non-autonomous case to an autonomous one in a space of a bigger dimension, Theorem 2.2 implies its own generalization:

Theorem 2.3. *Let X_t, $t \in \Delta$ be a C^1-smooth family of vector fields in a domain $\Omega \subset \mathbb{R}^n$. Then for any points $x_0 \in \Omega$ and $t_0 \in \Delta$ there exists $\epsilon > 0$ and a neighborhood $U \ni x_0, U \subset \Omega$, such that there exists an isotopy $f_t : U \to \Omega$, $t \in (t_0 - \epsilon, t_0 + \epsilon)$ which satisfies*

- $f_{t_0}(x) = x$ *for all* $x \in U$*;*
- $\frac{df_t(x)}{dt} = X_t(f_t(x)), x \in U, t \in (t_0 - \epsilon, t_0 + \epsilon)$

The next theorem shows that two *non-vanishing* smooth vector fields are *locally diffeomorphic*. More precisely,

Theorem 2.4. *Let X be a C^1-smooth vector field in a domain $\Omega \subset \mathbb{R}^n$. Suppose that that $X(a) \neq 0$ for some point $a \in \Omega$. Then there exists a local coordinate system $(y_1, \ldots, y_n)$ on a neighborhood $U \ni a$, $U \subset \Omega$, centered at the point a such that the vector field X on U is equal to $\frac{\partial}{\partial y_1}$.*

In particular,

Theorem 2.5. *Let λ be a C^1-smooth line field in a domain $\Omega \subset \mathbb{R}^n$. Then for any point $a \in \Omega$ there exists a neighborhood $U \ni a, U \subset \Omega$ and a local coordinate system $(y_1, \ldots, y_n)$ on U, centered at the point a such that the line field Y on U is spanned by the vector field $\frac{\partial}{\partial y_1}$.*

Proof of Theorem 2.4. We can assume without loss of generality that a is the origin of the Cartesian coordinate system, and the vector $X(a)$ coincides with the vector $\frac{\partial}{\partial x_1}$ at the point a. This could be achieved by rotating and scaling the original Cartesian system of coordinates. Let

$$D_\delta^{n-1} := \{x_1 = 0; \sum_2^n x_j^2 \leq \delta^2\}.$$

Suppose that ϵ is chosen so small that $D_\delta^{n-1} \subset U$, where U is the neighborhood provided by Theorem 2.2. Let $f_t : U \to \Omega, t \in (-\epsilon, \epsilon)$ be the local phase flow constructed in Theorem 2.2. Denote

$$H := \{|x_1| \le \epsilon, \sum_2^n x_j^2 \le \delta^2\}$$

and define a map $F : H \to \Omega$ given by the formula $F(x_1, x_2, \dots, x_n) = f_{x_1}(0, x_2, \dots, x_n))$.

The map F is an embedding, provided that ϵ, δ are small enough. Indeed, the differential of F at the origin is the identity map (why?), and hence by the implicit function theorem it is an embedding in a sufficiently small neighborhood of 0. But $F_*(\frac{\partial}{\partial x_1}) = X$, and hence, assuming that ϵ, δ are small enough, the coordinate system introduced on the neighborhood $U' = F(H)$ by the diffeomorphism $F^{-1} : U' \to H$ is the required one. ■

This theorem, in particular implies existence of the solution of a system $\dot{x} = X(x)$ for any initial data $x(t_0) = x_0$ on an interval $(t - \epsilon, t + \epsilon)$, provided that the vector field X is C^1-smooth. It also implies the uniqueness of solution with given initial data and its smooth dependence on the initial data.

2.4 Phase flow

Let X be a smooth vector field in a domain $\Omega \subset \mathbb{R}^n$. Choose $a \in \Omega$. Recall that according to Theorem 2.3 there exists a neighborhood $U \ni a$ in Ω and $\epsilon > 0$ such that there exists a local phase flow for the equation

$$\dot{x} = X(x),\ x \in \Omega, \tag{2.4.1}$$

i.e. an isotopy $f_t : U \to \Omega$, $t \in (-\epsilon, \epsilon)$, such that

- $f_0(x) = x$ for all $x \in U$;
- $\frac{df_t(x)}{dt} = X_t(f_t(x)), x \in U, t \in (-\epsilon, \epsilon)$.

Let us observe that that the interval $(-\epsilon, \epsilon)$ depends on the choice of an initial point $a \in \Omega$ and its neighborhood U. However, if the flow is defined on the whole Ω, i.e. it is a *diffeotopy* $f_t : \Omega \to \Omega$ then the flow is defined for all $t \in \mathbb{R}$.

Indeed, let $E = \sup \epsilon$ such that the flow is defined on $(-\epsilon, \epsilon)$. Suppose that $E < \infty$. Then the flow is defined on $(-E+\delta, E+\delta)$ for $\delta < \frac{\epsilon_0}{2}$ but then we can define it on $(-E', E')$, where $E' = E - \delta + \frac{3\epsilon_0}{4} > E$ by the formula $f_t := f_{\frac{3\epsilon_0}{4}} \circ f_{t-\frac{3\epsilon_0}{4}}$ for $t \in (E-\delta, E')$. This contradiction shows that $E = \infty$, i.e. the flow is defined for all $t \in \mathbb{R}$. The following lemma follows from the definition of the flow.

Lemma 2.6. *Suppose the flow $f_t : \Omega \to \Omega$ for a vector field X is defined for all $t \in \mathbb{R}$. Then*

1. $f_t \circ f_u = f_{t+u}$ *for all* $t, u \in \mathbb{R}$*;*
2. $f_0 = \mathrm{Id}$*;*
3. $f_{-t} = f_t^{-1}$.

One may express this lemma by saying that the *flow of an autonomous system which is defined for all $t \in \mathbb{R}$ forms a 1-parametric group of diffeomorphisms.*

Often for the flow f_t generated by a vector field X we will use the notation X^t instead of f_t.

Conversely, any 1-parametric group of diffeomorphisms $f_t : \Omega \to \Omega$ corresponds to a vector field X on Ω. Indeed, according to the formula (2.2.1) the isotopy f_t defines a *family* of vector fields $X_t(x) = \frac{df_t}{dt}(f_t^{-1}(x))$, $x \in \Omega, t \in \mathbb{R}$. But in this case, denoting $y = f_t^{-1}(x)$

$$X_t(x) = \frac{df_t}{dt}(y) = \lim_{u \to 0} \frac{f_{t+u}(y) - f_t(y)}{u} = \lim_{u \to 0} \frac{f_u(x) - f_t(x)}{u} = X_0(x),$$

i.e. X_t is independent of t.

Proposition 2.7. *Suppose that a vector field X on Ω integrates to a flow $X^t : \Omega \to \Omega$, $t \in \mathbb{R}$, and $f : \Omega \to \widetilde{\Omega}$ a diffeomorphism. Denote $\widetilde{X} := f_* X$. Then the vector field $\widetilde{X}$ integrates to a flow $\widetilde{X}^t$, $t \in \mathbb{R}$, on $\widetilde{\Omega}$ and*

$$\widetilde{X}^t = f \circ X^t \circ f^{-1}, \ t \in \mathbb{R}.$$

Proof. For any point $y = f(x) \in \widetilde{\Omega}$ we have

$$\begin{aligned}\frac{d}{dt}(\widetilde{X}^t(y))\big|_{t=0} &= \frac{d}{dt}(f \circ X^t \circ f^{-1}(y))\big|_{t=0}\\ &= \frac{d}{dt}(f \circ X^t(x))\big|_{t=0} = d_x f(\frac{d}{dt}(X^t(x)\big|_{t=0}) = d_x f(X(x))\\ &= f_* X(y) = \widetilde{X}(y).\end{aligned}$$

■

2.5 Symmetries

Let λ be a line field in $\Omega \subset \mathbb{R}^n$. A diffeomorphism $f : \Omega \to \Omega$ is called a *symmetry* of the line field λ if $f_*\lambda = \lambda$.

Lemma 2.8. *All symmetries of the line field λ form a group.*

Indeed, Id is a symmetry, if f, g are symmetries then $f \circ g$ is a symmetry and if f is a symmetry then f^{-1} is a symmetry.

Consider a differential equation

$$\dot{x} = X_t(x),\ x \in \Omega,\ t \in \Delta. \tag{2.5.1}$$

with the phase space $\Omega \subset \mathbb{R}^n$. Let λ be the corresponding line field on its extended phase space $\Omega \times \Delta$. Then any symmetry $f : \Omega \times \Delta \to \Omega \times \Delta$ of the line field λ is called the *symmetry of the equation* (2.5.1).

Let us stress the point that a symmetry is a diffeomorphism of an *extended* phase space, i.e. it acts on space-time domain, even in the case of an autonomous system. Of course, in the case of an autonomous system $\dot{x} = X(x)$, $x \in \Omega$, one can consider also more restricted class of symmetries, namely diffeomorphisms $h : \Omega \to \Omega$ preserving the vector field X, i.e. $h_*X = X$, as for instance, in the following

Proposition 2.9. *Consider an autonomous system $\dot{x} = X(x)$ on $\Omega \subset \mathbb{R}^n$. Suppose that it integrates to a phase flow $X^t : \Omega \to \Omega$. Then for each $s \in \mathbb{R}$ the diffeomorphism X^s is a symmetry of the equation.*

Proof. Let us compute $Y := X^s_*(X)$. By definition of the phase flow,

$$X(x) = \frac{d}{dt} X^t(x)\big|_{t=0}.$$

On the other hand, by the chain rule for any path $\gamma : (-\epsilon, \epsilon) \to \Omega$ such that $\gamma(0) = x$ and $\gamma'(0) = X(x)$ we have $\frac{d}{dt} f(\gamma(t))\big|_{t=0} = df_x(X(x)) = f_* X(f(x))$. Denote $f := X^s$. Then

$$f_* X(f(x)) = \frac{d}{dt} f \circ X^t(x)\big|_{t=0} = \frac{d}{dt} X^{s+t}(x)\big|_{t=0} = X(X^s(x)).$$

In other words, $f_* X(f(x)) = X(f(X))$, i.e. $f_* X = X$. ■

Theorem 2.10. *Let Y and λ be a vector field and a line field in Ω.*

- *Y integrates to a flow $Y^s : \Omega \to \Omega$;*
- *Y admits a transverse hypersurface Σ such that $\bigcup_{s \in \mathbb{R}} Y^s(\Sigma) = \Omega$ and either*
  - *(a) $Y^s(\Sigma) \neq Y^{s'}(\Sigma)$ for $s \neq s'$, or*
  - *(b) the flow Y^s is defined for all $s \in \mathbb{R}$ and either $Y^s(\Sigma) \cap Y^{s'}(\Sigma) = \varnothing$, or $Y^s(\Sigma) = Y^{s'}(\Sigma)$ (in the latter case the flow is periodic for $s, s' \in \mathbb{R}$).*

Suppose that Y^s is a symmetry of λ for all $s \in \mathbb{R}$. Then the order of the differential equation corresponding to λ can be reduced by 1. In particular, if $\dim \Omega = 2$ then the Pfaffian equation corresponding to λ can be reduced to an equation with separable variables, and hence solved in quadratures.

Proof. We consider below only the case $n = 2$. The proof in the general case follows a similar scheme. In this case Σ is a 1-dimensional manifold, and hence it is diffeomorphic either to $\mathbb{R}$ or to S^1. We will concentrate below on the case of $\mathbb{R}$. Consider a parameterization $\phi : \mathbb{R} \to \Sigma$. Define a map $\Phi : \mathbb{R}^2 \to \Omega$ by the formula

$$\Phi(u, v) = Y^v(\phi(u)).$$

We can think about (u, v) as curvilinear coordinates in Ω. The flow Y^s in these coordinates look like translation along the v-direction:

$$(u, v) \mapsto (u, v + s).$$

The line field λ in these coordinates can be defined by a 1-form $\alpha = P(u, v)du + Q(u, v)dv$. Let us assume that $P \neq 0$. In fact, at every point (u, v) either $P(u, v) \neq 0$ or $Q(u, v) \neq 0$. The case when $Q \neq 0$ can be considered similarly. Then we can define the line field λ by a Pfaffian equation $du + R(u, v)dv = 0$, where $R = \frac{Q}{P}$.

The fact that the line field λ is preserved by the flow Y^s means that

$$(Y^s)^*(du + R(u, v)dv) = f_s(u, v)(du + R(u, v)dv).$$

But $(Y^s)^*(du+R(u, v)dv) = du+R(u, v+s)dv$. Hence, $f_s(u, v) \equiv 1$ and $R(u, v+s) = R(u, v)$, i.e. the function R is independent of V, so we will just write $R(u)$.

Thus in coordinates (u, v) the equation takes the form

$$du + R(u)dv = 0$$

which is an equation with separable variables. ■

Let us notice that if we change the variables (u, v) to (u, V) where $v = h(V)$ then the variables will separate anyway. Indeed, the form $du + R(u)dv$ in coordinates (u, V) takes the form $du + R(u)h'(V)dV$. And thus the variables in the equation $du + R(u)h'(V)dV = 0$ separate as well.

