Ordinary differential equations

according to the text book with this title by V.I. Arnol'd.

Introductionary lectures held at Leipzig University during the summer semester 2011
by B. Herzog
Fr. 7.30 - 9.00, Ph. 2-18 (exercises)
Fr. 9.15 - 10.45, Ph. 2-18 (lectures)

Exercises and lectures from July 15, 2011 are moved to

Fr. July 22, 15.00-18.00, Math Institute, Felix-Klein-Hörsaal, Johannisgasse.
Mo. July 25, 15.00-18.00, Math Institute, Felix-Klein-Hörsaal, Johannisgasse.

Writen examination:
July 26, 8.00-11.00, Room 1-22, Mathematical Institute, Johannisgasse
Repetition:
September 26, 8.00-11.00, Room S 114, Augustusplaz

Conditions

- There will be a written test at the end of the semester.
- To enter this test you will have to solve problems every week.
- Each week you will usually have to deal with three problems, whose solutions you will have to return in written form at the beginning of the exicises at Friday moring.
- For each solution you can achieve up to three points. For the right to enter the written examination you need 50% of the maximum number of points.

Notation

∂X	boundary of the set X (in some topological space), see 2.2.2.
$\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x}, \frac{\partial}{\partial t}$	derivatives in the directions of, respectively, the x_i - , x-, and t- axes, which are identified with the standard unit vectors in the direction of these axes, see 1.4.2.
$\frac{dx}{dt}(t_0)$	the tangent vektor of the curve $x: I \longrightarrow M, t \mapsto x(t)$, at the point $x(t_0)$, see Example 4 of 1.4.3.
$\frac{\partial f}{\partial x}(p)$	the Jacobian matrix of a map $f: U \longrightarrow V$ between open sets U, V in euclidian space at the point $p \in U$, see Appendix.
$d_p\varphi$	differential of a map $\varphi: M \longrightarrow M'$ at the point $p \in M$, see 1.4.3, (which is a linear map $d_p\varphi: T_p(M) \longrightarrow T_{\varphi(p)}(M')$, in appropriately chosen coordinates it is multiplication by the Jacobian matrix of φ, see 1.4.3 Example 2).
$d\varphi$	differential form of a map $\varphi: M \longrightarrow M'$, see Remark (iii) of 1.4.7, (which is the map $\varphi: TM \longrightarrow TM'$ whose restriction to $T_p(M)$ is $d_p\varphi$ for every $p \in M$).
Aut(M)	transformation group of the set M, see 1.2.2.
$\mathcal{O}_{M,p}$	local ring of the manifold M at the point p, see 1.4.1.

TM	tangent bundle of a manifold, see 1.4.6.
$T_p(M)$	tangent space of the manifold M at the point p∈M, see 1.4.2.
Γ_f	graph of the map f: A ⟶ B, i.e., $\Gamma_f := \{(a,f(a)) \mid a\in A\}$.
S(I, A)	The space of solution $I \longrightarrow \mathbb{R}^n$ of the linear differential equation $\frac{dx}{dt} = A(t)\bullet x$, see 2.2.1. B.

1. Basic notions

1.1. A first definition

An ordinary differential equation (ODE) is an equation depending upon a function, say

$$x = x(t),$$

together with certain derivatives of this function and the variable t. In other words, an ordinary differential equation is something like this:

$$f(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \dots, \frac{d^nx}{dt^n}, t) = 0. \tag{1}$$

More precisely, such an equation is called n-th order ordinary differential equation.

The topic of the theory is to find solutions (one, as many as possible or all) of such an equation, i.e. all functions x such that this equation becomes an identity.

Moreover, one wants to find qualitative properties of the solutions.

Remarks

(i) We do not assume, that the values of the function x are real or complex numbers. They may be contained in any set, where differention is possible. Typically,

$$x(t) \in \mathbb{R}^n.$$

Likewise we allow f to be a vector valued function,

$$f(y_0, y_1, \dots, y_n) \in \mathbb{R}^m.$$

This way there will be not really a difference between one single differential equation and a system of such equations.

(ii) On the other hand, the variable t is always assumed to vary in an open subset of the real axis,

$$t \in U \subseteq \mathbb{R}, \text{ U open.}$$

Remember, open means, for every value t in U there is an open intervall around t, which is completely contained in U,

$$t \in U \Rightarrow t \in (t - \varepsilon , t + \varepsilon) \subseteq U \text{ for some } \varepsilon > 0.$$

(iii) Consider a massive particle (of mass equal to m) moving with the time t in 3-space, say

$$x(t) \in \mathbb{R}^3,$$

such that

$$\frac{d^2x(t)}{dt^2} - g/m = 0.$$

This ordinary differential equation is probably the first one has to solve in classical mechanics.

(vi) A considerable part of mathematics and physics is related to differential equations, for example

$$\frac{\partial^2 x}{\partial u^2} + \frac{\partial^2 x}{\partial v^2} + \frac{\partial^2 x}{\partial w^2} = 0\ ,\ x(u,\ v, w) \in \mathbb{R}, \qquad \text{(Laplace-Gleichung)} \quad (2)$$

or

$$\frac{\partial^2 x}{\partial u^2} + \frac{\partial^2 x}{\partial v^2} + \frac{\partial^2 x}{\partial w^2} - \frac{\partial^2 x}{\partial t^2} = 0\ , x\ (u,\ v, w) \in \mathbb{R}, \text{(Wellen-Gleichung)} \quad (3)$$

or

$$\frac{\partial f(u,v)}{\partial u} + i\bullet\frac{\partial f(u,v)}{\partial v} = 0, f(u,v) \in \mathbb{C} \qquad \text{(Cauchy-Riemann equations)} \quad (4)$$

Other examples are
the Hamiltopn equations in classical mechanics,
the Maxwell equations in electrodynamics,
the Schrödinger equation in Quantum mechanics
or various Yang-Mills equantions in Quantum field theory.

All these equations are not ordinary differential equations, since they contain partial derivatives with respect to different variables. They are partial differential equations (PDE's).

(vii) There are essential differences between ordinary and partial differential equations:

- The theory of ODE's is much easier than any theory of PDE's. If you want to understand partial differential equations you need a good knowledge of the behavior of the ordinary ones.
- One can even say there is a well understood theory about the behavior of all ODE's, while there is no common theory of PDE's. Two different PDE's usually require different theories. Even the questions one has to ask for are usually different for different PDE's.
- The theory of the equations of type (2) is called Potential theory in the real case or Hodge theory in the complex projective case.
- Those of type (3) lead to (classical) wave theory.
- Those of type (4) result in complex analysis.
- The theory of the (finitie dimensional) Hamilton equations is called simplectic geometry (or classical mechanics).
- etc.

(vii) Our topic, the ordinary differential equations, is much easier.

Convention: If not stated otherwise, we will always assume that the functions solving our equations are sufficiently often continuously differentiables.

Problem 2:
(viii) Note that every systems n-th order ODE's is equivalent to system of first order ODE's.

Two systems ODE's are called equivalent in case they have (essentially) the same solutions (i.e. the solutions of one system can be directly seen from the solutions of the other).

Equation (1) can be written as follows.

$$f(x_0, x_1, x_2, \dots, x_n, t) = 0.$$
$$x_1 \qquad = dx_0/dt$$

$x_2 \quad = dx_1/dt$

...

$x_n \quad = dx_{n-1}/dt$

We introduce a new notation:

$X := (x_0, x_1, x_2, \dots, x_n)$, $Y = (y_0, y_1, y_2, \dots, y_n)$

$F(X, Y, t) := (f(x_0, x_1, x_2, \dots, x_n, t), x_1 - y_0, x_2 - y_1, \dots, x_n - y_{n-1})$.

Then the above system reads

$$F(X, \frac{dX}{dt}, t) = 0.$$

This way we see that an n-th order ODE is equivalent to an first oder ODE.

To start with, we will give a quite different look at the theory of ordinary differential equations.

1.2 Phase spaces and phase flows

1.2.1 Finite dimensional deterministic processes

The theory of ordinary differential equations deals with deterministic, finite dimensional and differential processes evolving in time.

Finite dimensional means, the state of such a process is given by a point in a finite dimensional space (i.e. by a finite number of parameters). This space, consisting of all possible states, is called phase space.

Deterministic means, that all future and all past states are uniquely determined by the current state.

Formally, if M denotes the phase space, there is a map

$$g: \mathbb{R}\times M \longrightarrow M, (t, x) \mapsto g^t x,$$

such that $g^t x$ is the state of the process at time t whose state at time 0 is x.

In other word, given the state x at time 0, the state $g^t x$ at time t of this process is uniquely determined.

The map g is called the phase flow of the process.

The process is called differentiable, if M ist a space where differentiaton is defined (a differentiable manifold) and the map g is differentiable.

Remarks

(i) Let x be the state of the process at time 0. Than its state at time t is $g^t x$.
Similarly, if y is the state at time s, its state at time s+t is $g^s y$. Substituting $g^t x$ for y we obtain

$$g^{s+t}x = g^s g^t x \text{ for arbitrary } x \in M, s, t \in \mathbb{R}. \tag{1}$$

time	state
0	x
t	$g^t x$

s+t	$g^s g^t y$

Moreover, by the very definition of the phase flow,

$$g^0 x = x \text{ for each } x \in M. \tag{2}$$

The identities (1) and (2) are the most important properties of the phase flow.