Hence, it is not so important that the coordinate v along trajectories of Y coincides with the time-parameter, but what is crucial is that v is constant on translates of Σ under the flow Y^s.

2.6 Quasi-homogeneous equations

Consider in $\mathbb{R}^n$ the vector field

$$Y = \sum \alpha_i x_i \frac{\partial}{\partial x_i},$$

where $\alpha_1, \dots, \alpha_n$. It is called an *Euler field* with weights $\alpha_1, \dots, \alpha_n$, or just an Euler field, if all weights are equal to 1.

The vector field Y integrates to a 1-parametric group of linear transformations $Y^s : \mathbb{R}^n \to \mathbb{R}^n$ given by the formula

$$Y^s(x_1, \dots, x_n) = (e^{\alpha_1 s}x_1, \dots, e^{\alpha_n s}x_n).$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *quasi-homogeneous of degree d with weights* $\alpha_1, \dots, \alpha_n$ if $f(Y^s(x)) = e^{ds}F(x)$ for all $x \in \mathbb{R}^n, s \in \mathbb{R}$.

A line field λ in a domain Ω is called *quasi-homogeneous with weights* $\alpha_1, \dots, \alpha_n$) if $Y^s_*\lambda = \lambda$ for all s, i.e. transformations Y^s are symmetries of λ.[1]

A differential equation is called *quasi-homogeneous* if the corresponding line field in the extended phase space is quasi-homogeneous.

When all the weights are equal to 1 then the one uses the term *homogeneous* instead of quasi-homogeneous.

Exercise 2.11. 1. Consider a system of equations $\dot{x} = f(x),\ x \in \mathbb{R}^n$. Suppose that the coordinate functions f_i are quasi-homogeneous of degrees d_i with the same weights $\alpha_1, \dots, \alpha_n$. The corresponding line field λ in the extended phase space (x, t) is given by the system of Pfaffian equations

$$dx_1 = f_1(x_1, \dots, x_n)dt;$$

$$\dots$$

$$dx_n = f_n(x_1, \dots, x_n)dt.$$

Suppose $d_1 - \alpha_1 = \dots = d_n - \alpha_n$. Prove that the line field λ is quasi-homogeneous and find the weights. Let Y^s be the quasi-homogeneous flow $Y^s(x_1, \dots, x_n) = (e^{\alpha_1 s}x_1, \dots, e^{\alpha_n s}x_n)$. Compute the push-forward by Y^s of the vector field $X = \sum_1^n f_i \frac{\partial}{\partial x_i}$.

[1] Note that the above definition implies, among other things that domain Ω itself is *invariant* with respect to Y^s, i.e. $Y^s(\Omega) = \Omega$ for all $s \in \mathbb{R}$.

2. Consider equation of k-th order with respect to 1 unknown function:

$$\frac{d^k y}{dx^k} = f(x, y).$$

Suppose that $f(x, y)$ is a quasi-homogeneous function of degree d with weights α, β. Find a relation between α, β and d which ensures that the line field representing the system in its extended $(k + 1)$-dimensional phase space is quasi-homogeneous (and find weights).

2.7 Digression: Differential forms

2.7.1 Multilinear functions

A function $l(X_1, X_2, \dots, X_k)$ of k vector arguments $X_1, \dots, X_k \in V$ (i.e. a function $l : \underbrace{V \times \dots \times V}_{k} \to \mathbb{R}$) is called *$k$-linear* (or multilinear) if it is linear with respect to each argument when all other arguments are fixed. We say *bilinear* instead of 2-linear. Multilinear functions are also called *tensors*. Sometimes, one may also say a "k-linear *form*", or simply k-form instead of a "k-linear functions". However, we will reserve the term k-form for a skew-symmetric tensors which will be defined in Section 2.7.2 below.

If one fixes a basis $v_1 \dots v_n$ in the space V then with each bilinear function $f(X, Y)$ one can associate a square $n \times n$ matrix as follows. Set $a_{ij} = f(v_i, v_j)$. Then $A = (a_{ij})_{i,j=1,\dots,n}$ is called *the matrix of the function f in the basis* $v_1, \dots, v_n$. For any 2 vectors

$$X = \sum_1^n x_i v_i,\ Y = \sum_1^n y_j v_j$$

we have

$$f(X, Y) = f\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{i,j=1}^n x_i y_j f(v_i, v_j) = \sum_{i,j=1}^n a_{ij} x_i y_j = X^T A Y\,.$$

Similarly, with a k-linear function $f(X_1, \dots, X_k)$ on V and a basis $v_1, \dots, v_n$ one can associate a "k-dimensional" matrix

$$A = \{a_{i_1 i_2 \dots i_k};\ \ 1 \le i_1, \dots, i_k \le n\},$$

where

$$a_{i_1 i_2 \dots i_k} = f(v_{i_1}, \dots, v_{i_k}) \, .$$

If $X_i = \sum_{j=1}^{n} x_{ij} v_j, \;\; i = 1, \dots, k$, then

$$f(X_1, \dots, X_k) = \sum_{i_1, i_2, \dots i_k = 1}^{n} a_{i_1 i_2 \dots i_k} x_{1 i_1} x_{2 i_2} \dots x_{k i_k} \, .$$

2.7.2 Symmetric and skew-symmetric tensors

A multilinear function (tensor) is called *symmetric* if it remains unchanged under the transposition of any two of its arguments:

$$f(X_1, \dots, X_i, \dots, X_j, \dots, X_k) = f(X_1, \dots, X_j, \dots, X_i, \dots, X_k)$$

Equivalently, one can say that a k-tensor f is symmetric if

$$f(X_{i_1}, \dots, X_{i_k}) = f(X_1, \dots, X_k)$$

for any permutation $i_1, \dots, i_k$ of indices $1, \dots, k$.

Exercise 2.12. *Show that a bilinear function $f(X,Y)$ is symmetric if and only if its matrix (in any basis) is symmetric.*

.

A tensor is called *skew-symmetric (or anti-symmetric)* if it changes its sign when one transposes any two of its arguments:

$$f(X_1, \dots, X_i, \dots, X_j, \dots, X_k) = -f(X_1, \dots, X_j, \dots, X_i, \dots, X_k).$$

The matrix A of a bilinear skew-symmetric function is skew-symmetric, i.e.

$$A^T = -A.$$

Any linear function is (trivially) symmetric and skew-symmetric.

Example 2.13. *The determinant* $\det(X_1, \dots, X_n)$ *(considered as a function of columns* $X_1, \dots, X_n$ *of a matrix) is a skew-symmetric* n*-linear function.*

Exercise 2.14. *Prove that any* n*-linear skew-symmetric function on* $\mathbb{R}^n$ *is proportional to the determinant.*

The space of skew-symmetric k-linear functions on a vector space V is denoted by $\Lambda^k(V^*)$. Note that $\Lambda^1(V^*) = V^*$. Note that if $k > \dim V$ then any skew-symmetric k-linear function on V is ideentically equal to 0.

2.7.3 Exterior product

Given k linear functions $l_1, \dots, l_k$ on V we define its exterior product $l_1 \wedge \dots \wedge l_k$ as a k-linear skew-symmetric function whose value on vectors $A_1, \dots, A_k \in V$ is given by the formula

$$l_1 \wedge \dots \wedge l_k(A_1, \dots, A_k) = \begin{vmatrix} l_1(A_1 & l_1(A_2) & \dots & l_1(A_k) \\ l_2(A_1 & l_2(A_2) & \dots & l_2(A_k) \\ & & & \\ l_k(A_1 & l_k(A_2) & \dots & l_k(A_k) \end{vmatrix}.$$

Exercise 2.15. *Show that* $l_1 \wedge \dots \wedge l_k = 0$ *if and only if linear functions* $l_1, \dots, l_k$ *are linearly dependent.*

In particular one can take an exterior product $x_{i_1} \wedge x_{i_2} \wedge \dots \wedge x_{i_k}$ of any k out of n, $k \leq n$, coordinate functions $x_{i_1}, \dots, x_{i_2}, \dots x_{i_k}$, $i_1 < \dots < i_k$. Its value on vectors $A_1, \dots, A_k$ is the determinant of the $k \times k$-matrix formed by coordinates of vector $A_1, \dots, A_k$ with numbers $i_1, \dots, i_k$.

Exercise 2.16. *Show that* k*-forms* $x_{i_1} \wedge x_{i_2} \wedge \dots \wedge x_{i_k}$, $i_1 < \dots < i_k$ *form a basis. In particular* $\dim \Lambda^k(V^*) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$.

We extend the definition of the exterior product by linearity to any forms. Given a k-form $\alpha = \sum\limits_{i_1 < \dots < i_k} a_{i_1 \dots i_k} x_{i_1} \wedge x_{i_2} \wedge \dots \wedge x_{i_k}$ and a l-form $\beta = \sum\limits_{j_1 < \dots < j_l} a_{i_1 \dots i_k} x_{j_l} \wedge x_{j_2} \wedge \dots \wedge x_{j_l}$ we define $\alpha \wedge \beta$ as a $(k+l)$-form

$$\alpha \wedge \beta = \sum_{i_1 < \cdots < i_k, j_1 < \cdots < j_l} a_{i_1 \dots i_k} a_{i_1 \dots i_k} x_{i_1} \wedge x_{i_2} \wedge \cdots \wedge x_{i_k} \wedge x_{j_l} \wedge x_{j_2} \wedge \cdots \wedge x_{j_l}.$$

Of course, in this sum all terms with repeated coordinates are 0, and terms which differ by a permutation differ by an appropriate sign .

Exercise 2.17. *Prove the the exterior product has the following properies:*

- *(associativity)* $\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma$.
- *(skew-commutativity) for a k-form α and an l-form β we have* $\alpha \wedge \beta = (-1)^{kl} \beta \wedge \alpha$.

2.7.4 Differential forms

Generalizing the notion of a differential 1-form we define a *differential k-form* on a domain U in a vector space V as a field of k-linear skew-symmetric functions α_x on V_x, $x \in U$.

For instance in $\mathbb{R}^n$ with coordinates $x_1, \dots, x_n$ we can consider differential n-form $dx_1 \wedge \cdots \wedge dx_n$. Its value on any n vectors $A_1, \dots, A_n \mathbb{R}^n_x$ is the determinant of the matrix formed by these vectors as columns. This determinant is an *oriented volume* of the parallelepiped spanned by these vectors. The difference between the n-linear form $x_1 \wedge \cdots \wedge x_n$ and the differential n-form $dx_1 \wedge \cdots \wedge dx_n$, that the former one can be only applied to the vectors originated at the origin, while the latter one can be applied to vectors originated at any point $x \in \mathbb{R}^n$. But the result would be the same as to first parallel transport the vectors to the origin, and then apply $x_1 \wedge \cdots \wedge x_n$.

Any differential k-form α on a domain $U \subset \mathbb{R}^n$ can be written as $\alpha = \sum\limits_{i_1 < \cdots < i_k} a_{i_1 \dots i_k} dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}$, where the coefficients $a_{i_1 \dots i_k}$ are functions on U.

Note that a 0-form is just a function.

2.7.5 Pull-back of a differential form

Suppose we are given two domains $U \subset V$ and $U' \subset V'$ and a diffeomorphism $f : U \to U'$. Then for any differential k-form α on U' we define the differential form $f^*\alpha$ on U, called *pull-back* of α) by the formula

$$(f^*\alpha)_x(X_1, \dots, X_k) = \alpha_{f(x)}(d_x f(X_1), \dots, d_x f(X_k).$$

Here $X_1, \dots X_k \in T_x V$ are vectors originated at $x \in U \subset V$ and $d_f : T_x V \to T_{f(x)} V$ is the differential of the map f and x.

If one think about the diffeomorphism f as a change of coordinatesm theen the pull-back operator just rewrite the form in new coordinates.

An important property of the pull-back operator is that it preserves the exterior product:

Proposition 2.18. $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$.

Proposition 2.18 together with the chain rule implies

Proposition 2.19. *Let $f : U \to U'$ be a diffeomorphism, $x_1, \dots x_n$ are coordinates in V' and $u_1, \dots, u_m$ are coordinates in U, so that the map f is given in these coordinates as*

$$f(u_1, \dots, u_m) = (f_1(u_1, \dots, u_m), \dots, f_n(u_1, \dots, u_m)).$$

Let $\alpha = \sum\limits_{i_1 < \dots < i_k} a_{i_1 \dots i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}$ be a differential form on U'. Then

$$f^*\alpha = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} \circ f df_{i_1} \wedge df_{i_2} \wedge \dots \wedge df_{i_k},$$

Indeed,

$$f^*\alpha = f^*(\sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k})$$

$$\sum_{i_1 < \dots < i_k} f^* a_{i_1 \dots i_k} (f^* dx_{i_1}) \wedge f^*(dx_{i_2}) \wedge \dots \wedge f^*(dx_{i_k}) = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k} \circ f df_{i_1} \wedge df_{i_2} \wedge \dots \wedge df_{i_k}.$$

In other words, to change coordinates in a differential form one just need to replace each coordinate by its expression through new coordinate. This proposition makes changing coordinates in a differential form a simple exercise.

2.7.6 Exterior differential

Given a k-form $\alpha = \sum\limits_{i_1<\dots<i_k} a_{i_1\dots i_k} dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}$ we define its *exterior differential* $d\alpha$ as a $(k+1)$-form

$$d\alpha = \sum_{i_1<\dots<i_k} da_{i_1\dots i_k} \wedge dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}.$$

For instance for a 0-form, i.e. a function f, the exterior differential is just the usual differential: $df = \sum\limits_1^n \frac{\partial f}{\partial x_i} dx_i$. For a differential 1-form $\alpha = \sum\limits_1^n f_i dx_i$ we have

$$d\alpha = \sum_{1\le i<j\le n} \left(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} \right) dx_i \wedge dx_j.$$

For a differential n-form $\alpha = \sum\limits_1^n f_i dx_1 \wedge \dots (\check{dx_i}) \dots \wedge dx_n$ (dx_i is missing) we have

$$d\alpha = \left(\sum_1^n (-1)^{i-1} \frac{\partial f_i}{\partial x_i} \right) dx_1 \wedge \dots \wedge dx_n.$$

For any n form α on an n-dimensional space we have $d\alpha = 0$.

Proposition 2.20. *Let α be a differential k-form, and β a differential l-form*

1. *$d^2 = 0$, i.e. for any differential form α we have $d(d\alpha) = 0$.*

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

3. *Let $f : U \to U'$ be a diffeomorphism, α a differential k-form on U'. Then $df^*\alpha = f^* d\alpha$.*

A differential k-form α is called *closed* if $d\alpha = 0$, and it is called *exact* if there exists a differential $(k-1)$-form β such that $\alpha = d\beta$.

Any exact form is closed, as it follows from Proposition 2.20.3.

Locally *any closed form is exact*, but globally this is not true, and depends on the topology of the domain U. For instance, in $\mathbb{R}^n$ any closed form is exact but on $\mathbb{R}^n \setminus 0$ the $(n-1)$-form

$$\theta = \sum_1^n \frac{(-1)^{i-1} x_i dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n}{r^n},$$

where $r = \sqrt{\sum\limits_1^n x_i^2}$, is closed but not exact.

2.8 Directional derivative revisited

Let X be a smooth vector field defined on a domain $U \subset \mathbb{R}^n$ (more generally we can assume that U is any n-dimensional manifold). Given a function $f : U \to \mathbb{R}$ we can define the *directional derivative* $L_X f$ of f along X:

$$L_X f = \lim_{s \to 0} \frac{f(x + tX) - f(x)}{t}. \tag{2.8.1}$$

The directional derivative has many other notation: $D_X(f), \frac{\partial f}{\partial X}, df(X), \dots$.

Let us denote by $X^t : U' \to U$, $t \in (-\epsilon, \epsilon)$, the local phase flow of X^t defined on a neighborhood $U' \subset U$ of a point $a \in U$.

Let us observe that the directional derivative can be also defined by the formula

$$L_X f(a) = \frac{d}{ds} f \circ X^s \Big|_{s=0} (a). \tag{2.8.2}$$

It turns out that formula (2.8.2) can be generalized to define an analog of directional derivatives for differential forms and vector fields, which is the *Lie derivative*.

2.9 Lie derivative of a differential form

Let ω be a differential k-form. We define the Lie derivative $L_X\omega$ of ω along a vector field X as

$$L_X\, \omega = \frac{d}{ds} (X^s)^* \omega \Big|_{s=0}. \tag{2.9.1}$$

Note that if ω is a 0-form, i.e. a function f, then $(X^s)^* f = f \circ X^s$, and hence, in this case definitions (2.8.2) and (2.9.1) coincide, i.e. for functions the Lie derivative is the same as the directional derivative.

Proposition 2.21. *The following identities hold*

1. $L_X(\omega_1 \wedge \omega_2) = (L_X\omega_1) \wedge \omega_2 + \omega_1 \wedge L_X\omega_2$.

2. $L_X(d\omega) = d(L_X\omega)$.

Proof.

1. $L_X(\omega_1 \wedge \omega_2) = \frac{d}{ds}(X^s)^*(\omega_1 \wedge \omega_2)\Big|_{s=0} = \frac{d}{ds}\big((X^s)^*\omega_1 \wedge (X^s)^*\omega_2\big)\Big|_{s=0}$
$= \frac{d}{ds}\big((X^s)^*\omega_1\big)\Big|_{s=0} \wedge \omega_2 + \omega_1 \wedge \frac{d}{ds}\big((X^s)^*\omega_2\big)\Big|_{s=0} = (L_X\omega_1) \wedge \omega_2 + \omega_1 \wedge L_X\omega_2.$

2. $L_X(d\omega) = \frac{d}{ds}\big((X^s)^*d\omega\big)\Big|_{s=0} = \frac{d}{ds}\big(d(X^s)^*\omega\big)\Big|_{s=0} = d\Big(\frac{d}{ds}(X^s)^*\omega\Big|_{s=0}\Big) = L_X(d\omega).$

■

The following formula of Élie Cartan provides an effective way for computing the Lie derivative of a differential form.

Theorem 2.22. *Let X be a vector field and ω a differential k-form. Then*

$$L_X\omega = d(X \lrcorner\, \omega) + X \lrcorner\, d\omega. \tag{2.9.2}$$

Proof. Suppose first that $\omega = f$ is a 0-form. Then $L_X f = df(X) = X \lrcorner df$, which is equivalent to formula (2.9.2), because in this case the first term in the formula is equal to 0. Then, using Proposition 2.212) we get

$$L_X df = dL_X f = d(df(X)) = d(X \lrcorner\, df),$$

which is again equivalent to (2.9.2) because in this case $ddf = 0$. Next we note that if the formula (2.8.1) holds for ω_1 and ω_2 then it holds also for $\omega_1 \wedge \omega_2$. Indeed, we have

$$(\star)\quad L_X(\omega_1 \wedge \omega_2) = (L_X\omega_1) \wedge \omega_2 + \omega_1 \wedge L_X\omega_2$$
$$= (X \lrcorner\, d\omega_1 + d(X \lrcorner\, \omega_1)) \wedge \omega_2 + \omega_1 \wedge (X \lrcorner\, d\omega_2 + d(X \lrcorner\, \omega_2))$$
$$= (X \lrcorner\, d\omega_1) \wedge \omega_2 + \omega_1 \wedge (X \lrcorner\, d\omega_2) + d(X \lrcorner\, \omega_1) \wedge \omega_2 + \omega_1 \wedge d(X \lrcorner\, \omega_2)$$

On the other hand, denoting by d_1 and d_2 the degrees of ω_1 and ω_2, we get

$$(\star\star)\quad X \lrcorner\, d(\omega_1 \wedge \omega_2) + d(X \lrcorner\, (\omega_1 \wedge \omega_2))$$
$$= X \lrcorner\, (d\omega_1 \wedge \omega_2 + (-1)^{d_1}\omega_1 \wedge d\omega_2) + d((X \lrcorner\, \omega_1) \wedge \omega_2 + (-1)^{d_1}\omega_1 \wedge (X \lrcorner\, \omega_2))$$
$$= (X \lrcorner\, d\omega_1) \wedge \omega_2 + (-1)^{d_1+1} d\omega_1 \wedge (X \lrcorner\, \omega_2) + (-1)^{d_1}(X \lrcorner\, \omega_1) \wedge d\omega_2 + \omega_1 \wedge (X \lrcorner\, d\omega_2)$$
$$+ d(X \lrcorner\, \omega_1) \wedge \omega_2 + (-1)^{d_1+1} X \lrcorner\, \omega_1 \wedge d\omega_2 + (-1)^{d_1} d\omega_1 \wedge (X \lrcorner\, \omega_2) + \omega_1 \wedge (d(X \lrcorner\, \omega_2))$$
$$= (X \lrcorner\, d\omega_1) \wedge \omega_2 + \omega_1 \wedge (X \lrcorner\, d\omega_2) + d(X \lrcorner\, \omega_1) \wedge \omega_2 + \omega_1 \wedge d(X \lrcorner\, \omega_2)\,.$$

Comparing the computation in $(\star)$ and $(\star\star)$ we conclude that

$$L_X(\omega_1 \wedge \omega_2) = X \lrcorner\, d(\omega_1 \wedge \omega_2) + d(X \lrcorner\, (\omega_1 \wedge \omega_2)).$$

By induction we can prove a similar formulas for an exterior product of any number of forms.

Finally we observe that any differential k-form ω can be written in coordinates as

$$\sum_{1 \le i_1 < \cdots < i_k \le n} f_{i_1 \ldots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k},$$

i.e. ω is a sum of exterior products of functions (0-forms) and exact 1-forms, and hence Cartan's formula follows. ■

Proposition 2.23. *We have*

$$L_X\omega = 0 \iff (X^s)^*\,\omega = \omega \ \textit{ for all } \ s \in \mathbb{R}.$$

Proof. If $(X^s)^*\,\omega \equiv \omega$ then $L_X\omega = \frac{d}{ds}(X^s)^*\omega\Big|_{s=0} = 0$. To prove the converse we note that

$$\frac{d}{ds}(X^s)^*\omega\Big|_{s=s_0} = \lim_{t\to 0}\frac{(X^{s_0+t})^*\omega - (X^{s_0})^*\omega}{t} = (X^{s_0})^*\left(\lim_{t\to 0}\frac{(X^t)^*\omega - \omega}{t}\right)$$
$$= (X^{s_0})^*\,(L_X\omega)\,,$$

and hence if $L_X\omega = 0$ then $(X^s)^*\,\omega = \omega$. ■

2.10 Lie bracket of vector fields

Let $A, B \in \mathrm{Vect}(U)$ be two vector fields on a domain $U \subset \mathbb{R}^n$. As it was shown in 52H, there is a vector field $C \in \mathrm{Vect}(V)$, called the *Lie bracket* of the vector fields A and B and denoted by $C = [A, B]$, which is characterized by the following property: *for any smooth function $\phi : U \to \mathbb{R}$ one has*

$$L_C\phi = (L_A L_B - L_B L_A)\phi.$$

A surprising fact here is that though the right-hand side of this equation seems to be the second order differential operator, the left-hand side is the first order operator, so the second derivatives on the right side cancel each other.

Recall that the bracket $[A, B]$ has the following properties

- Lie bracket is a bilinear operation;
- $[A, B] = -[B, A]$ (skew-symmetricity);
- $[[A, B]C] + [[B, C], A] + [[C, A], B] = 0$ (Jacobi identity);
- If $A = \sum\limits_1^n a_j \frac{\partial}{\partial x_j}$ and $B = \sum\limits_1^n b_j \frac{\partial}{\partial x_j}$ then

$$[A, B] = \sum_{i=1}^n \left(\sum_{j=1}^n a_j \frac{\partial b_i}{\partial x_j} - b_j \frac{\partial a_i}{\partial x_j} \right) \frac{\partial}{\partial x_i}\,. \qquad (2.10.1)$$

In this section we will give a new interpretation of the Lie bracket $[A, B]$.

Recall that given a diffeomorphism $f : U \to V$ we can define the push-forward map

$$f_* : \mathrm{Vect}(U) \to \mathrm{Vect}(V).$$

We can also define the *pull back* map

$$f^* : \mathrm{Vect}(V) \to \mathrm{Vect}(U)$$

by the formula $f^* := f_*^{-1}$. Note that we also have $f^* = f_*^{-1}$.