(ii) For each $t \in \mathbb{R}$ the phase flow defines a map

$$g^t: M \longrightarrow M,\ x \mapsto g^t x.$$

Thus we have a family of maps

$$\{g^t\}_{t\in\mathbb{R}}$$

Note that the composition of any two maps in this family is again a map in this family:

$$g^s g^t = g^{s+t}.$$

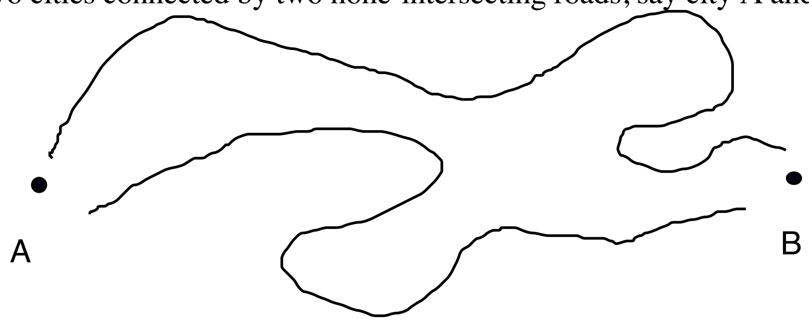
The family $\{g^t\}_{t\in\mathbb{R}}$ is called the <u>one-parameter group</u> of transformations associated with the phase flow g.

Problem 1: Prove that each map of the family $\{g^t\}_{t\in\mathbb{R}}$ is bijective and that these maps form a commutative group.

(iii) The notion of phase space is the most important notion in our context. Often the introduction of the phase space of a problem is already the solution of the problem.

Example

Consider two cities connected by two none-intersecting roads, say city A and city B.



About these two roads the following fact is known:

Assume that there are two cars in city A, which are connected by a cord of length $< 2\cdot\ell$. Then we know that it is possible for these cars to travel to city B on different roads without breaking the cord.

Question:

Assume that there are two circular wagons (moving discs) of radius ℓ, one in city A and one in city B. Is it possible to move these two wagons simultaneously to the other city without touching each other ?

Intuitively one would expect that this is impossible. But how to prove this ?

It turns out that the solution is to form the phase space describing the situation. The movement of each car or wagon it given by the single parameter, namely the distance it has moved from its city on its road.

Denote the distance on the first road by x and the distance on the second by y. We may assume that both roads have the same length 1 (we may use for each road its own unit to measure the distances). Thus both parameters vary between 0 and 1,

$$0 \le x \le 1, 0 \le y \le 1.$$

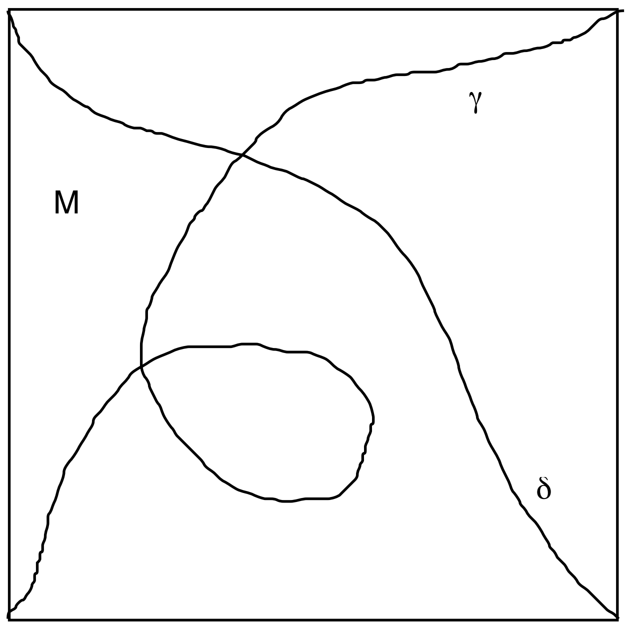
Therefore, the positions of two bodies on the two roads is given by a point in the square

$$M := \{ (x, y) \in \mathbb{R}^2 \mid 0 \le x \le 1, 0 \le y \le 1 \}$$

Assume that both roads start in city A and end in city B, i.e. the coordinate 0 describes in both roads the starting point in city A and the coordinate 1 describes city B.

The movement from A to B of the two cars is given by a curve γ from (0,0) to (1,1), since both start in the same city A and end in the same city B.

Similarly, the movement of the two wagons is given by a curve δ from (1,0) to (0,1). The situation tells us that the two curves must intersect.[1]



Let (x,y) be a point of intersection. This point describes a situation, where the two cars are in the same position like the two wagons. But the distance of the two cars is be assumption less than $2\cdot\ell$, while the two wagons should be at a distance greater than $2\cdot\ell$, which is impossible.

There are no ODE's involved in the above problem. But we see that the pure construction of the phase space already solves a not quite trivial problem !

1.2.2 One-parameter groups of transformations

Let M be a set. A <u>transformation</u> (or <u>automorphism</u>) of M ist simply a bijektive map

$$M \longrightarrow M.$$

The transformations of M form a group, denoted

[1] Each curve divides the square in two components, and the other curve connects points of different components.

$$\mathrm{Aut}(M),$$

whose group law is the usual composition of maps. A one-parameter group of transformations (or automorphisms) is by definition a group homomorphisms

$$h:\mathbb{R} \longrightarrow \mathrm{Aut}(M),\ t \mapsto h^t$$

from the additive group of real numbers to the transformation group Aut(M). Equivalently, a one-paramter group is given by a map

$$\mathbb{R}\times M \longrightarrow M,\ (t, x) \mapsto h^t(x),$$

such that

1. $h^0(x) = x$ for every $x \in M$
2. $h^s(h^t(x)) = h^{s+t}(x)$ for every $x \in M$ and arbitrary $s,t \in \mathbb{R}$.

Remarks

(i) The one-parameter group of a phase flow is obviously a one-parameter group in the sense just defined.

(ii) In case M is a (finite dimensional) differentiable manifold, and h is a differentiable map, than h is nothing else but a phase flow.

1.2.3 Phase curves and fixed points of the phase flow

Let M be a phase space with phase flow

$$g: \mathbb{R}\times M \longrightarrow M,\ (t, x) \mapsto g^t(x).$$

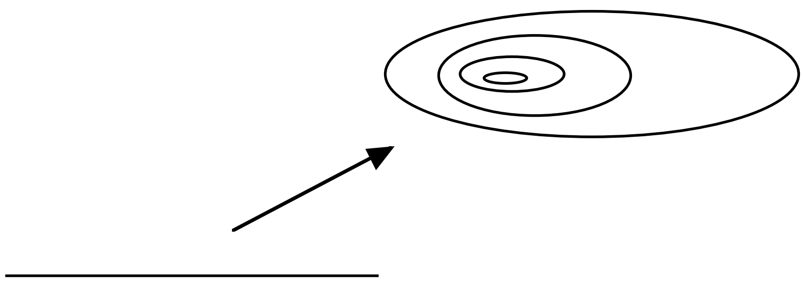
Then for each point x the map

$$\mathbb{R} \longrightarrow M,\ t \mapsto g^t(x),$$

is called the motion of the point x under the effect of the phase flow g. The image of this map is called the phase curve through x or the trajectory through x.
In case the phase curve through x consists of one single point, i.e.,

$$g^t(x) = g^0(x) = x \text{ for every } t \in \mathbb{R},$$

the point x is called equilibrium or fixed point of the phase flow.



The space $\mathbb{R}\times M$ is also called extended phase space of the given flow. An integral curve of the flow is by definition the graph of a motion of the flow, i.e. a set in the extended phase space of the type

$$\{ (t, g^t(x) \mid t \in \mathbb{R}\}$$

Remark

The line $\mathbb{R}\times\{x\}$ is an integral curve if and only if the point x is a fixed point of the flow.

The $\mathbb{R}\times\{x\}$ is an integral curve if and only if there is some point $x' \in M$ such that

$$\mathbb{R}\times\{x\} = \{ (t, g^t(x') \mid t \in \mathbb{R}\},$$

i..e.,

$$g^t(x') = x \text{ for every } t \in \mathbb{R}.$$

For t = 0 this latter condition gives $x' = g^0 x' = x$. Thus:

$$\mathbb{R} \times \{x\} \text{ is an integral curve} \Leftrightarrow g^t(x) = x \text{ for every } t \in \mathbb{R}.$$

The condition on the right hand side means that x is a fixed point of the flow.

Problem 2
Prove that there is precisely one integral curce through each point of the extended phase space.