We define the *Lie derivative* $L_A B$ of the vector field B along the vector field A in a similar way as we defined in Section 2.9 the Lie derivative of a differential form. Namely,

$$L_A B = \frac{d\left(A^s\right)^* B}{ds}\Big|_{s=0}. \tag{2.10.2}$$

More explicitly,

$$L_A B(x) = \lim_{s \to 0} \frac{d_{A^s(x)}\left(A^{-s}\right)\left(B(A^s(x)) - B(x)\right.}{s}.$$

Similarly, to Proposition 2.23 we have

Proposition 2.24.

$$L_A B = 0 \iff \left(A^s\right)^* B \equiv B \;\; \textit{for all} \;\; s \in \mathbb{R}.$$

Proof. We have

$$\begin{aligned} \frac{d\left(A^s\right)^* B}{ds}\Big|_{s=s_0} &= \lim_{s \to 0} \frac{\left(A^{s+s_0}\right)^* B - \left(A^{s_0}\right)^* B}{s} \\ &= \lim_{s \to 0} \left(A^{s_0}\right)^* \left(\frac{\left(A^s\right)^* B - B}{s}\right) = \left(A^{s_0}\right)^* \left(\lim_{s \to 0} \frac{\left(A^s\right)^* B - B}{s}\right) \\ &= \left(A^{s_0}\right)^* \left(L_A B\right). \end{aligned}$$

Hence, if $L_A B = 0$ then $\frac{d(A^s)^* B}{ds}$ for all s and hence $\left(A^s\right)^* B = \left(A^0\right)^* B = B$. The converse is obvious. ■

Theorem 2.25. *For any two vector fields* $A, B \in \mathrm{Vect}(U)$

$$L_A B = [A, B].$$

Proof. Note that $A^s(x) = x + sA(x) + o(s)$. Hence, we can write

$$d_y A^{-s} = \mathrm{Id} - s d_y A + o(s),$$

where we view here A as a map $\mathbb{R}^n \to \mathbb{R}^n$. Furthermore, plugging $y = A^s(x)$ we get

$$d_{A^s(x)} A^{-s} = \mathrm{Id} - s d_x A + o(s).$$

Indeed, $d_{A^s(x)}A - d_xA \underset{s\to 0}{\to} 0$ and hence $s(d_yA - d_xA) = o(s)$. We also have

$$B(A^s(x)) = B(x + sA(x) + o(x)) = B(x) + sd_xB(A(x)) + o(s).$$

Thus, ignoring $o(s)$-terms we get

$$\begin{aligned} L_AB &= \lim_{s\to 0} \frac{1}{s}\left(d_{A^s(x)}\left(A^{-s}\right)(B(A^s(x))) - B(x)\right) \\ &= \lim_{s\to 0} \frac{1}{s}\left((\mathrm{Id} - sd_xA)\right)(B(x) + sd_xB(A(x))) - B(x)) \\ &= \lim_{s\to 0} \frac{1}{s}\left(B(x) - sd_xA(B) + sd_xB(A) - B(x)\right) = d_xB(A) - d_xA(B). \end{aligned}$$

But the right-hand-side expression written in coordinates has the form

$$d_xB(A) - d_xA(B) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_j\frac{\partial b_i}{\partial x_j} - b_j\frac{\partial a_i}{\partial x_j}\right)\frac{\partial}{\partial x_i}$$

which coincides with the expression (2.10.1) for the Lie bracket. ∎

Exercise 2.26. Prove that for any smooth function ϕ one has

$$L_{[A,B]}\phi = \frac{\partial^2(\phi\circ A^s\circ B^t)}{\partial s\partial t}.$$

If $[A,B]=0$ then one says that the vector field A and B *commute*.

Lemma 2.27. *Suppose two commuting vector fields A, B on Ω can be integrated into phase flows A^t, B^s. Then*

$$A^t\circ B^s = B^s\circ A^t,$$

$t, s\in\mathbb{R}$, i.e. the flows of commuting vector fields. Conversely, if two flows A^t, B^s commute for all $t, s\in\mathbb{R}$ then $[A,B]=0$.

Proof. We have $[A,B] = L_AB$. Then according to Proposition 2.24 we have

$$(A^s)^*B = B. \tag{2.10.3}$$

Recall from Proposition 2.7 that for any diffeomorphism $f : \Omega \to \Omega$ if $f^*B = C$ then

$$C^t = f^{-1} \circ B^t \circ f, \quad t \in \mathbb{R}.$$

Applying this to $f = A^s$ and using (2.10.3) we conclude

$$B^t = A^{-s} \circ B^t \circ A^s,$$

or

$$A^s \circ B^t = B^t \circ A^s, \quad s, t \in \mathbb{R}.$$

■

2.11 First integrals

Suppose we are given a differential equation

$$\dot{x} = A(x), \tag{2.11.1}$$

where A is a vector field on the domain $U \subset \mathbb{R}^n$ A function $\phi : U \to \mathbb{R}$ is called a *first integral*, or simply an *integral* of equation (2.11.1) if it is constant on solutions of this equation, or equivalently on integral curves of the vector field A.

Clearly, a necessary and sufficient condition for ϕ to be an integral is to satisfy the equation $L_A\phi = 0$. Here $L_A\phi$ denotes the directional derivative of ϕ along A.

If ϕ is an integral of (2.10.2) then the solutions are contained in the level sets of the function ϕ, and hence, this allows us to reduce the order of equation by 1. If (2.10.2) has two integrals ϕ_1, ϕ_2, then the solutions lie in the intersection of level sets $\{\phi_1 = c_1\}$ and $\{\phi_2 = c_2\}$, $c_1, c_2 \in \mathbb{R}$. Hence, if these level sets transverse to each other (which means that the differential $d\phi_1$ and $d\phi_2$ are linearly independent at every point of the intersection), then the solutions lie in $\{\phi_1 = c_1\} \cap \{\phi_2 = c_2\}$, which allows to further reduce the order of the system. If the order is reduced to 1 then the equation can be explicitly integrated in quadratures. Such systems are called *completely intregrable.*

Some important examples of integrals which come from Mechanics are discussed in the next section.

2.12 Hamiltonian vector fields

Consider the vector space $\mathbb{R}^{2n}$ with coordinates $(p_1, \ldots, p_n, q_1, \ldots, q_n)$ and a closed differential 2-form $\omega = \sum\limits_1^n dp_i \wedge dq_i$. Note that this form is non-degenerate, i.e. its matrix is non-degenerate at every point. Therefore, the map $J : \mathrm{Vect}(\mathbb{R}^{2n}) \to \Omega^1(\mathbb{R}^{2n})$ given by the formula $X \mapsto X \lrcorner \omega$ is an isomorphism between the space $\mathrm{Vect}(\mathbb{R}^{2n})$ of vector fields and the space $\Omega^1(\mathbb{R}^{2n})$ of differential 1-forms on $\mathbb{R}^n$. In coordinates the map J associates with a vector field $\sum\limits_1^n P_i \frac{\partial}{\partial P_i} + \sum\limits_1^n Q_i \frac{\partial}{\partial Q_i}$ the differential form $\sum\limits_1^n P_i dq_i - Q_i dp_i$.

Lemma 2.28. *Given a vector field A on $\mathbb{R}^{2n}$ the differential 1-form $J(A) = A \lrcorner \omega$ is closed if and only if $L_A \omega = 0$.*

Proof. Indeed, according to Cartan's formula (2.9.2) we have $L_A \omega = d(A \lrcorner \omega) = dJ(A)$ because ω is closed. ■

Given a function $H : \mathbb{R}^{2n} \to \mathbb{R}$ we denote by X_H the vector field $-J^{-1}(dH)$. Vector fields obtained by this construction are called *Hamiltonian*.

To find a coordinate expression for X_H we write $X_H = \sum\limits_1^n a_i \frac{\partial}{\partial p_i} + b_i \frac{\partial}{\partial q_i}$. Then

$$X_H \lrcorner \omega = \left(\sum_1^n a_i \frac{\partial}{\partial p_i} + b_i \frac{\partial}{\partial q_i} \right) \lrcorner \sum_1^n dp_i \wedge dq_i = \sum_1^n -b_i dp_i + a_i dq_i.$$

Hence, the equation

$$X_H \lrcorner \omega = -dH = -\sum_1^n \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i$$

implies $a_i = -\frac{\partial H}{\partial q_i}, b_i = \frac{\partial H}{\partial p_i}$, $i = 1, \ldots, n$. Thus,

$$X_H = \sum_1^n -\frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i}.$$

In a shorter form, omitting indices we will write

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

Thus the system of differential equations corresponding to the vector field X_H has the form

$$\begin{aligned} \dot{p} &= -\frac{\partial H}{\partial q} \\ \dot{q} &= \frac{\partial H}{\partial p}. \end{aligned} \tag{2.12.1}$$

These equations play an important role in Mechanics, and called *Hamilton canonical equations.* They describe the phase flow of a mechanical system. Here the coordinates $q = (q_1, \ldots, q_n)$ determine a position of the system, or a point in the *configuration space* of the mechanical system. The coordinates $p = (p_1, \ldots, p_n)$ are called *momenta* and can be viewed as vectors of the cotangent bundle to the configuration space. The function H is the full energy of the system expressed through coordinates and momenta.

Lemma 2.29. *The function H is a first integral of the equation* (2.12.1)*, i.e.* $L_{X_H}H = 0$.

Proof.

$$L_{X_H}H = dH(X_H) = -\frac{\partial H}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} = 0.$$

■

Example 2.30. Consider Newton equations

$$\ddot{q}_i = -\frac{\partial U}{\partial q_i}, i = 1, \ldots, n,$$

or in shorter notation

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

Reducing it to a system of first order equation we get

$$\dot{p} = -\frac{\partial U}{\partial q} \tag{2.12.2}$$

$$\dot{q} = p. \tag{2.12.3}$$

Consider the *full energy* $H(p,q) = \sum_{1}^{n} \frac{p_i^2}{2} + U(q) = \frac{1}{2}p^2 + U(q)$. Then $\frac{\partial H}{\partial q} = \frac{\partial U}{\partial q}$ and $\frac{\partial H}{\partial p} = p$, and hence equation (2.12.2) takes the form (2.12.1) with this Hamiltonian function H. Lemma 2.29 is the law of conservation law of energy.

Lemma 2.31. *Let X_H be a Hamiltonian vector field and X_H^s the phase flow it generates. Then $(X_H^s)^* \omega = \omega$ for all $s \in \mathbb{R}$. In other words, the flow of a Hamiltonian vector field preserves the form ω.*

Proof. It is sufficient to prove that $L_{X_H}\omega = 0$. Using Theorem 2.22 we get

$$L_{X_H}\omega = d(X_H \lrcorner\, \omega) + X_H \lrcorner\, d\omega.$$

But ω is closed, and hence $d\omega = 0$, while $X_H \lrcorner\, \omega = dH$. Thus, $L_{X_H}\omega = ddH = 0$. ■

2.13 Canonical transformations

The equations (2.12.1) are called canonical because they are invariant with respect to a large group of transformation of the phase space. Let us call a diffeomorphism $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ a *symplectomorphism* (or alternatively a *canonical transformation*) if it preserves the form ω. Then it preserves also the form of the equations (2.12.1). Indeed, suppose $f(p,q) = (\widetilde{p}, \widetilde{q})$. Then $f^*(\omega) = f^*(dp \wedge dq) = d\widetilde{p} \wedge d\widetilde{q} = \omega = dp \wedge dq$. Thus if we express the function $H(p,q)$ through the coordinates $\widetilde{p}, \widetilde{q}$, $H(p,q) = \widetilde{H}(\widetilde{p}, \widetilde{q})$ then the equation (2.12.1) will take the same form in coordinates $(\widetilde{p}, \widetilde{q})$:

$$\begin{aligned} \dot{\widetilde{p}} &= -\frac{\partial \widetilde{H}}{\partial \widetilde{q}} \\ \dot{\widetilde{q}} &= \frac{\partial \widetilde{H}}{\partial \widetilde{p}}. \end{aligned} \tag{2.13.1}$$

The following proposition provides an important class of canonical transformations,

Proposition 2.32. *Consider any diffeomorphism $f : U \to V$ between two domains $U, V \subset \mathbb{R}^n$. Let Df be the Jacobi matrix of the map U. Then the map*

$$(p,q) \mapsto \left((Df)^{-1}\right)^T p, f(q))$$

is a symplectomorphism $\widehat{f}$ of the domain $\widehat{U} = \{p \in \mathbb{R}^n, q \in U\}$ to the domain $\widehat{V} = \{p \in \mathbb{R}^n, q \in V\}$. Here $\left((Df)^{-1}\right)^T$ is the matrix transpose to inverse of the Jacobi matrix Df.

In other words, *any change of q-coordinates extends to a canonical change of the (p,q)-coordinates.*

Proof. Let us denote the elements of the matrix $(Df)^{-1}$ by $g_{ij},\ i,j = 1,\dots,n$. Thus, $\sum\limits_i^n g_{ji}\frac{\partial f_i}{\partial q_k} = \delta_{jk}$, $\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ if $j \neq k$.

Let us compute $\widehat{f}^*(pdq) = \widehat{f}^*\left(\sum\limits_1^n p_i dq_i\right)$. We have

$$\widehat{f}(p_1,\dots,p_n,q_1,\dots,q_n) = \left(\sum_1^n g_{j1}p_j,\dots,\sum_1^n g_{jn}p_j, f_1(q),\dots,f_n(q)\right).$$

Hence,

$$\begin{aligned}\widehat{f}^*(pdq) = \widehat{f}^*\left(\sum_1^n p_i dq_i\right) &= \sum_{i=1}^n\sum_{j=1}^n g_{ji}p_j df_i \\ &= \sum_{i,j,k=1}^n g_{ji}\frac{\partial f_i}{\partial q_k}p_j dq_k = \sum_{j,k=1}^n \delta_{jk}p_j dq_k \\ &= \sum_1^n p_k dq_k = pdq.\end{aligned}$$

Hence,

$$\widehat{f}^*\omega = \widehat{f}^* dp \wedge dq = d(\widehat{f}^*(pdq)) = d(pdq) = dp \wedge dq = \omega.$$

■

Corollary 2.33. . *Suppose that there exists a change of coordinates $\widetilde{q} = f(q)$ such that in new coordinates the Hamiltonian function H is independent of the coordinate $\widetilde{q}_1$. Then $\widetilde{p}_1 = \sum\limits_1^n g_{j1}p_j$ is a first integral of the system* (2.12.1). *Here the notation g_{ij} stands for the elements of the matrix $(Df)^{-1}$.*

Proof. Let us extend the coordinate change $q \mapsto \widetilde{q} = f(q)$ to a canonical change of coordinates $(p,q) \mapsto (\widetilde{p},\widetilde{q}) = \widetilde{f}(p,q)$ as in Proposition 2.32. Then the equation in the new coordinates $(\widetilde{p}, \widetilde{q})$ also has the canonical Hamiltonian form (2.13.1). Then $\dot{\widetilde{p}}_1 = \frac{\partial H}{\partial \widetilde{q}_1} = 0$ because by assumption the Hamiltonian is independent of the coordinate $\widetilde{q}_1$. Hence $\widetilde{p}_1 = \sum_1^n g_{j1} p_j$ is constant along trajectories, i.e. it is a first integral. ■

2.14 Example: angular momentum

Consider a Newton equation

$$\ddot{q} = -\nabla U(q), \quad q \in \mathbb{R}^3, \tag{2.14.1}$$

which describes the motion of a particle of mass 1 in a field with a potential energy function $U(q)$. Suppose there exists an axis l in $\mathbb{R}^3$ such that the function $U(q)$ remains invariant with respect to rotations around l.