Problem 3
Prove that for every $s \in \mathbb{R}$, the following translation of the extended phase space maps integral curves into integral curves.

$$h^s: \mathbb{R} \times M \longrightarrow \mathbb{R} \times M, (t, x) \mapsto (t + s, x).$$

1.2.4 Phase flows and direction fields in the plane

Let $g: \mathbb{R} \times M \longrightarrow M$ a phase flow with $M = \mathbb{R}$. Through every point

$$(t_0, x_0) \in \mathbb{R}^2$$

of the extented phase space goes presicely one integral curve

$$C(t_0, x_0) = \{(t, g^t g^{-t_0} x_0) \mid t \in \mathbb{R}\}.$$

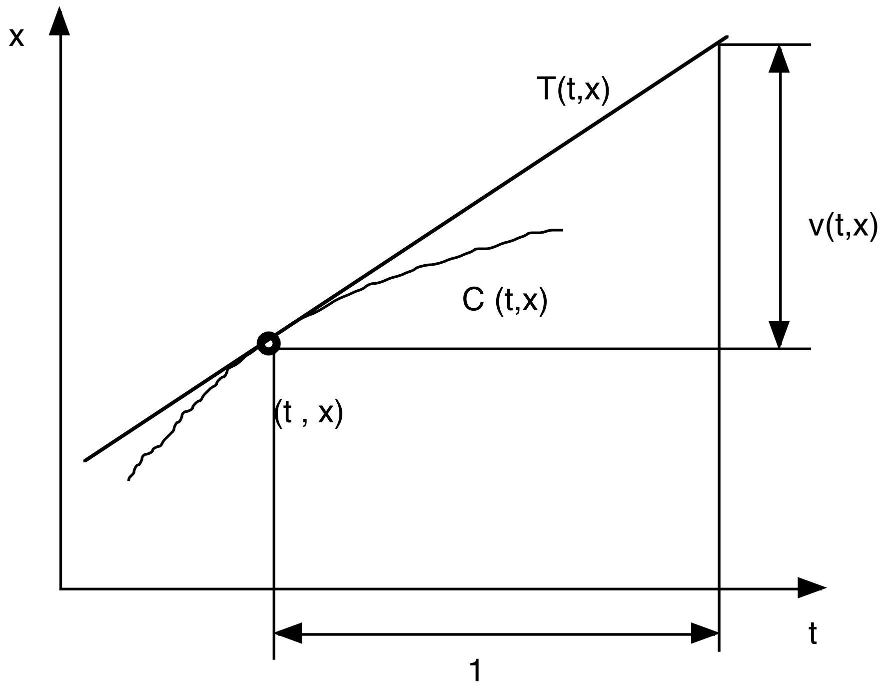
Let

$$T(t_0, x_0)$$

denote the tangent to this curve at the given point (t_0, x_0). This way we get a map

$$\mathbb{T}: \mathbb{R}^2 \longrightarrow \{\text{lines in } \mathbb{R}^2\}$$

which associates with each point in $\mathbb{R}^2$ a lin going through this point. Such a map is called direction field.



Problem: Is it possible to recalculate the integral curves of the flow from this direction field ?

Generalization of the problem: forget that the direction field comes from a flow.

A curve is called integral curve of a direction field if the tangent line at any point of the curve is equal to the line of the direction field at that point.

Generalized problem: how to find integral curves of direction fields.

More precisely, let

$$v(t, x)$$

denote the slope of the direction field line associated with the point (t,x). The problem is to recalculate the integral curves from the function

$$v: \mathbb{R}^2 \longrightarrow \mathbb{R} \cup \{\infty\}.$$

Equivalently, solve the differential equation

$$\frac{dx}{dt} = v(t, x).$$

A trivial case

Assume the the following additional conditions are satisfied

1) There is no direction in the field parallel to the x-axis.
2) Translations in the direction of the x-axis map the lines of the direction field into lines of that field.

- first condition means, the the values of v are always finite and the integral curves are graphs of functions, say
$$x = x(t).$$
- The second condition means,
$$v = v(t)$$
does not depent upon x.
- The graph of the function $x = x(t)$ is an integral curve, if and only if
$$\frac{dx}{dt} = v(t).$$

Integration yields

$$\int_{t_0}^{t} \frac{dx}{dt} dt = \int_{t_0}^{t} v(t)\, dt$$

hence

$$x(t) = x(t_0) + \int_{t_0}^{t} v(t)\, dt.$$

A less trivial case: $\frac{dx}{dt} = v(x).$

This is not the type of equation corresponding to the above trivial case. But it asks for the integral curves of a given direction field.

But there is another problem with the same solution: look for a function $t = t(x)$, such that the graph is an integral curve of the given direction field. The differential equation for this problem,

$$\frac{dt}{dx} = \frac{1}{v(x)},$$

can be solved:

$$t(x) = t(x_0) + \int_{x_0}^{x} \frac{dx}{v(x)}$$

Condition: v(x) should be continuos and none-zero.
Example: The equation of normal reproduction

$$\frac{dx}{dt} = x.$$

(Remove the t-axis x = 0 from $\mathbb{R}^2$).

Solution:

$$
\begin{aligned}
t &= t_0 + \int_{x_0}^{x} \frac{dx}{x} \\
&= t_0 + [\ln |x|]_{x_0}^{x} \\
&= t_0 + (\ln |x| - \ln |x_0|) \\
&= t_0 + \ln |x/x_0| \\
e^{t-t_0} &= |x/x_0| \\
x &= x_0 \cdot e^{t-t_0}
\end{aligned}
$$

The graph of this curve is the ingral curve through the point (t_0, x_0).

The general (phase flow) case:
As we know, translations in the diretion of the t-axis map integral curves of phase flows into integral curves. This means that the function v(t,x) does not depend upton t. Thus the above solved problem is already the general case.

Remarks

(i) Liouville has proved that the equation

$$\frac{dx}{dt} = x^2 - t.$$

cannot be solved just by integration.

(ii) A slight generalization of the above example,

$$\frac{dx}{dt} = k \cdot x$$

decribes population growth[2] for k > 0 and radioactive decay for k < 0 (or air density)[3]. Calculations like above give:

$$x = x_0 \cdot e^{k(t-t_0)}$$

The most remarkable property of this function is, that there is a constant time period after which the absolute value of x doubles:

$$T = k^{-1} \cdot \ln(2)$$

(T is negative if so is k, and |T| is called <u>half life</u> in this case).

(iii) The above equation of normal reproduction describes population growth only for a restricted time period. At some point the population might become too large, so that the way of development changes due to food problems. One way to get a model for this phenomen is to assume that the constant k above depends upon the size of the population,

$$k = k(x).$$

In the simplest type of model one assumes the k is a linear function of x, say

---

[2] Population growth is proportional to the number of individuals.

[3] Air density is half as large at Mt. Elbrus in abour 5,6 km height compared with the density at sea level.

$$k = a - bx.$$

An easy transformation leads to the differential equation

$$\frac{dx}{dt} = (1 - x)x.$$

This equation has an instable[4] equilibrium at

$$x = 0.$$

and a stable[5] equilibrium at

$$x = 1$$

All integral curves in the upper half plane are asymptotic to the curve x = 0.

1.3 Differentiable Manifolds

1.3.1 Aboud the notion of differentiable process

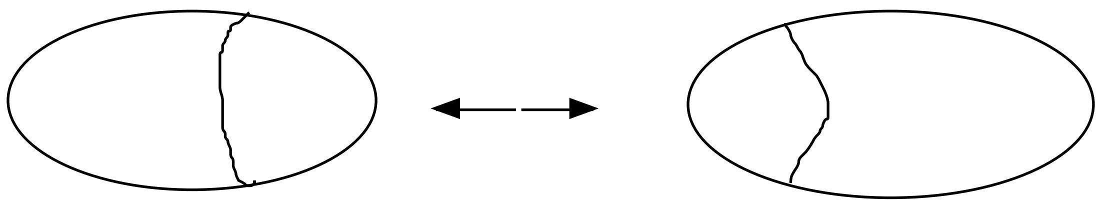
The definitions of the previous sections formally define the notion of deterministic process. The notions of finiteness and differentiability have been treated until now in a rather uncertain way. We want to be now somewhat more precise.

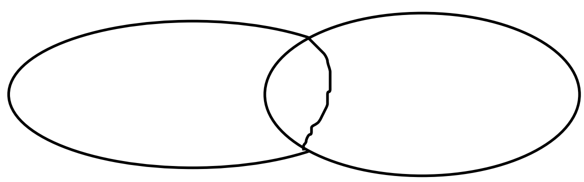
Together these conditions mean that the phase space is a finite dimensional differentiable manifold and the phase flow is a differentiable map. Hence we have to define the notions of differentiable manifold and differentiable map.

Examples
of differentialbe manifolds are the euclidian spaces, open subsets of euclidian spaces, circles, spheres, toruses, etc.
Examples of differentiable maps are map between differentiable manifolds whose coordinate functions are differentiable.

More generally, a differentiable manifold (of finite dimension) is a space which looks locally like an open set in euclidian space and is such that the notion differentiable function is well-defined and such that the derivative of such functions can be formed.

Differentiable manifolds can be obtained by gluing together open sets of an euclidian space of a fixed (finite) dimension identifying commen open subsets.





[4] An integral curve x = x(t) with a state close to 0 but different from zero increases the absolute value of its state with increasing time, i.e. it moves away from the equilibrium state.
[5] An integral curve x = x(t) with a state close to 1 but different from 1 develops with increasing time such that its state approaches the vale 1, i.e. comes closer and closer to this state.

The identification has to be done in such a way that the notions of differentiability on both piecs coicide, i.e. one has to use bijective maps which are differentiable in both directions.

To give a more formal definition, we have first to introduce the notion of topological space.

1.3.2 Topological spaces

A topological space is a set X together with a family T(X) of subsets of X, called open sets of X, such that the following conditions are satisfied.

(i) The empty set and the whole space are open sets of X, i.e.,

$$\emptyset \in T(X) \text{ and } X \in T(X).$$

(ii) The intersection of any two open sets is open, i.e.,

$$U, V \in T(X) \Rightarrow U \cap V \in T(X).$$

(iii) The union of any family of open sets is open, i.e.,

$$U_i \in T(X) \text{ for every } i \in I \Rightarrow \bigcup_{i\in I} U_i \in T(X).$$

In this situation, the family T(X) is called the topology of X. An open set $U \in T(X)$ containing a given point $x \in X$ is also called (open) neightbourhood of x.

A topological space X is called Hausdorff space , if there exist, for any two given different points $x, y \in X$, open subsets $U, V \subseteq X$ such that

$$x \in U, y \in V \text{ and } U \cap V = \emptyset.$$

A map f: X $\longrightarrow$ between topological spaces X and X' is called continuous, if for every open set $U' \in T(X')$ the set

$$f^{-1}(U') := \{x \in X \mid f(x) \in U'\}$$

is an open set of X.

The map f: X $\longrightarrow$ X' is called a homeomorphism, if it is bijective and both f and f^{-1} are continuous. In this case X and X' are called homeomorph. As topological spaces they are consindered as essentially equal.

Example
For every set X the family

$$T(X) := \{U \mid U \subseteq X\}$$

of all subsets defines on X the structure of a topological space. This topology T(X) is called discrete topology of X. The topological space whose topology is the discrete one is called discrete topological space. A discrete topological space is always a Hausdorff space.

Example
For every set X the family

$$T(X) = \{\emptyset, X\}$$

defines a topology on X. If X is equipped with this topology and has at least two points, it is never a Hausdorff space (each point is in every neighbourhood.

Example
For every two points $x = (x_1,...,x_n)$ and $y = (y_1,...,y_n)$ in $\mathbb{R}^n$ write

$$d(x, y) := \sqrt{(x_1-y_1)^2+...+(x_n-y_n)^2}$$

for the distance between x and y und denote by

$$U_\varepsilon(x) := \{y \in \mathbb{R}^n \mid d(y, x) < \varepsilon\}$$

the ε–neightbourhood of all points of distance less than ε from x. A subset

$$U \subseteq \mathbb{R}^n$$

is called open, if there is, for every $x \in U$ there is some positive real number ε such that $U_\varepsilon(x)$ is completely containted in U,

$$x \in U \Rightarrow U_\varepsilon(x) \subseteq U \text{ for some positive real } \varepsilon.$$

The set $T(\mathbb{R}^n)$ of all open sets of $\mathbb{R}^n$ as just defined is called the euclidian topology of $\mathbb{R}^n$. One can easily prove that a map

$$f: \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

where both spaces are equipped with the euklidian topology, is continuous in the above defined sense if and only if it safisfies the conditions of the δ–ε-criterion.

A examples of an open set in $\mathbb{R}^2$ is the open disc of radius r around the origin o := (0,0),

$$U_r(o) \text{ is open in } \mathbb{R}^2.$$

The closed disc (including the border)

$$\{ y \in \mathbb{R}^2 \mid d(x,o) \le r\} \qquad (1)$$

is not an open set. To see this, take a point y on the border of this set. Then there is no ε-neightbourhood around this point, which is completely contained in (1).

1.3.3 Manifolds

An n-dimensional manifold or n-manifold is a topological Hausdorff space X which looks locally like an open set in $\mathbb{R}^n$, i.e., for every point $x \in M$ there is an open neightbourhood $U \subseteq M$ of x, an open subset $V \subseteq \mathbb{R}^n$ and a homeomorphism

$$\varphi: U \longrightarrow V$$

(which identifies the points of U with the point of V). These homeomorphisms are are called charts of M. The domain of definition of a chart is called coordinate neightbourhood of M. A family of charts such that the coordinate neightbourhoods cover M is called an atlas of M.

Example

Take two copies of the complex plane, say

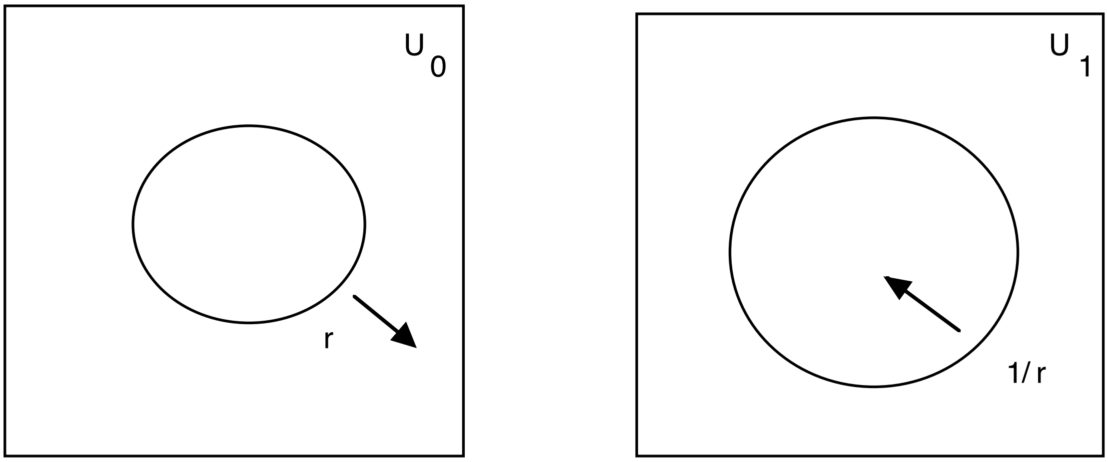
$$U_0 := \mathbb{C} \text{ und } U_1 := \mathbb{C}.$$

On both copies fix an open subset,

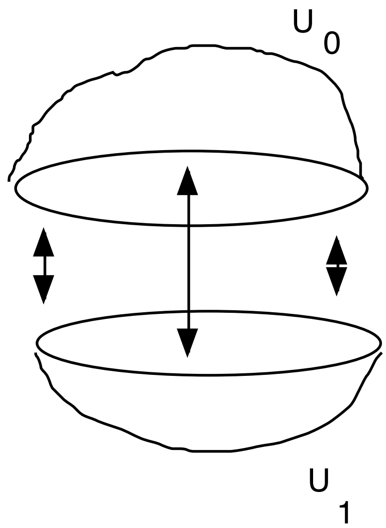
$$U' := \{z \in U_0 \mid z \neq 0\}$$

$$U'' := \{z \in U_1 \mid z \neq 0\}$$

Now identify each point $z \in U' - \{0\}$ with the point $1/z \in U''$.



Note that increasing circles around the origing in U_0 are identified with decreasing circles around the origin in U_1. The unit circle in U_0 is identified with the unit circle in U_1 (with the orientation reversed).



Thus, the union

$$\mathbb{P}^1_{\mathbb{C}} := U_0 \cup U_1$$

can be identified with a 2-dimensional sphere, the so-called Riemann sphere. Another way to describe this set is to identify it with the set of 1-dimensional complex linear subspaces in complex 2-space $\mathbb{C}^2$.[6]

The identifying functions

$$g' : U' \longrightarrow U'' \text{ , } z \mapsto 1/z$$

$$g'' : U'' \longrightarrow U' \text{ , } z \mapsto 1/z$$

are obviously differentiable in each point of their domain of definition. Therefore, a function

$$f: U' \longrightarrow \mathbb{R}$$

is differentiable in a given point $p \in U'$, if and only if the composition

$$f \circ g'' : U'' \longrightarrow \mathbb{R}$$

[6] Hence the notation $\mathbb{P}^1_{\mathbb{C}}$:.

is differentiable in the corresponding point of U". In other words, the notions of differentiability of a function is independent upon wheter it is considered as a function on U' ore a function on U". This way one has the notion of differentiable function on the Riemann sphere.
Note that the two identifying functions g', g" are even complex analytic, such that the notion of complex analytic function is defined on the Riemann sphere, i.e., the Riemann sphere even has the structure of a complex manifold.

The open subsets U_0 and U_1 are called charts ore local coordinate systems of the manifold $\mathbb{P}^1_{\mathbb{C}}$. Since their union is equal to the whole manifold, the set $\{U_0, U_1\}$ is called an atlas of the Riemann sphere.

For arbitrary manifolds these notions are defined in a similar way.

1.3.4 Differentiable manifolds

A differentiable manifold of dimension n is a topological (Hausdoff-) space M equipped with a covering by open sets,

$$M = \cup_{i\in I} U_i ,$$

such that each U_i can be identified with an open subset $V_i \subseteq \mathbb{R}^n$, i.e. there are bijective maps

$$\varphi_i: U_i \longrightarrow V_i$$

such that

(i) φ_i is homeomorphic, i.e., φ_i and φ_i^{-1} are continuous functions for every i.

(ii) For every pair $i, j \in I$ such that $U_i \cap U_j$ is none-empty the following map between open subsets in euclician space is continuously differentiable[7]

$$\varphi_i(U_i \cap U_j) \longrightarrow \varphi_j (U_i \cap U_j), x \mapsto \varphi_j(\varphi_i^{-1}(x)).$$

A chart or local coordinate system of this maninfold M is by definition a bijective map

$$\varphi: U \longrightarrow V$$

of an open subset $U \subseteq M$ onto an open subset $V \subseteq \mathbb{R}^n$ such that the conditions (i) and (ii) above continue to be satisfied if φ is added to the familie of φ_i used in the definition above. In this situation, U is called the coordinate neighbourhood of the chart φ.