The system (2.14.1) can be rewritten in the Hamiltonian form (2.12.1) with the Hamiltonian function $H = \frac{p^2}{2} + U(q) = \frac{p_1^2}{2} + \frac{p_2^2}{2} + \frac{p_3^2}{2} + U(q_1, q_2, q_3)$. Let us assume for simplicity that the q_3-axis coincides with the axis l.

Let us change coordinates (q_1, q_2, q_3) to cylindrical coordinates (ϕ, r, z):

$$q_1 = r\cos\phi, q_2 = r\sin\phi, q_3 = z.$$

Equivalently,

$$\phi = \arctan\frac{q_2}{q_1}, r = \sqrt{q_1^2 + q_2^2}, z = q_3.$$

Computing the Jacobi matrix $\frac{D(\phi, r, z)}{D(q_1, q_2, q_3)}$ we get

$$\begin{pmatrix} \frac{\partial\phi}{\partial q_1} & \frac{\partial\phi}{\partial q_2} & \frac{\partial\phi}{\partial q_3} \\ \frac{\partial r}{\partial q_1} & \frac{\partial r}{\partial q_2} & \frac{\partial r}{\partial q_3} \\ \frac{\partial z}{\partial q_1} & \frac{\partial z}{\partial q_2} & \frac{\partial z}{\partial q_3} \end{pmatrix} = \begin{pmatrix} -\frac{q_2}{q_1^2+q_2^2} & \frac{q_1}{q_1^2+q_2^2} & 0 \\ \frac{q_1}{\sqrt{q_1^2+q_2^2}} & \frac{q_2}{\sqrt{q_1^2+q_2^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then the inverse matrix is equal to

$$\begin{pmatrix} -q_2 & \frac{q_1}{\sqrt{q_1^2+q_2^2}} & 0 \\ q_1 & \frac{q_2}{\sqrt{q_1^2+q_2^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Let us extend the coordinate change $(q_1, q_2, q_3) \mapsto (r, \phi, z)$ to a canonical coordinate change

$$(q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (\phi, r, z, p_\phi, p_r, p_z),$$

where we denoted by p_r, p_ϕ, p_z momenta variables corresponding to new coordinates (r, ϕ, z). In fact, we need only the coordinate p_ϕ which is given by $p_\phi = -p_1 q_2 + q_1 p_2$. Thus, the function $-p_1 q_2 + p_2 q_1$ is the first integral. It is called *the angular momentum* around the q_3-axis.

Recall that along trajectories we have $p_i = \dot{q}_i$, $i = 1, 2, 3$. Hence, $q_1 \dot{q}_2 - \dot{q}_1 q_2$ is constant along the trajectories. But this is exactly the projection M_3 of the cross-product $M = q \times \dot{q}$ to the q_3-axis which is the axis of rotational symmetry. Introducing cylindrical coordinates $(r, \phi, z)$with the axis q_3 as z, then we get $M_3 = r^2 \dot{\phi}$.

In particular, if $U(q)$ is invariant under *all* rotations, i.e. it depends only on the distance $r = ||q||$ from the origin, then all components of the angular momentum vector $M = q \times \dot{q}$, and hence, the angular momentum vector M is constant along trajectories. Note that $q\dot{M} = 0$, and hence the motion happens in the plane orthogonal to the vector M. In the cylindrical coordinates with M at its axis, the absolute value of the angular momentum,

$$||M|| = r^2 \dot{\phi}$$

is preserved.

Chapter 3

Solving one first order partial differential equation

3.1 Jet spaces

When studying functions on $\mathbb{R}^n$, or a domain in $\mathbb{R}^n$ it is useful to consider their graphs which live in $\mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$, i.e. for $u : \mathbb{R}^n \to \mathbb{R}$ its graph

$$\Gamma_u := \{z = u(x_1, \dots, x_n)\} \subset \mathbb{R}^{n+1}.$$

Similarly, when studying first order partial differential equations with respect to a function on $\mathbb{R}^n$ it is useful to consider a *simultaneous graph of a function and all its derivatives*:

$$\Lambda_u = \{z = u(x), p_1 = \frac{\partial u}{\partial x_1}(x), \dots p_n = \frac{\partial u}{\partial x_n}(x),\ x = (x_1, \dots, x_n) \in \mathbb{R}^n\} \subset \mathbb{R}^{2n+1},$$

where we denoted coordinates in $\mathbb{R}^{2n+1} = \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ by (x, p, z), $x, p \in \mathbb{R}^n, z \in \mathbb{R}$. The coordinate z is reserved for graphing the value of a function u and $p_1, \dots, p_n$ for the corresponding first partial derivatives.

The space R^{2n+1} in this context is called the 1*-jet space of functions on* $\mathbb{R}^n$ and usually denoted by $J^1(\mathbb{R}^n)$. We denote by π the projection $J^1(\mathbb{R}^n) \to \mathbb{R}^n \to \mathbb{R}^n$ given by the formula

$$\pi(x, p, z) = x,\ (x, y, z) \in J^1(\mathbb{R}^n) = \mathbb{R}^n \times \mathbb{R}^n \times R.$$

A map $s : \mathbb{R}^n \to J^1(\mathbb{R}^n)$ is called a *section* if $\pi \circ s = \mathrm{Id} : \mathbb{R}^n \times \mathbb{R}^n$. In other words, if $s(x) = (x, v(x), u(x)) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ for $x \in \mathbb{R}^n$. With every function $u : \mathbb{R}^n \to \mathbb{R}$ one can associate a very special section. Namely,

$$x \mapsto \left(x, \frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_n}(x), u(x)\right),\ x \in \mathbb{R}^n,$$

which maps $\mathbb{R}^n$ onto the simultaneous graph of the function u and all its first partial derivatives. Sections of this type are called *holonomic*. We note that most of the sections are not holonomic..

The following lemma gives a necessary and sufficient condition for a section $s : \mathbb{R}^n \to \mathbb{R}^{2n+1}$ to be holonomic. Denote by λ the differential 1-form

$$\lambda := dz - \sum_1^n p_i dx_i,$$

and by ξ the hyperplne field defined by the Pfaffian equation $\lambda = 0$. This hyperplane field is called a *contact structure*.

Lemma 3.1. *A section $s : \mathbb{R}^n \to J^1(\mathbb{R}^n)$ is holonomic if and only if $s^*\lambda = 0$. In other words, s is holonomic if its image is tangent to the contact structure ξ.*

Proof. We have $s(x) = (x, p = v(x), z = u(x))$, and hence the equation

$$0 = s^*\lambda = s^*(dz - pdx) = du - vdx$$

is equivalent to

$$v_1(x) = \frac{\partial u}{\partial x_1}(x), \ldots, v_n(x) = \frac{\partial u}{\partial x_n}(x)$$

which is the definition of a holonomic section.

Submanifolds of dimension n which are tangent to ξ are called *Legendrian*. We note that a general Legendrian submanifold need not be necessarily graphical.

Exercise 3.2. Give an example of a non-graphical Legendrian submanifold $\Lambda \subset J^1(\mathbb{R}^n)$.

3.2 The case $n = 1$

When $n = 1$ then the 1-jet space is 3-dimensional, $J^1(\mathbb{R}) = \mathbb{R}^3$. A holonomic section $s : \mathbb{R} \to J^1(\mathbb{R})$ is a simultaneous graph of a function and its derivative:

$$s(x) = (x, p = f'(x), z = f(x)).$$

The contact structure ξ is the 2-dimensional plane field given by a Pfaffian equation $dz - pdx = 0$.

Let $\Sigma \subset J^1(\mathbb{R})$ be a 2-dimensional submanifold. Suppose that for $a \in \Sigma$ the tangent plane $T_a\Sigma$ is transverse to the contact plane ξ_a. Then the line $\ell_a = T_a\Sigma \cap \xi_a$ is called the *characteristic line.* If Σ is transverse to ξ everywhere, then $\ell = \{\ell_a\}_{a\in\Sigma}$ is a tangent line field to Σ (which is called the *characteristic line field). The integral curves of this line field are called* characteristics.

Lemma 3.3. *Characteristics are Legendrian submanifolds. In particular, if a characteristic $\Lambda \subset \mathbb{R}$ is graphiical with respect to the projection $J^1(\mathbb{R}) \to \mathbb{R}$ then it is a holonomic, i.e. there exists a function $h : (a, b) \to \mathbb{R}$ such that $s(x) = (c, h'(x), h(x)), x \in (a, b)$.*

3.3 Characteristics in the n-dimensional case

Let $\Sigma \subset J^1(\mathbb{R}^n)$ be a hypersurface. A point $a \in \Sigma$ is called *singular* if $T_a\Sigma = \xi_a$. Otherwise, i.e. if $T_a\Sigma$ is transverse to ξ_a, it is called *regular.* At a regular point $a \in \Sigma$ the intersection $\Pi_a = T_a\Sigma \cap \xi_a$ is an $(2n-1)$-dimensional subspace. Here are some conditions which guarantees transversality of $\Sigma \subset J^1(\mathbb{R}^n)$ and $\xi = \{\lambda = 0\}$, i.e. regularity of all points of Σ.

Example 3.4.

1. Suppose a $\Sigma = \{F = 0\}$ where for every point $a \in \Sigma$ there exists $i = 1, \dots, n$ such that $\frac{\partial F}{\partial p_i}(a) \neq 0$. Then Σ is transverse to ξ.

2. Suppose the hypersurface Σ is tangent to the z-directions (e.g. the defining it function F is independent of z. Then Σ is transverse to ξ.

Lemma 3.5. *Suppose Σ is transverse to ξ. Then for any point $a \in \Sigma$ there exists a unique line $\ell_a \subset \Pi_a = \xi_a \cap T_a$ which is characterized by the following condition. Given any vectors $v \in \ell_a$ and $w \in \Pi_a$ we have*

$$d\lambda(v, w) = 0.$$

In other words, ℓ_a is the kernel of the form $d\lambda|_{\Pi_a}$.

Proof. The contact hyperplane field $\xi - \{dz - pdx = 0\}$ is transverse to the z-axis, and hence the form $d\lambda = dp \wedge dx|_\xi$ has the maximal rank $2n$. Therefore the restriction of this form to the codimension 1 subspace $\Pi_a \subset \xi_a$ has rank $2n - 1$, because the rank cannot drop more than by 1, but on the other hand the rank of a skew-symmetric form is always even. Hence, there exists a 1-dimensional kernel $\ell_a \subset \Pi_a$ of the form $d\lambda|_{\Pi_a}$, i.e. $d\lambda(v, w) = 0$ for any vectors $v \in \ell_a, w \in \Pi_a$.

The line field $\ell = \{\ell_a\}_{a \in \Sigma}$ which is tangent to Σ is called the *characteristic line field*, and its integral curves are called *characteristics*.

The next lemma gives an explicit expression for a vector field directing the line field ℓ.

Lemma 3.6. *Suppose $\Sigma = \{F(x, p, z) = 0\}$ and $a = (x, p, z) \in \Sigma$ a regular point. Then the line ℓ_a is generated by the vector*

$$v = \sum_1^n F_{p_i} \frac{\partial}{\partial x_i} - \sum_1^n (F_{x_i} + p_i F_z) \frac{\partial}{\partial p_i} + \sum_1^n p_i F_{p_i} \frac{\partial}{\partial z}. \tag{3.3.1}$$

Proof. Given any vector $w = (X, Y, Z) \in \Pi_a = \xi_a \cap T_a\Sigma$ its coordinates should satisfies the following conditions. The equation $dF_a(w) = 0$ takes the form

$$F_x X + F_p P + F_z Z = 0. \tag{3.3.2}$$

The equation $\lambda(w) = 0$ takes the form

$$Z - pX = 0. \tag{3.3.3}$$

Hence, vectors in ξ_a have the form (X, P, pX), and the necessary and sufficient condition for a vector w to be $\xi_a \cap T_a\Sigma$ is that it satisfies the equation

$$(F_x + pF_z)X + F_pP = 0.$$

Let $v = (\widetilde{X}, \widetilde{Y}, \widetilde{Z})$ be a non-zero vector given by (3.3.1). We have $\widetilde{Z} = p\widetilde{X} = \sum_1^n p_x \widetilde{X}_i$ and $(F_x + pF_z)\widetilde{X} + F_p\widetilde{P} = (F_x + pF_z)F_p - F_p(F_x + pF_z) = 0$, and hence $v \in \Pi_a$. We also have

$$v \lrcorner d\lambda = v \lrcorner dp \wedge dx = \widetilde{P}dx - \widetilde{X}dp,$$

and for any vector $w = (X, P, pX) \in \Pi_a$, we have

$$P\widetilde{X} - \widetilde{P}X = (F_x + pF_z)X + F_pP = 0. \tag{3.3.4}$$

Lemma 3.7. *Let $\Sigma \subset J^1(\mathbb{R}^n)$ be a hypersurface transverse to ξ, and ℓ the characteristic line field. Let $L \subset \Sigma$ be a submanifold such that $\lambda|_L = 0$ and L is transverse to ℓ. Let $\widehat{L}$ denote the union of all trajectories of the characteristic foliation intersecting L. Then $\lambda|_{\widehat{L}} = 0$.*

In other words, if we flow a k-dimensional submanifold of Σ tangent to ξ along the characteristics, then it swaps a $(k+1)$-dimensional submanifold of Σ tangent to ξ.

Proof. Choose a non-vanishing vector field $v \in \ell$. At a point $a \in L$ the tangent $T_a\widehat{L} \subset \Pi_a$ is spanned by T_aL and the vector $v(a)$. Note that $d\lambda_{T_a\widehat{L}} = 0$ because $d\lambda|_{T_aL} = 0$ by assumption, and $d\lambda(v(a), w) = 0$ for all $w \in T_a\widehat{L}$ because $v(a) \in \ell_a = \operatorname{Ker} d\lambda|_{\Pi_a}$. We also note that the flow of the vector field v on $\widehat{L}$ preserves the form $\mu := \lambda|_{\widehat{L}}$. Indeed, the Lie derivative $L_v(\lambda|_{\widehat{L}}) = d(\lambda(v)) + v \lrcorner d\lambda) = 0$. Here the first term vanishes because $v \in \ell \subset \xi$, and the second one vanishes because $v \in \ell = \operatorname{Ker}(d\lambda|_\Pi)$. Therefore, if λ vanishes in one point of a trajectory of v, then it vanishes at every point of this trajectory. But by definition any trajectory of v on $\widehat{L}$ intersects L, and as we had seen above λ vanishes on $\widehat{L}$ at the points of L. Hence, it vanishes, everywhere.

Lemma 3.8. *Let $\Sigma \subset J^1(\mathbb{R}^n)$ be a hypersurface transverse to ξ, and ℓ the characteristic line field. Then any Legendrian submanifold $L \subset \Sigma$ is tangent to ℓ.*

Proof. Recall that a Legendrian submanifold is an n-dimensional submanifold tangent to Σ. Suppose that for a point $a \in \Sigma$ the characteristic line ℓ_a is transverse to T_aL. Consider the $(n+1)$-dimensional space $S := \mathrm{Span}(T_aL, v)$. We have $S \subset \Pi_a \subset \xi_a$. On the other hand, $d\lambda|_S = 0$. Indeed, $d\lambda|_{T_aL} = 0$ by assumption, and $d\lambda(v, w) = 0$ for all $w \in S$ and $v \in \ell_a$ because $\ell_a = \mathrm{Ker}\, d\lambda|_{\Pi_a}$. But $d\lambda$ is a non-degenerate form on a $2n$-dimensional space ξ_a. Hence, it cannot vanish on a subspace of dimension $> n$.

Theorem 3.9. *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^{n-1} = \{x_n = 0\}$ be two bounded open domains such that $\overline{\Omega}_1 \subset \Omega_2$, and $\phi : \overline{\Omega}_2 \to \mathbb{R}$ a smooth function. Consider a Cauchy problem*

$$\begin{aligned} &\frac{\partial u}{\partial x_n} = f\left(x_1, \ldots, x_{n-1}, \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_{n-1}}, u\right) \\ &u(x_1, \ldots, x_{n-1}, 0) = \phi(x_1, \ldots, x_{n-1}). \end{aligned} \tag{3.3.5}$$

with respect to a function $u : \mathbb{R}^n \to \mathbb{R}$. Then for a sufficiently small $\epsilon > 0$ the Cauchy problem (3.3.5) *has a unique solution for $(x_1, \ldots, x_{n-1} \in \Omega_1$, $|x_n| \leq \epsilon$. This solution can be found using the following procedure. Consider a system of ordinary differential equations*

$$\begin{aligned} &\dot{x}_i = -\frac{\partial f}{\partial p_i}(x, p_1, \ldots, p_{n-1}, z), \quad i = 1, \ldots, n-1, \\ &\dot{x}_n = 1, \\ &\dot{p}_i = \frac{\partial f}{\partial x_i}(x, p_1, \ldots, p_{n-1}, z) + p_i \frac{\partial f}{\partial z}(x, p_1, \ldots, p_{n-1}, z), \\ &\dot{z} = f(x, p_1, \ldots, p_{n-1}, z) - \sum_1^{n-1} p_i \frac{\partial f}{\partial p_i}(x, p_1, \ldots, p_{n-1}, z), \end{aligned} \tag{3.3.6}$$

Let

$$x_j = \alpha_j(c_1, \ldots, c_{n-1}, t),\ j = 1, \ldots, n-1,\ x_n = t, \tag{3.3.7}$$

$$p_j = \beta_j(c_1, \ldots, c_{n-1}, t),\ j = 1, \ldots, n-1, \tag{3.3.8}$$

$$z = \gamma(c_1, \ldots, c_{n-1}, t), \tag{3.3.9}$$

be the solution of system (3.3.6) *with initial data*

$$x_j(0) = c_j,\ j = 1, \ldots, n-1,\ (c_1, \ldots, c_{n-1}) \in \Omega_2,$$
$$x_n(0) = 0.,$$
$$p_j(0) = \frac{\partial \phi}{\partial x_j}(c_1, \ldots, c_{n-1}),\ j = 1, \ldots, n-1,$$
$$z(0) = \phi(c_1, \ldots, c_{n-1}).$$

The system of algebraic equations (3.3.7) *can be resolved with respect to* c_i, $i = 1, \ldots, n-1$*:*

$$c_j = \delta_j(x_1, \ldots, x_n),\ j = 1, \ldots, n-1,$$

for sufficiently small values of x_n. *Then the function*

$$u(x_1, \ldots, x_n) := \gamma(\delta_1(x_1, \ldots, x_n), \ldots, \delta_{n-1}(x_1, \ldots, x_n), t)$$

is the solution of the Cauchy problem for (3.3.5).

3.4 Integrable systems

3.4.1 Generating functions

Consider a canonical transformation (symplectomorphism) $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$. We endow the source the space by canonical coordinates (p, q) and the target space by
with the canonical coordinates (P, Q). Suppose that the map f is given by coordinate function

$$P = P(p, q), \tag{3.4.1}$$
$$Q = Q(p, q). \tag{3.4.2}$$

Then $dp \wedge dq = dP \wedge dQ$, or

$$d(pdq + QdP) = 0.$$

If (3.4.1) can be resolved with respect to the variables, p, Q, i.e. if we can express from (3.4.1) p and Q as functions of q and P:

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

then the differential 1-form

$$\lambda : -p(q, P)dq + Q(q, P)dP$$

is closed, and hence *exact*, because all closed forms in the whole Euclidean space are exact. Therefore, there exists a function $S(q, P)$, such that $dS = \lambda$, or

$$p = \frac{\partial S}{\partial q}(q, P), \tag{3.4.3}$$

$$Q = \frac{\partial S}{\partial P}(q, P). \tag{3.4.4}$$

Conversely, any function $S(q, P)$ defines via formulas (3.4.3) a canonical transformation *if* (and this is a very big "IF") equations (3.4.3) can be resolved with respect to the variables P and Q. The function $S(q, P)$ is called a *generating function* for the canonical transformation (3.4.1). Given a transformation its generating function, if exists is defined up an additive constant.

3.4.2 Polarizations

The above construction of a generating function is a special case of a more general phenomenon. We begin with the following Linear Algebra lemma. Consider the standard symplectic space $\mathbb{R}^{2n}$. Note that the n-dimensional coordinate subspaces $L_q = \{p = 0\}$ and $L_p = \{q = 0\}$ are Lagrangian and they intersect at one point, the origin. This is an example of a *polarization*. In general, a polarization is any pair $L_1, L_2 \subset \mathbb{R}^{2n}$ of Lagrangian subspaces of $\mathbb{R}^{2n}$ which are transverse, i.e. intersect only at the origin.

Lemma 3.10. *For any polarization $L_1, L_2 \subset \mathbb{R}^{2n}$ there is a linear canonical transform $F : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that $F(L_1) = L_q$ and $F(L_2) = L_p$. In other words, any two polarizations are equivalent under a linear symplectic change of coordinates.*

Proof. To simplify the notation, we will prove it only for the case $n = 2$. The general case is similar. Let e_1, f_1, e_2, f_2 be the standard symplectic basis of $\mathbb{R}^4$, i.e.

$$\omega(e_1, f_1) = \omega(e_2, f_2) = 1, \omega(e_1, f_2) = \omega(e_2, f_1) = \omega(f_1, f_2) = \omega(e_1, e_2) = 0.$$

We have $L_q = \mathrm{Span}(e_1, e_2), L_p = \mathrm{Span}(f_1, f_2)$.

Take any basis $v_1, v_2 \in L_1$. Consider ω-orthogonal complements $v_1^{\perp_\omega}, v_2^{\perp_\omega}$ of vectors v_1 and v_2. Then $\supset L_1, v_2^{\perp_\omega} \supset L_1$, and $\dim(v_1^{\perp_\omega}) = \dim(v_2^{\perp_\omega}) = 3$ and $v_1^{\perp_\omega} \cap v_2^{\perp_\omega} = L_1$.

There exists a non-zero vector $w_1 \in L_2 \cap v_2^{\perp_\omega}$. Then $w_1 \notin v_1^{\perp_\omega}$ (because $v_1^{\perp_\omega} \cap v_2^{\perp_\omega} = L_1$), and therefore $\omega(v_1, w_1) \neq 0$. By scaling w_1 with a scalar factor we can arrange that $\omega(v_1, w_1)$. Similarly we can find a vector $w_2 \in L_2 \cap v_1^{\perp_\omega}$, such that $\omega(v_2, w_2) = 1$. Summarizing our construction we get

$$\omega(v_1, w_1) = \omega(v_2, w_2) = 1, \omega(v_1, w_2) = \omega(v_2, w_1) = \omega(v_1, v_2) = \omega(w_1, w_2) = 0.$$

Hence the linear transformation which sends the basis v_1, w_1, v_2, w_2 to the basis e_1, f_1, e_2, f_2 preserves the symplectic form ω and maps the polarization L_1, L_2 onto the polarization L_q, L_p.

Let us now revisit the generating function construction for canonical transfornations. Consider a canonical transformation $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ given by the formulas (3.4.1) and take its graph

$$\Gamma_f = \{(p, q, P = P(p, q), Q = Q(p, q))\}.$$

Then Γ_F is Lagrangian for the symplectic form

$$\Omega := dp \wedge dq - dP \wedge dQ = dp \wedge dq + dQ \wedge dP.$$

Note that the Lagrangian coordinate planes $L_{q,P}$ and $L_{p,Q}$ form a polarization of $\mathbb{R}^{4n} = \mathbb{R}^{2n} \times \mathbb{R}^{2n}$. But as it follows from Lemma 3.10 any other pair of transverse Lagrangian planes also form a polarization, and because all polarizations are symplectically equivalent one can associates a generating function with any polarization (L_1, L_2) for which the projection of Γ_f onto L_1 along L_2 is non-degenerate. Hence, if the polarization $L_{q,P}$ and $L_{p,Q}$ does not satisfy this condition, then one might want to search for another polarization for which this condition holds. For instance, one can try the polarization $L_{q,Q}, L_{p,P}$. Unfortunately, for sufficiently complicated transformation the regularity of projection condition does not hold for any polarization. A hint of what can be done in this case is given in Section 3.4.6 below.