An atlas of M is a family of charts of M such that the associated coordinate neighbourhoods cover M.

Two differentiable manifolds with the same underlying set M but possibly different families $\{\varphi_i: U_i \longrightarrow V_i\}$ are considered equal if a bijection $\varphi: U \longrightarrow V$ is a chart with respect to the first manifold if and only if it is a chart with respect to the other.

Remark

---

[7] We will usually assume that these maps are sufficiently often continuously differentiable, say r times. In this case M is called a C^r-manifold. In case $r = \infty$ one also says, M is a smooth manifold. If these functions are analytic (given by power series), on writes $r = \omega$ and says that M is an analytic manifold.

The condition in (ii) of being continuously differentiable can be replaced by the condition to be r times continously differentiable (r = 0, 1, ... , ∞, ω) where the case r = ω means that the involved functions are analytic, i.e., can be locally expanded into power series. The resulting manifold is than called a C^r-manifold (smooth manifold in case r = ∞ and analytic manifold in case r = ω).

Example

Let

$$S^1 := \{ z \in \mathbb{C} \mid |z| = 1\}$$

denote the unit circle of the complex plane. We want to show that S^1 is an analytic manifold in the sense of the above definition. The construct an atlas, it is usefull to consider the complex exponential function. It defines a surjective map

$$f: \mathbb{R} \longrightarrow S^1, r \mapsto e^{2\pi ir} = \cos 2\pi r + i\bullet\sin 2\pi r,$$

such that

$$f(\rho') = f(\rho'') \Leftrightarrow \rho' - \rho'' \in \mathbb{Z}. \qquad (1)$$

In particular, the restriction

$$f_r := f|_{(r-\varepsilon,r+\varepsilon)}: (r - \varepsilon, r + \varepsilon) \longrightarrow S^1, 0 < \varepsilon < \frac{1}{2}$$

of f to any open intervall of length < 1 is injective. Write

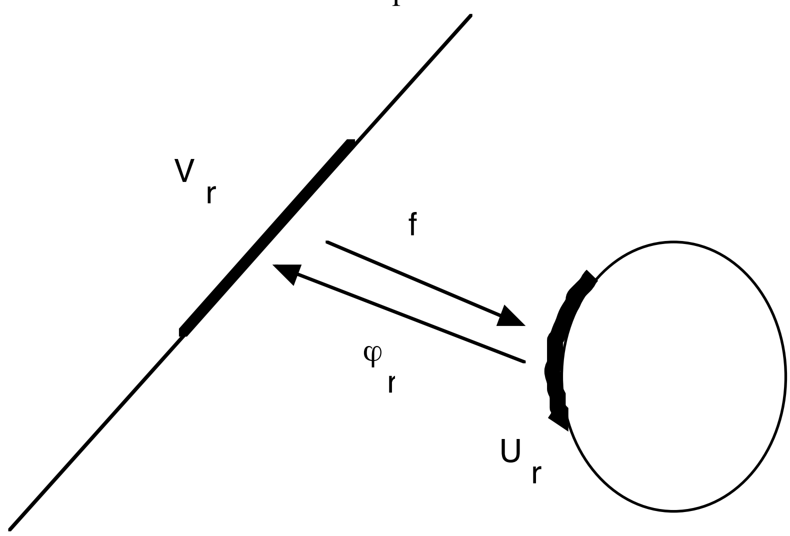
$$V_r := (r - \varepsilon, r + \varepsilon) \subseteq \mathbb{R}$$

$$U_r := \mathrm{Im}(f_r) = f_r(V_r) \subseteq S^1.$$

Then

$$\varphi_r := f_r^{-1}: U_r \longrightarrow V_r$$

is a bijective map for every r, and the sets U_r cover S^1 as r varies in the real numbers.



Now let $\varepsilon = \frac{1}{2}$. Every complex number $z \in S^1$ of absolute value 1 can be written

$$z = e^{2\pi ir} = \cos 2\pi r + i\bullet \sin 2\pi r = f(r) \text{ with } -\frac{1}{2} \le r \le +\frac{1}{2}.$$

Therefore,[8]

[8] Note that U_0 contains every point except $e^{\pi i}$ = f(1/2).

$$S^1 = U_0 \bigcup U_{1/2}$$

Assume $p \in U_0 \bigcap U_{1/2}$. and write

$$p = f_0 (x_0) = f_{1/2} (x_{1/2}) \qquad (2)$$

with

$$x_0 \in V_0 = (-\frac{1}{2}, +\frac{1}{2}) \text{ and } x_{1/2} \in V_{1/2} = (0, 1).$$

From (1) we see $x_0 - x_{1/2} \in \mathbb{Z}$, i.e., either[9]

$$x_0 = x_{1/2} \in (0, \frac{1}{2})$$

or

$$x_{1/2} = x_0 + 1 \text{ and } x_0 \in (-\frac{1}{2}, 0)$$

i.e,

$$\varphi_{1/2}(p) = \varphi_0(p) \text{ for } p \in (0, \frac{1}{2}),$$

$$\varphi_{1/2}(p) = \varphi_0(p) + 1 \text{ for } p \in (-\frac{1}{2}, 0),$$

Note that $p = \varphi_0^{-1}(x_0)$. The above identieis write

$$\varphi_{1/2}(\varphi_0^{-1}(x_0)) = x_0 \text{ for } x_0 \in \varphi_0(0, \frac{1}{2}),$$

$$\varphi_{1/2}(\varphi_0^{-1}(x_0)) = x_0 + 1 \text{ for } x_0 \in \varphi_0 (-\frac{1}{2}, 0),$$

Therefore, the map

$$\varphi_{1/2} \circ \varphi_0^{-1} : \varphi_0(U_0 \bigcap U_{1/2}) \longrightarrow \varphi_{1/2}(U_0 \bigcap U_{1/2})$$

is locally a translation by an integer constant, hence analytic. Similarly one proves that the same is true for

$$\varphi_0 \circ \varphi_{1/2}^{-1}: \varphi_{1/2}(U_0 \bigcap U_{1/2}) \longrightarrow \varphi_0(U_0 \bigcap U_{1/2}).$$

Example

To prove that the torus is an analytic manifold, one can identify it with the direct product

$$S^1 \times S^1$$

and prove that the direct product of to analytic manifolds is an analytic manifold.

Example

An alternative method to prove that the 2-sphere

$$S^2 := \{ x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$$

is an analytic manifold, is to use the stereographic projection.

---

[9] Either one of the points is in the intersection $V_0 \bigcap V_{1/2}$ = (0, 1/2) or none is in this intersection. In the first case the other must be also in this intersection (since a shift be ± 1 is outside the union of V_0 and $V_{1/2}$). In the second case one point must be in the left part of V_0 and the other in the right part of $V_{1/2}$.

1.3.5 Diffeomorphisms

Let

$$f: M \longrightarrow M'$$

be a map between C^r-manifolds. This map is called (r times)[10] <u>differentiable</u>, if its descriptions in terms of local coordinate systems gives differentiable maps (between open subsets of euklidian spaces). More precisely, we require:

(i) f is continuous.

(ii) For every point $x \in M$ there are charts $\varphi: U \longrightarrow V$ and $\varphi': U' \longrightarrow V'$ of M and M' respectively, such that the following conditions are satisfied.

1. $x \in U$.
2. $f(U) \subseteq U'$
3. The uniquely determined map $g: V \longrightarrow V'$ ´such that the diagramm

$$\begin{array}{ccc} U & \xrightarrow{f|U} & U' \\ \downarrow\varphi & & \varphi'\downarrow \\ V & \xrightarrow{g} & V' \end{array}$$

is commutative, is (r times) differentiable.

In case $r = \infty$ we say that f is smooth, and in case $r = \omega$, f is analytic.

A <u>diffeomorphism</u> is a bijective map $f: M \longrightarrow M'$ between differentiable manifolds such that f and f^{-1} are differentiable.

Problem 4

Let M and M' differentiable manifolds such that there is a diffeomorphism

$$f: M \longrightarrow M'.$$

Prove that M and M' have equal dimension.
Hint: Use the implicite function theorem.

Problem 5

Decide which ones of the following maps $f: \mathbb{R} \longrightarrow \mathbb{R}$ are diffeomorphisms.

$$f(x) = 2x, x^2, x^3, e^x, e^x + x.$$

1.3.6 One parameter groups of diffeomorphisms

Let M be a differentiable manifold. A <u>one-parameter group of diffeomorphisms</u> on M is a map

$$g: \mathbb{R}\times M \longrightarrow M, (t, x) \mapsto g^t x ,$$

such that the following conditions are satisfied.

(i) g is a differentiable map.

(ii) The map $g^t: M \longrightarrow M, x \mapsto g^t x$, is a diffeomorphism for every $t \in \mathbb{R}$.

(iii) The map $\mathbb{R} \longrightarrow Aut(M), t \mapsto g^t$, is a one-parameter group of transformations of M.

Problem 6

Prove that condition (ii) follows from the other two conditions.