3.4.3 First order systems

Consider a first order Hamiltonian system with the Hamiltonian

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

We will assume that the

$$0 \leq U(q) \to |q| \to \infty\infty,$$

that $U(q) = 0$ and that $q = 0$ is the only critical point of the potential U. The Hamiltonian equations are

$$dotp = -U'(q),$$
$$\dot{q} = p.$$

The Hamiltonian is an integral of the system, so the (unparametrized) trajectories are energy levels

$$M_h := \left\{ \frac{p^2}{2} + U(q) \right\}, h \in \mathbb{R}.$$

The integral

$$I(h) := \frac{1}{2\pi} \int\limits_{M_h} pdq.$$

is called the *action* of the trajectory. We have

$$I(h) = \frac{1}{2\pi} \iint\limits_{H \leq h} dp \wedge dq,$$

and hence the action $I(h)$ is up to the factor $\frac{1}{2\pi}$ just the area of the domain enclosed by the trajectory M_h. Note that due to our assumption on U the function $I(h)$ is monotonically increasing and we have $I'(h) > 0$. Hence, we can parameterize the level sets M_h by the parameter I instead of h, so we will write $\widetilde{M_I}$ instead of M_h when $I(h) = I$. We will also denote by $g(I)$ the inverse function to $I(h)$, i.e. $g(I(h)) = h$.

Our goal to find a variable ϕ, valued in $\mathbb{R}/2\pi\mathbb{Z}$ (i.e. defined up to addition of a multiple of 2π like the angular coordinate), and such that

$$(p, q) \mapsto (I, \phi)$$

is a canonical change of coordinates, i.e. $dp \wedge dq = dI \wedge d\phi$. To do that we will try to find the generating function $S(I, q)$ for this canonical transformation, i.e. the function which satisfies

$$p = \frac{\partial S(I,q)}{\partial q},$$
$$\phi = \frac{\partial S(I,q)}{\partial I}.$$

Choose the point $q_0(I) = (0, -\sqrt{g(I)}) \in \widetilde{M_I}$ and denote by $\gamma_{I,q}$ a path which is contained in the level set M_h with $I(h) = I$ and which projects to the interval $[0, q]$ on the q-axis. This path is not unique, but we ignore this for a moment. Define the generating function $S(I, q)$ by the formula

$$S(I,q) = \int\limits_{\gamma(I,q)} p dq = -\int\limits_0^q \sqrt{2g(I) - 2U(q)} dq. \tag{3.4.5}$$

Then we have

$$\frac{\partial S(I,q)}{\partial q} = p$$

and *we define*

$$\phi = \frac{\partial S(I,q)}{\partial I} = \frac{g'(I)}{2} \int\limits_0^q \frac{dq}{\sqrt{2g(I) - 2U(q)}}.$$

We note that the path $S(I, q)$ is defined up to adding a full loop around the level set $\widetilde{M_I}$. But $\int\limits_{\widetilde{M_I}} pdq = 2\pi I$, and therefore the variable ϕ is defined up to addition of a multiple of $\frac{\partial(2\pi I))}{\partial I} = 2\pi$.

Example 3.11. Suppose $U(q) = \frac{q^2}{2}$. then $I(h) = \frac{1}{2\pi} \int\limits_{p^2+q^2=2h} pdq = h$, and hence $g(I) = I$, $\widetilde{M}_I = M_h = \{p^2 + q^2 = 2I\}$ We then have

$$S(I,q) = -\int\limits_0^q \sqrt{2I - q^2} dq,$$

$$\frac{\partial S(I,q)}{\partial I} = -\int\limits_0^q \frac{dq}{\sqrt{2I - q^2}} = -\int\limits_0^{\frac{q}{\sqrt{2I}}} \frac{2du}{\sqrt{1-u^2}} = -\arcsin\left(\frac{q}{\sqrt{2I}}\right).$$

But this integral coincides up to a constant with the angular polar coordinate.

3.4.4 Tori and Lagrangian tori

Before considering the general case we discuss n-dimensional tori. An n-dimensional torus is a submanifold diffeomorphic to the product of n circles: $T = S^1 \times \cdots \times S^1$. Thinking of the circle S^1 as the quotient $\mathbb{R}/2\pi/Z$, i.e. as parameterized by real number? up to multiples of 2π, we can think of the torus as a tuples of n cyclic coordinates $\phi_i \in \mathbb{R}/2\pi/Z$. The decomposition of a torus into a product of n circles is not unique. For instance, given any integer-valued $(n \times n)$-matrix A with $\det A = 1$ the map $x \mapsto Ax$ of $\mathbb{R}^n$ induces a transformation $\widehat{A} : T^n \to T^n$. For instance, when $n = 2$ and $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, the transformation $\widehat{A}$ sends the meridians $\phi_1 = \text{const}$ to latitudes, but send the parallels $\phi_2 = \text{const}\}$ to curves which run once around meridians and once around parallels.

Consider now a *Lagrangian* torus T in the symplectic space $(\mathbb{R}^{2n}, dp \wedge dq)$, $dp \wedge dq|_T = 0$. Equivalently, this means that the differential for $pdq|_T$ is closed.

Example 3.12. The torus $T_{a_1,\ldots,a_n} := \{|z_1| = a_1, |z_2| = a_2, |z_n| = a_n\} \subset \mathbb{C}^n = \mathbb{R}^{2n}$ is Lagrangian.

Suppose we fixed a splitting of a Lagrangian torus $T \subset \mathbb{R}^{2n}$ into the product $S^1 \times \cdots \times S^1$. Let ϕ_i be the corresponding cyclic coordinates. The numbers

$$I_j := \int\limits_{\phi_k = c_k,\ k \neq j} pdq$$

are called *periods* of the Lagrangian torus T. They depend on the choice of the splitting, but independent of any constants c_k and any continuous deformation of the coordinate system.

Example 3.13. Periods of the torus $T_{a_1,\dots,a_n} := \{|z_1| = a_1, |z_2| = a_2, |z_n| = a_n\} \subset \mathbb{C}^n = \mathbb{R}^{2n}$ are equal to $I_j = \pi a_j^2$.

3.4.5 The general case

Suppose now we are given an integrable Hamiltonian system

$$\ddot{p} = -\frac{\partial H}{\partial q};$$
$$\ddot{q} = \frac{\partial H}{\partial p}, \quad p, q \in \mathbb{R}^n.$$

which has n independent integrals $G_1 = H, G_2, \dots, G_n$ in involution:

$$\{G_i, G_j\} = 0, \quad i, j = 1, \dots n.$$

Here independence means the linear independence of the differentials $dG_1, \dots, dG_n$ at every point, or equivalently, that the Jacobi matrix of the map $(G_1, \dots, G_n) : \mathbb{R}^{2n} \to \mathbb{R}^n$ has the rank n.

According to the Liouville-Arnold theorem the *compact* common level sets $T_a = \{G_a = a_1, \dots, G_n = a_n\}$ are n-dimensional tori, and the motion on these tori are quasi-periodic, i.e. given by equations

$$\dot{\phi}_1 = \omega_1,$$
$$\dots,$$
$$\dot{\phi}_n = \omega_n,$$

for some angular coordinates $\phi_1, \dots \phi_n \in \mathbb{R}/2\pi\mathbb{Z}$, defined up to multiples of 2
pi. The constants $\omega_1, \dots, \omega_n$ depends on the tori T_a and vary continuously with a.

The goal of this section is to explain how these angular coordinates could be found. For this we will try to find a canonical change of coordinates

$$(p, q) \mapsto (I, \phi), \quad I = h(a).$$

Choose some $c \in \mathbb{R}^n$. There exists a diffeomorphism F of the product $(B_\epsilon(c) := \{a \in \mathbb{R}^n; ||a - c|| < \epsilon\}) \times (T = S^1 \times \cdots \times S^1)$ onto a neighborhood of the torus T_c, so that $a \times T$ is mapped onto the Lagrangian torus T_a. For each $a \in B_\epsilon(c)$ we define I_j as periods of T_a with respect to the above splitting. One can show that the map $a \mapsto I(a) = (I_1(a), \ldots, I_n(a))$ is 1-1 in $B_\epsilon(c)$, and hence we can invert this map: $a = g(I)$.

We will also choose a base point $z_0 \in T = S^1 \times \cdots \times S^1$ Then $F(a) = (p(a), q(a)) \in T_a$. Let us assume that the projection of T_a onto the coordinate q-space is $1-1$ in a neighborhood of the point $F(a)$. Pick any path $\gamma^a_{q(a),q}$ in this neighborhood connecting the point $F(a) = (p(a), q(a))$ with the point which projects to the point q and define

$$S(q, I) = \int\limits_{\gamma^{g(I)}_{q(g(I)),q}} p dq.$$

Then we have $\frac{\partial S}{\partial q} = p$, and hence if we view S as a generating function, and define

$$\phi := \frac{\partial S}{\partial I},$$

then the transformation $(p, q) \mapsto (I, \phi)$ is canonical. In fact the same formulas define the coordinates ϕ globally, but only up to multiples of 2π because going around one of factors of the torus by definition increases the action function S by $2\pi I_j$, and increases $\frac{\partial S}{\partial I_j}$ by 2π.

The coordinates I, ϕ are called *action-angle coordinates*. In action angles coordinates the Hamiltonian depends only the action coordinates I_j, and hence I_j are integrals, and the Hamiltonian system takes a simple form $\dot{\phi} = \frac{\partial H}{\partial I}$.

Exercise 3.14 (Geodesics on an ellipsoid of revolution)**.** Take an ellipsoid $E := \{x^2 + y^2 + 2z^2 = 1\}$ and consider the problem of a free particle moving on the surface of the ellipsoid.

The Lagrangian of this system is just the kinetic energy

$$L = T = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

It is clearly useful to pass to the cylindrical coordinates:

$$x = r\cos\theta, y = r\sin\theta, z.$$

Then

$$\dot{x} = \cos\theta\dot{r} - r\sin\theta\dot{\theta},$$
$$\dot{y} = \sin\theta\dot{r} + r\cos\theta\dot{\theta}.$$

We also have

$$z = \pm\frac{1}{\sqrt{2}}\sqrt{1 - x^2 - y^2},$$
$$\dot{z} = \mp\frac{1}{\sqrt{2}}\frac{x\dot{x} + y\dot{y}}{\sqrt{1 - x^2 - y^2}} =$$
$$\mp\frac{1}{\sqrt{2}}\frac{r\dot{r}}{\sqrt{1 - r^2}}.$$

Hence,

$$T = \frac{1}{2}\left((\dot{r})^2 + r^2(\dot{\theta})^2 + (\dot{z})^2\right) = \frac{1}{2}\left((\dot{r})^2 + r^2(\dot{\theta})^2 + \frac{r^2(\dot{r})^2}{1 - r^2}\right) = \frac{1}{2}\left(\frac{(\dot{r})^2}{1 - r^2} + r^2(\dot{\theta})^2\right).$$

Making the Legendre transform we get the Hamiltonian

$$H = \frac{1}{2}\left((1 - r^2)p_r^2 + \frac{p_\theta^2}{r^2}\right).$$

The Hamiltonian is independent of θ and hence, p_θ is an integral (it is the angular momentum). The integrals $G_1 = H$ and $G_2 = \theta$ Poisson commute, and hence the system is integrable. The invariant tori are

$$T_a = \{G_1 = a_1^2, G_2 = a_2^2\} = \left\{|p_\theta| = a_2^2, |p_r| = \sqrt{\frac{a_1^2 - \frac{a_2^2}{r^2}}{1 - r^2}}\right\}, \quad a_1, a_2 \geq 0$$

To have these tori non-empty and non-degenerate $a_2^2 < a_1^2$ and in this case we have $r \in (\frac{a_2}{a_1}, 1)$. *Find action-angle coordinates and integrate the system explicitly.*

3.4.6 Fixed points of canonical transformation

In this section we illustrate the power of the method of generating functions for the problem concerning existence of fixed points of symplectomorphisms. The following conjecture formulated by V.I. Arnold in 1960s stimulated the development of a new subject of symplectic topology.

Conjecture 3.15 (Arnold's fixed points conjecture)**.** *Let (M, ω) be a closed symplectic manifold and $f : M \to M$ a Hamiltonian diffeomorphism. Then f has at least as many fixed points, as the minimal number of critical points of a smooth function $\phi : M \to \mathbb{R}$.*

Remark 3.16. In this generality the conjecture is still open. For the case of 2-torus and other surfaces it was first proven by myself in [3]. For the case of an n-dimensional torus it was proven by C. Conley and E. Zehnder in [1]. It was generalized to other manifolds by the work of many people: M. Gromov ([6]), A. Floer ([5]), K. Fukaya and K Ono ([7]) and others. It is now known for all symplectic manifolds, but the lower bound for the manifold is not quite as good as predicted by Conjecture 3.15.

We note that this minimal number is at least 2 because any function on a closed manifold has at least two critical points, the minimum and the maximum. In fact, it is usually larger. For instance, for the 2-dimensional torus this number is 3, if one allows fixed points to be degenerate, and 4 in the non-degenerate case.

3.5 Proof of Arnold's conjecture for the 2-torus

Lemma 3.17. *Given a loop $\gamma : S^1 \to M$ consider a map $F : S^1 \times [0,1] \to M$ given by the formula $\gamma(u,t) = f_t(\gamma(u))$, $u \in S^1, t \in [0,1]$. Then $\int\limits_{S^1 \times [0,1]} F^*\omega = 0$.*

Proof. The tangent space to $S^1 \times [0,1]$ is generated by the vector fields $\frac{\partial}{\partial u}$ and $\frac{\partial}{\partial t}$. We have

$$\frac{\partial}{\partial t} \lrcorner F^*\omega = \frac{\partial F}{\partial t} \lrcorner \omega = X_{H_t} \lrcorner \omega = -dH_t.$$

Hence,

$$\int\limits_{S^1\times[0,1]} F^*\omega = \int\limits_0^1 \left(\frac{\partial}{\partial t} \lrcorner F^*\omega\right) dt = -\int\limits_0^1 \left(\int\limits_{f_t(\gamma)} dH_t\right) dt = 0,$$

because the integral of the exact 1-form dH_t over a closed curve $f_t(\gamma)$ is equal to 0.