Example

---

[10] $r = 0, 1, 2, \dots , \infty , \omega$.

$$M = \mathbb{R}, g^t(x) = x + vt \; (v \in \mathbb{R}).$$[11]

Remark

Out next aim is the definition of the notion of tangent vector to a manifold at a given point. The problem is, that there is no canonical coordinate system, so that our definition must avoid coordinates.

The idea of the definition is that, for a given vector $X = (X_1, ..., X_n)$ at a given point p there is a derivative in the direction of this vector,

$$X(f) = \frac{df(p+t\bullet X)}{dt}\Big|_{t=0} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)\bullet X_i \,.$$

It turns out,

- these derivatives can be considered quite indepently upon any coordinate system. they are just function such that

  $X(f\bullet g) = f(p)\bullet X(g) + g(p)\bullet X(f)$ if f and g are defined near p

  $X(c) = 0$ if c is a constant near p.

  $X(c\bullet f + d\bullet g) = c\bullet X(f)+d\bullet X(g)$ if c and d are constant near p.

- the vector X is uniquely determined by the derivative $f \mapsto X(f)$, for example, the i-th coordinate with respect to the given coordinate system is $X_i = X(x_i)$.

This allows to identify the vector X with its associated derivative. For example, the unit vector in the direction of the x_i-axis of a given coordinate system will be identified with the derivative

$$\frac{\partial}{\partial x_i}$$

in the direction of this axis. Thus, tangent vectors will be defined to be operators

$$f \mapsto X(f)$$

satisfying the above conditions. The most complicated part of this setting is the description of the domain of definition of these operators, called the local ring of the manifold at the given point.

1.4 Tangent spaces and vector fields

1.4.1 The local ring of a manifold at a given point

Let $p \in M$ be a point on a C^r manifold M. Denote by

$$\mathcal{O}_{M,p}$$

the set of all C^r functions

$$f:U \longrightarrow \mathbb{R}$$

which are defined on an open neightbourhood $U \subseteq M$ of the point x. Different elements f of this set can be defined on different open sets U (but every U should contain the given point p.

We are interested in the behavior of these functions close to the point p and will ignore everything which happens far away from p. Therefore we make the following

---

[11] Condition (i) is trivially satisfied. Therefore it is sufficient to check for condition (iii). One has

$$g^s(g^t(x) = g^s\,(x + vt) = (x+vt) + vs = x + v(s+t) = g^{s+t}(x).$$

Aggreement: two elements of $\mathcal{O}_{M,x}$, say

$$f': U' \longrightarrow \mathbb{R} \text{ and } f'': U'' \longrightarrow \mathbb{R},$$

are considered to be “equal”, if there is an open neightbourhood $U \subseteq M$ of p such that

$$U \subseteq U' \cap U'' \text{ and } f'|_U = f''|_U .$$

In other words, functions which are equal in a neightbourhood of p are considered equal.

This way $\mathcal{O}_{M,x}$ consists of equivalence classes of functions rather than functions.

With this aggreement elements of $\mathcal{O}_{M,p}$ can be added and multiplied. Given to functions, say

$$f': U' \longrightarrow \mathbb{R} \text{ and } f'': U'' \longrightarrow \mathbb{R},$$

the sum f'+f'' is given by the function

$$f'+f'': U' \cap U'' \longrightarrow \mathbb{R}, u \mapsto f'(u) + f''(u),$$

and their product f'•f'' is given by

$$f' \bullet f'': U' \cap U'' \longrightarrow \mathbb{R}, u \mapsto f'(u) \bullet f''(u).$$

With these operations, the set $\mathcal{O}_{M,p}$ is a commutative ring with 1.
This ring is called local ring of M at p, its elements are called germs of functions on M near p.

Example

Let $M = U \subseteq \mathbb{R}^n$ be an open set in euclidian space and consider M as an analytic manifold. To each analytic function defined in a neighbourhood of the point $p \in U$ we associate its power series expansion around p,

$$f \mapsto f(p) + \sum_{i=1}^{n} \frac{\partial f(p)}{\partial x_i}(x_i - p_i) + \sum_{i,j=1}^{n} \frac{\partial^2 f(p)}{\partial x_i \partial x_j}(x_i - p_i)(x_j - p_j) + \text{higher terms}$$

Two analytic functions are mapped to the same power series if and only if they are equal in a neighbourhood of p, i.e., if they are equal as elements of $\mathcal{O}_{M,p}$. In other words,

$$\mathcal{O}_{M,p}$$

can be identified with the ring power series at p.

In the general case $\mathcal{O}_{M,p}$ can be considered as a replacement for the power series ring, even in the case of functions which do not allow any power series expansions.

1.4.2 Tangent vectors on manifolds

Let M be a C^r manifold with $r \geq 2$ and $p \in M$ be some Point. A tangent vektor of M at p is a map (or operator)

$$X: \mathcal{O}_{M,p} \longrightarrow \mathbb{R}$$

such that the following conditions are satisfied.

(i) X is $\mathbb{R}$-linear, i.e.,

$$X(c \bullet f + d \bullet g) = c \bullet X(f) + d \bullet X(g)$$

for arbitrary $c,d \in \mathbb{R}$ and $f,g \in \mathcal{O}_{M,p}$.

(ii) X is a derivation, i.e.

$$X(f\bullet g) = f(p)\bullet X(g) + g(p)\bullet X(f)$$

for arbitrary $f, g \in \mathcal{O}_{M,p}$.

(iii) $X(c) = 0$ für c constant near p.

The set of all tangent vectors of M at p is denoted by

$$T_p(M)$$

and is called tangent space of M at p.

Remarks

(i) The tangent space $T_p(M)$ is obviously a real vector space, since for every two tangent vectors

$$X, Y : \mathcal{O}_{M,p} \longrightarrow \mathbb{R}$$

and every two real numbers c, d the linear combination

$$c\bullet X + d\bullet Y: \mathcal{O}_{M,p} \longrightarrow \mathbb{R}\ , f \mapsto c\bullet X(f) + d\bullet Y(f),$$

satisfies again the above conditions (i) - (iii).

(ii) Let

$$\varphi : U \longrightarrow V \subseteq \mathbb{R}^n$$

be a chart near the given point $p \in M$. The tangent space $T_p(M)$ does not change, if M is replaced by the open set $U \subseteq M$ containing the point p (sind the local ring does not change).

Consider the map[12]

$$\alpha:\mathbb{R}^n \longrightarrow T_p(U) = T_p(M), X = \begin{pmatrix} X_1 \\ \dots \\ X_n \end{pmatrix} \mapsto \sum_{i=1}^{n} X_i\bullet\frac{\partial}{\partial x_i}\Big|_p \qquad (1)$$

mapping each vector $X \in \mathbb{R}^n$ to the associated operator, i.e. to the derivate at p in the direction of X, i.e.,

$$X(f) = \sum_{i=1}^{n} X_i\bullet\frac{\partial(f\circ\varphi)}{\partial x_i}\Big|_p$$

for every smooth funtion f defined near $p \in U$.[13]

---

[12] Below we will see that $\alpha = d_p\varphi$ is just the differential of the chart φ at the point p.

[13] We use the chart φ: $U \longrightarrow V$ to identify U with V and hence identify each function

$$f: U' \longrightarrow \mathbb{R}$$

defined in a neighbourhood $U' \subseteq U$ of p with the corresponding function

$$f\circ\varphi: V' \longrightarrow \mathbb{R}$$

defined in the corresponding neighbourhood $V' := \varphi(U')$ of $\varphi(p)$.

This is an isomorphism of real vector spaces.

Usually we will use this map α to identify the tangent space $T_p(M)$ with euclidian n-space $\mathbb{R}^n$. In particular, we will identify the standard unit vector e_i the the derivative $\frac{\partial}{\partial x_i}$.

(iii) Example. Let

$M = \mathbb{R}^2$

and denote by

$x: M \longrightarrow \mathbb{R}, (u, v) \mapsto u,$

$y: M \longrightarrow \mathbb{R}, (u, v) \mapsto v,$

the two coordinate functions on $\mathbb{R}^2$. The map φ with coordinate functions x and y,

$\varphi = (x,y): M \longrightarrow \mathbb{R}^2, p = (u, v) \mapsto (x(p), y(p)) = (u , v)$

is (global) chart of M. For every point $p \in M$ the tangent space at p is identified with $\mathbb{R}^2$ via the map

$$\mathbb{R}^n \longrightarrow T_p(M), \begin{pmatrix} u \\ v \end{pmatrix} \mapsto u\bullet\frac{\partial}{\partial x}\Big|_p + v\bullet\frac{\partial}{\partial y}\Big|_p ,$$

i.e. the Point $\begin{pmatrix} u \\ v \end{pmatrix}$ is identified with the operator $u\bullet\frac{\partial}{\partial x}\Big|_p + v\bullet\frac{\partial}{\partial y}\Big|_p$. In particular the two standard unit vectors $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are identified with the derivatives at p in the directions of the x- and y-axis, resp. Thus

$$e_1(f) = \frac{\partial f}{\partial x}(p) , e_2(f) = \frac{\partial f}{\partial y}(p)$$

for every differentiable function f defined near p.
Other charts may lead to other identifications.

(iv) Example. Let

$M = \mathbb{R}$

and

$t: M \longrightarrow \mathbb{R}, u \mapsto u,$

the coordinate function on $\mathbb{R}$. Then

$\varphi = t: M \longrightarrow \mathbb{R}$

is a (global) chart of M. For every point $p \in M$ the tangent space at p is identified with $\mathbb{R}$ via the map

$$\mathbb{R} \longrightarrow T_p(M), u \mapsto u\bullet\frac{\partial}{\partial t}\Big|_p$$

In particular, the unit vector 1 of $\mathbb{R}$ at $p \in \mathbb{R}$ is identified with the derivative $\frac{\partial}{\partial t} = \frac{d}{dt}$ at the point p.

Prove of remark (ii).

First of all we see that this map is $\mathbb{R}$-linear: the sum of two vectors is mapped to the sum of the two associated operators, a real multiple of a vector is mapped to the real multiple of the corresponding operator:

$$\alpha(\begin{pmatrix} X_1 \\ \ldots \\ X_n \end{pmatrix} + \begin{pmatrix} X'_1 \\ \ldots \\ X'_n \end{pmatrix}) = \alpha(\begin{pmatrix} X_1+X'_1 \\ \ldots \\ X_n+X'_n \end{pmatrix}) = \sum_{i=1}^{n} (X_i + X'_i)\bullet\frac{\partial}{\partial x_i}|_p$$

$$= \sum_{i=1}^{n} X_i\bullet\frac{\partial}{\partial x_i}|_p + \sum_{i=1}^{n} X'_i\bullet\frac{\partial}{\partial x_i}|_p$$

$$= \alpha(\begin{pmatrix} X_1 \\ \ldots \\ X_n \end{pmatrix}) + \alpha(\begin{pmatrix} X'_1 \\ \ldots \\ X'_n \end{pmatrix})$$

and similarly

$$\alpha(c\bullet\begin{pmatrix} X_1 \\ \ldots \\ X_n \end{pmatrix}) = \alpha(\begin{pmatrix} cX_1 \\ \ldots \\ cX_n \end{pmatrix}) = c\bullet \sum_{i=1}^{n} X_i\bullet\frac{\partial}{\partial x_i}|_p = c\bullet\alpha(\begin{pmatrix} X_1 \\ \ldots \\ X_n \end{pmatrix})$$

Further, the linear map α ist injective, since its kernel is trivial:

$$\alpha(\begin{pmatrix} X_1 \\ \ldots \\ X_n \end{pmatrix}) = 0$$

implies $\sum_{i=1}^{n} X_i\bullet\frac{\partial f}{\partial x_i}|_p = 0$ for every differentiable function f defined near p. In particular, if f is the j-th coordinate function $f = x_j$ we obtain

$$0 = \sum_{i=1}^{n} X_i\bullet\frac{\partial x_j}{\partial x_i}|_p = X_j .$$

Since this is true for every j, we see that $\begin{pmatrix} X_1 \\ \ldots \\ X_n \end{pmatrix} = 0$ is the zero vector.

To prove surjectivity of α it will be sufficient ot prove that the tangent space has at most dimension n,

$$\dim T_p(M) \le n.$$

For this, it will be sufficient to prove that the tangent space is generated by the operators

$$\frac{\partial}{\partial x_i}|_p \text{ with } i = 1, 2, \ldots , n.$$

We will prove more. We will even show that the following identity is true for every element $D \in T_p(M)$:

$$D = D(x_1)\cdot\frac{\partial}{\partial x_1}|_p + ... + D(x_n)\cdot\frac{\partial}{\partial x_n}|_p$$

i.e.,

$$D(f) = D(x_1)\cdot\frac{\partial f}{\partial x_1}(p) + ... + D(x_n)\cdot\frac{\partial f}{\partial x_n}(p) \quad (1)$$

for every C^r-function f (with $r \geq 2$) defined near p.

To simplify notation, we may assume that p is the origin.

$$p = (0, ... , 0) = o.$$

Consider the first order Taylor expansion of f at p = 0.

$$f = c + \sum_{i=1}^{n} f_i x_i + \sum_{i,j=1}^{n} f_{ij} x_i x_j \text{ with } f_i = \frac{\partial f}{\partial x_i}|_{x=p}.$$

Applying the operator D gives

$$D(f) = \sum_{i=1}^{n} (f_i(p)D(x_i) + x_i(p)D(f_i)) + \sum_{i,j=1}^{n} (f_{ij}(p)x_i(p)D(x_j) + x_j(p)D(f_{ij}x_i))$$

Since $x_i(p) = 0$ for every i,

$$D(f) = \sum_{i=1}^{n} f_i(p)D(x_i) = \sum_{i=1}^{n} D(x_i)\frac{\partial f}{\partial x_i}(p) = (\sum_{i=1}^{n} D(x_i)\frac{\partial}{\partial x_i}) (f).$$

But this is the claim.

Convention

For simplicity we will assume below that all manifolds and all maps between manifolds are smooth (rather than being of type C^r), despite of the fact that most constructions will also work for C^r with r sufficiently large.

1.4.3 The differential at a given point

Let M, M' be a C^r manifolds ($r \geq 2$),

$$\varphi: M \longrightarrow M'$$

a C^r map and $p \in M$ be a point. Then the following map is well-defined and $\mathbb{R}$-linear. It is called differential of φ at the point p.

$$d_p\varphi: T_p(M) \longrightarrow T_{\varphi(p)}(M'), X \mapsto \varphi_*(X)$$

such that

$$\varphi_*(X)\,(f) := X(f\circ\varphi) \quad (1)$$

for every germ f on M' near φ(p). Note that the composition

$$f\circ\varphi$$

is a germ[14] on M defined near p, i.e., $X(f\circ\varphi)$ is a well-defined real number. Moreover, the map

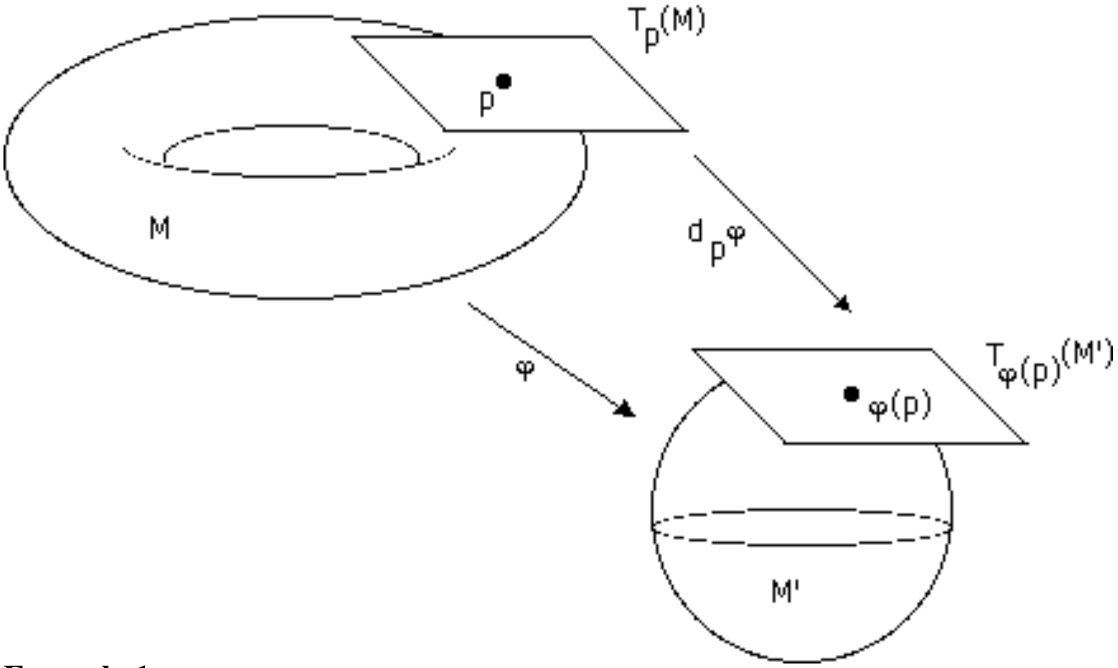
$$\mathcal{O}_{\mathbb{R},\varphi(p)} \longrightarrow \mathbb{R}, f \mapsto X(f\circ\varphi),$$

defined this way, is a tangential vector at φ(p). For, it is obviously $\mathbb{R}$-linear and vanishes on constant functions f. Moreover,

[14] i.e. an element of $\mathcal{O}_{M,p}$.

$$X((f\bullet g)\circ\varphi) = X((f\circ\varphi)\bullet(g\circ\varphi) = f(\varphi(p))X(g\circ\varphi) + g(\varphi(p))X(f\circ\varphi).$$

We have proved, the differential $d_p\varphi$ is well-defined. Its linearity can be directly seen from the definition (1).