We prove below Arnold's fixed point conjecture for the 2-torus, but we will only prove existence of 1 fixed point. A slightly more precise argument allows to prove existence of at least 3 fixed points. The current proof was first given in [4].

What is remarkable about this proof that it could be given by H. Poincaré. In fact, the first half of the proof almost precisely follows the first page of Poincaré's paper [2].

Theorem 3.18 (C. Conley and E. Zehnder,[1]). *Any Hamiltonian diffeomorphism f of the 2-torus $(T^2,\omega) = (\mathbb{R}^2/\mathbb{Z}^2, dp\wedge dq)$ must have at least 1 fixed point.*

Proof. We view the torus T^2 as the quotient $\mathbb{R}^2/\mathbb{Z}^2$, i.e. the set of points $(p,q)\in\mathbb{R}^2$ up to addition of a vector with integer coordinates. Let us denote by π the projection $\mathbb{R}^2\to T^2$. The area form $\Omega = dp\wedge dq$ on $\mathbb{R}^2$ descends to the an area form ω on $\mathbb{R}^2$, i.e. $\pi^*\omega = \Omega$.

The Hamiltonian isotopy $f_t : T^2 \to T^2$ lifts to a Hamiltonian isotopy $F_t : \mathbb{R}^2 \to \mathbb{R}^2$ such that $F_0 = \mathrm{Id}$ and $\phi\circ F_t = f_t$ for all $t\in[0,1]$.

We have $F(p,q) = (P(p,q), Q(p,q))$ and $dP\wedge dQ = dp\wedge dq$. Let us first assume that F is C^1-close to the identity. Then its graph

$$\Gamma_F = \{(p,q,P,Q)|\ P = P(p,q), Q = Q(p,q)\} \subset \mathbb{R}^4$$

is graphical with respect to the splitting of $\mathbb{R}^{4n}$ into the (q,P)- and (p,Q)-coordinate subspaces, i.e.

$$\Gamma_F = \{p = p(q,P), Q = Q(q,P)\},$$

and hence the equation $dp\wedge dq = dP\wedge dQ$ is equivalent to the existence of a function $G(q,P)$ such that $pdq + QdP = dG$. Fixed points $p = P, Q = q$ of F are zeroes of the 1-form $(p-P)dq + (Q-q)dP = d(G-qP)$. In other words, fixed points are exactly the critical points of the function $\widetilde{G}(q,P) := G(q,P) - qP$.

Lemma 3.19. *The function* $\widetilde{G}$ (called a generating function of the canonical transformation F) *is* 1*-periodic in variables* q, P, *i.e.* $\widetilde{G}(q+1,P)=\widetilde{G}(q,P+1)=\widetilde{G}(q,P)$.

Proof. Take a path γ in the coordinate plane (p,q) connecting points $\gamma(0)=(q_0,p_0)$ and $\gamma(1)=(q_0+1,p_0)$. Note that the projection $\widetilde{\gamma}:=\pi\circ\gamma$ of this path to the torus T^2 is a loop. Consider a family of paths $\delta_s:[0,1]\to\mathbb{R}^4=\mathbb{R}^2\times\mathbb{R}^2$, $s\in[0,1]$, defined by the formula $(p,q)=\delta(t),(P,Q)=F_s(\delta(t))$, so that the path δ_s lies on the graph Γ_{F_s}. Denote $(P_s,Q_s):=F_s(p_0,q_0)$. Then $F_s(p_0,q_0+1)=(P_s,Q_s+1)$. Thus, $\delta_s(0)=(p_0,q_0,P_s,Q_s)$, $\delta_s(1)=(p_0,q_0+1,P_s,Q_s+1)$. Then by Stokes' formula

$$\widetilde{G}(q_0+1,P_1)-\widetilde{G}(q_0,P_1)=\int\limits_{\delta_1}d\widetilde{G}=\int\limits_{\delta_1}(p-P)dq+(Q-q)dP.$$

But $(p-P)dq+(Q-q)dP=pdq-PdQ+d(P(Q-q))$. Hence,

$$\widetilde{G}(q_0+1,P_0)-\widetilde{G}(q_0,P_0)=\int\limits_{\delta_1}pdq-PdQ+d(P(Q-q))$$
$$=\int\limits_{\gamma}pdq-\int\limits_{F\circ\gamma}pdq+\int\limits_{\delta_1}d(P(Q-q)),$$

But the latter integral is equal to 0 because the function $P(Q-q)$ is equal to 0 at the end points of the path δ_1. On the other hand, $\int\limits_{\gamma}pdq=\int\limits_{F\circ\gamma}pdq$. Indeed, denote $\beta(s):=(P_s,Q_s)$ $\overline{\beta}(s):=(P_s,Q_s+1)$, Then, $\int\limits_{\beta}pdq-\int\limits_{\overline{\beta}}pdq$. Hence,

$$\int\limits_{\gamma}pdq-\int\limits_{F\circ\gamma}pdq=\int\limits_{\gamma}pdq+\int\limits_{\overline{\beta}}pdq-\int\limits_{F\circ\gamma}pdq-\int\limits_{\beta}pdq.$$

Consider a square $A=[0,1]\times[0,1]$ and a map $\Phi:A\to\mathbb{R}^2$ defined by the formula $\Phi(t,s)=F_s(\gamma(t))$. Then

$$\int\limits_{\gamma}pdq+\int\limits_{\overline{\beta}}pdq-\int\limits_{F\circ\gamma}pdq-\int\limits_{\beta}pdq=\int\limits_{\partial A}\Phi^*(pdq)=\int\limits_{A}\Phi^*(dp\wedge dq).$$

Denote $\overline{\Phi} := \pi \circ \Phi : A \to T^2$. Recall that the projection $\pi : \mathbb{R}^2 \to T^2$ satisfies $\pi(p,q) = \pi(p,q)+1)$. Therefore, $\overline{\Phi}(0,s) = \overline{\Phi}(1,s)$ for $s \in [0,1]$. We have $dp \wedge dq = \pi^*\omega$, and hence $\Phi^*(dp \wedge dq) = \overline{\Phi}^*\omega$. But by Lemma 3.17 we have $\int\limits_A \overline{\Phi}^*\omega = 0$. Hence,

$$\widetilde{G}(q_0+1, P_0) - \widetilde{G}(q_0, P_0) = \int\limits_A \Phi^*(dp \wedge dq) = \int\limits_A \overline{\Phi}^*\omega = 0.$$

We similarly check that $\widetilde{G}(q_0, P_0+1) = wtG(q_0, P_0)$.

Thus the function $\widetilde{G}$ hence descends to the torus T^2, and hence must have at least 2 critical points, the maximum and the minimum. In fact, one can show that it has to have at least 3 critical points. But Its critical points are in 1-1 correspondence with the fixed points of f, and therefore, f has as many fixed points. This concludes the proof of Arnold's conjecture for the 2-torus for the case when f (and hence F) is C^1-small

Consider now the of the general F. Recall that the Hamiltonian isotopy F_t connects $F_0 = \mathrm{Id}$ with $F_1 = F$. For any integer $N > 0$ we can present F as a composition $F = \widetilde{F}_N \circ \dots \widetilde{F}_1$, where we denote

$$\widetilde{F}_k = F_{\frac{k}{N}},\ k = 1, \dots, N.$$

By taking N sufficiently large we can make all the diffeomorphisms $\widetilde{F}_k$ arbitrarily C^1-small.

We consider below the case $N = 2$, the general case differs only in the notation.

As above, we can conclude, that the product $\Gamma := \Gamma_{\widetilde{F}_1} \times \Gamma_{\widetilde{F}_2} \subset \mathbb{R}^8$ of the graphs of $\widetilde{F}_1$ and $\widetilde{F}_2$ is given by the equations

$$p_1 = p_1(q_1, P_1), Q_1 = Q_1(q_1, P_1), p_2 = p_2(q_2, P_2), Q_2 = Q_2(q_2, P_2).$$

Furthermore, we have $p_i dq_i + Q_i dP_i = dG_i$ and the functions $\widetilde{G}_i = G_i - q_i P_i$ are $\mathbb{Z}^2$-periodic, $i = 1, 2$. Set $\widetilde{G}(q_1, P_1, q_2, P_2) := G_1(q_1, P_1) + G_2(q_2, P_2)$. Fixed points of F are in 1-1 correspondence with the intersection $\Gamma \cap \{p_2 = P_1, Q_1 = q_2, p_1 = P_2, Q_2 = q_1\}$, i.e. with the *zeroes of the 1-form*

$$\begin{aligned}\alpha :=&(p_1 - P_2)dq_1 + (Q_1 - q_2)dP_1 + (p_2 - P_1)dq_2 + (Q_2 - q_1)dP_2 \\ =&dG(q_1, q_2, P_1, P_2) + d\Big((P_1 - P_2)(q_1 - q_2)\Big).\end{aligned}$$

Changing the variables $(q_1, q_2, P_1, P_2) \mapsto (q_1, u_1 := q_2 - q_1, P_1, U_1 := P_2 - P_1)$ we get

$$\alpha = d(\widehat{G} + u_1 U_1), \text{ where } \widehat{G}(q_1, u_1, P_1, U_1) := \widetilde{G}(q_1, q_1 + u_1, P_1, P_1 + U_1).$$

Similarly to the proof of Lemma 3.19, one can check that the function $\widehat{G}$ is periodic with respect to all variables, and in particular, in variables (q_1, P_1), and hence it descends to a function

$$T^2 \times \mathbb{R}^2 = \mathbb{R}^2/\{q_1 \sim q_1 + 1, P_1 \sim P_1 + 1\} \to \mathbb{R}.$$

Note also that this function and its derivatives are bounded. Then the following lemma implies that the function $\widetilde{G}(q_1, P_1, u_1, U_1)$ must have some critical points, which, as we showed above, corresponds to fixed points of F.

Lemma 3.20. *Let M be a closed manifold, $C : \mathbb{R}^n \to \mathbb{R}$ a non-degenerate quadratic form, and $\phi : M \times \mathbb{R}^n \to \mathbb{R}$ a smooth function which is bounded and has bounded 1st derivatives. Then the function $\psi(x, y) = \phi(x, y) + C(y),\ x \in M, y \in \mathbb{R}^n$ has at least 1 critical point.*

Proof. [Sketch of the proof] We can assume that $C(y) = \sum_1^k y_j^2 - \sum_{k+1}^n y_j^2$. Suppose that $k \neq 0$ (if $k = 0$ we can change the sign of the function ψ). Consider a map $h : \mathbb{R}^k \to M \times \mathbb{R}^n$ such that $h(y_1, \ldots, y_k) = (x_0, y_1, \ldots, y_k, 0, \ldots, 0)$ when $||y||^2 = \sum_1^k y_j^2$ is large enough. Let us denote by $\mathcal{H}$ the space of all maps h with this property. For any $h \in \mathcal{H}$ the function $\psi \circ h : \mathbb{R}^k \to \mathbb{R}$ is bounded below and achieves its minimal value at a point $a_h \in \mathbb{R}^k$. Indeed, $\lim_{||y|| \to \infty} \psi \circ h = +\infty$. Denote $b_h = h(a_h) \in M \times \mathbb{R}^n$. There exists a point $b \in M \times \mathbb{R}^n$ such that $\psi(b) \geq \psi(b_h)$ for all $h \in \mathcal{H}$ (why?). Then b is a critical point of ψ (why?).

Bibliography

[1] C. Conley and E. Zehnder, The Birkhoff-Lewis fixed point theorem and a conjecture of V.I. Arnold, *Invent. Math.*, **73**(1983), 33–49.

[2] H. Poincaré, Sur une Th'eorème de géométrie, *Rend. Circ. Mat. Palermo*, **33**(1912(, 375–507.

[3] Y. Eliashberg, Estimates on the numeber of fixed points of area preserving transformations, *Proc. VINITI*, 1979 (in Russian).

[4] Y. Eliashberg and M. Gromov, Lagrangian intersection theory. Finite-dimensional approach, in *Geom. Diff eq.*, AMS, 1998, 123–135.

[5] A. Floer, The unregularized gradient flow of the symplectic action, *Comm. Pure Appl. Math.*, **41**(1988), 775-813.

[6] M. Gromov, Pseudoholomorphic curves in symplectic manifolds, *Invent. Math.*, **82**(1985), 307–347.

[7] K. Fukaya and K. Ono, Arnold conjecture and Gromov-Witten invariants for general symplectic manifolds, *Fields Inst. Comm*, **24**(1999), 173–190.