Example 1

Let

$$M = M' = \mathbb{R}, p \in M,$$

and

$$\varphi: \mathbb{R} \longrightarrow \mathbb{R}, t \longrightarrow \varphi(t),$$

a C^r map. For every C^r function f defined near $\varphi(p)$, the composition $f\circ\varphi$ is a C^r function defined near p and

$$\frac{\partial(f\circ\varphi)}{\partial t}(p) = \frac{\partial f(t)}{\partial t}(\varphi(p)) \bullet \frac{\partial\varphi(t)}{\partial t}(p)$$

i.e.

$$d_p\varphi(\frac{\partial}{\partial t}|_p)(f) = (\frac{\partial\varphi(t)}{\partial t}(p)\bullet\frac{\partial}{\partial t}|_{\varphi(p)})(f)$$

i.e.,

$$d_p\varphi(\frac{\partial}{\partial t}|_p) = \frac{\partial\varphi(t)}{\partial t}(p)\bullet\frac{\partial}{\partial t}|_{\varphi(p)}$$

The linear map

$$d_p\varphi: \mathbb{R}\bullet\frac{\partial}{\partial t}|_p \longrightarrow \mathbb{R}\bullet\frac{\partial}{\partial t}|_{\varphi(p)}, c\bullet\frac{\partial}{\partial t}|_p \mapsto c\bullet\frac{\partial\varphi(t)}{\partial t}(p)\bullet\frac{\partial}{\partial t}|_{\varphi(p)},$$

is just multiplication by $\frac{\partial\varphi(t)}{\partial t}(p)$. If we identity the two tangent spaces with $\mathbb{R}$, the differential can be written

$$d_p\varphi: \mathbb{R}\longrightarrow\mathbb{R}, c\bullet \mapsto c\bullet\frac{\partial\varphi}{\partial t}(p),$$

To simplify formulas, assume p = 0 is the origin. If φ has Taylor expansion

$$\varphi(t) = \varphi(p) + \frac{\partial\varphi}{\partial t}(p)\bullet t + \frac{1}{2}\frac{\partial^2\varphi}{\partial t^2}(p)\bullet t^2 + \ldots$$

at p, then

$$d_p\varphi(t) = \frac{\partial\varphi}{\partial t}(p)\cdot t,$$

i.e. $d_p\varphi$ is the linear part of the Taylor expansion. On also says that $d_p\varphi$ is the linearization of φ at p.

Example 2

Let

$M = \mathbb{R}^m$

$M' = \mathbb{R}^n$

$p \in M$

and

$$\varphi: \mathbb{R}^m \longrightarrow \mathbb{R}^n, x = \begin{pmatrix} x_1 \\ \ldots \\ x_m \end{pmatrix} \mapsto \begin{pmatrix} \varphi_1(x) \\ \ldots \\ \varphi_n(x) \end{pmatrix},$$

a C^r map. For every C^r function f defined near $\varphi(p)$, the composition $f\circ\varphi$ is a C^r function defined near p and

$$\frac{\partial(f\circ\varphi)}{\partial x_i}(p) = \sum_{j=1}^{n} \frac{\partial f(y)}{\partial y_j}(\varphi(p)) \cdot \frac{\partial\varphi_j(x)}{\partial x_i}(p)$$

i.e.

$$d_p\varphi\left(\frac{\partial}{\partial x_i}\Big|_p\right)(f) = \left(\sum_{j=1}^{n} \frac{\partial\varphi_j(x)}{\partial x_i}(p)\cdot \frac{\partial}{\partial y_j}\Big|_{\varphi(p)}\right)(f)$$

i.e.

$$d_p\varphi\left(\frac{\partial}{\partial x_i}\Big|_p\right) = \left(\sum_{j=1}^{n} \frac{\partial\varphi_j(x)}{\partial x_i}(p)\cdot \frac{\partial}{\partial y_j}\Big|_{\varphi(p)}\right)$$

The linear map

$$d_p\varphi: T_p(\mathbb{R}^m) \longrightarrow T_p(\mathbb{R}^n), \frac{\partial}{\partial x_i}\Big|_p \mapsto \sum_{j=1}^{n} \frac{\partial\varphi_j(x)}{\partial x_i}(p)\cdot \frac{\partial}{\partial y_j}\Big|_{\varphi(p)}$$

maps the i-th basis vector $\frac{\partial}{\partial x_i}\Big|_p$ to the linear combination $\sum_{j=1}^{n} \frac{\partial\varphi_j(x)}{\partial x_i}(p)\cdot \frac{\partial}{\partial y_j}\Big|_{\varphi(p)}$ of the basis vectors $\frac{\partial}{\partial y_j}\Big|_{\varphi(p)}$. If we use the two sets of basis vectors to identifiy the tangent spaces with euclidian spaces the differential can be written as follows.

$$d_p\varphi: \mathbb{R}^m \longrightarrow \mathbb{R}^n, e_i \mapsto \sum_{j=1}^{n} \frac{\partial\varphi_j}{\partial x_i}(p)\cdot e_j,$$

i.e, the i-th standard unit vector e_i is mapped to the i-th column of the Jacobian matrix

$$\frac{\partial\varphi}{\partial x}(p).$$

Equivalently, $d_p\varphi$ is just multiplicatoin by the Jacobian matrix,

$$d_p\varphi: \mathbb{R}^m \longrightarrow \mathbb{R}^n, X = \begin{pmatrix} X_1 \\ \ldots \\ X_m \end{pmatrix} \longrightarrow \frac{\partial\varphi}{\partial x}(p)\cdot X = . \sum_{i=1}^{n} \frac{\partial\varphi}{\partial x_i}(p)\cdot X_i$$

Comparing with the Tayler expansion of φ at p we see again that $d_p\varphi$ is again the linear part of this Taylor expansion (and is therefore also called linearization of φ at p).

Note that the expression on the right is obtained from

$$d\varphi = \sum_{i=1}^{n} \frac{\partial\varphi}{\partial x_i} dx_i$$

when dx_i is replaced with X_i for every i. Thus, $d_p\varphi$ is essentially the same like the total differential of φ at p.

Special case n = 1.
We assume that φ is a map

$$\varphi = \varphi_1 : \mathbb{R}^m \longrightarrow \mathbb{R}$$

Then $d_p\varphi$ is the linear map

$$d_p\varphi: T_p(\mathbb{R}^m) \longrightarrow T_p(\mathbb{R}) = \mathbb{R}, \frac{\partial}{\partial x_i}|_p \mapsto \frac{\partial\varphi}{\partial x_i}(p)$$

which maps the i-th generator of the tangent space to the i-th derivative of φ at p, for short

$$d_p\varphi: T_p(\mathbb{R}^m) \longrightarrow \mathbb{R}, \frac{\partial}{\partial x_i} \mapsto \frac{\partial\varphi}{\partial x_i}.$$

Note that $d_p\varphi$ is a linear functional on the tangent space $T_p(\mathbb{R}^m)$, i.e., an element of the dual space,

$$d_p\varphi \in T_p(\mathbb{R}^m)^*.$$

Consider the case that φ is the j-th coordinate function,

$$\varphi = x_j : \mathbb{R}^m \longrightarrow \mathbb{R}, \begin{pmatrix} u_1 \\ \dots \\ u_m \end{pmatrix} \mapsto u_j ,$$

we get

$$d_p x_j (\frac{\partial}{\partial x_i}) = \delta_{ij} .$$

We see that the differentials

$$d_p x_1 , \dots , d_p x_m T_p(\mathbb{R}^m)^*$$

form a dual base of the base

$$\frac{\partial}{\partial x_1}, \dots , \frac{\partial}{\partial x_m} \in T_p(\mathbb{R}^m).$$

The dual tangent space $T_p(M)^*$ to a manifold M at a point p is also called cotangent space of M at p. Its elements are called cotangent vectors or covectors to M at p.[15]

Example 3

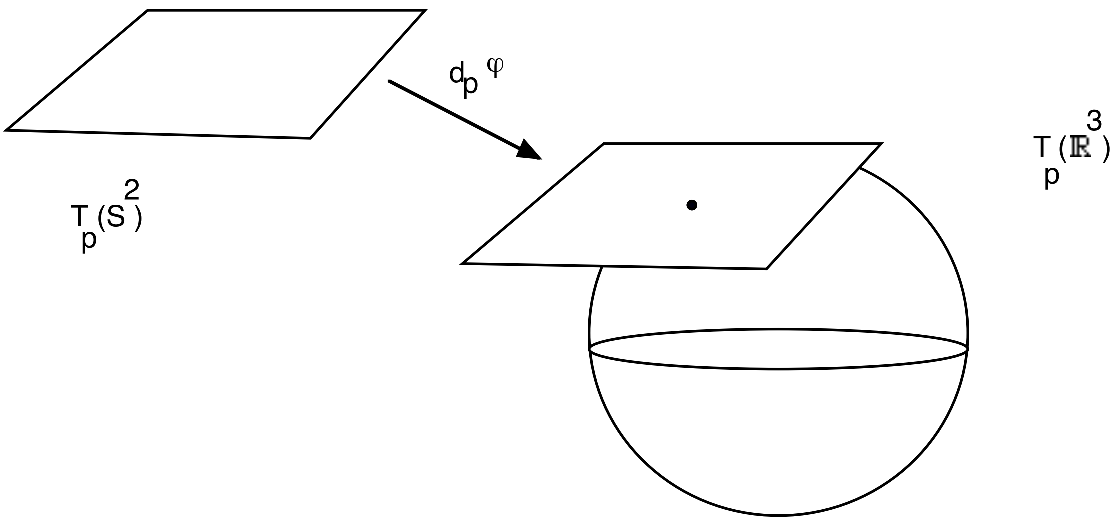
Consider the 2-sphere in 3-space,

$$S^2 := \{ x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid F(x) = 0 \} , F(x) = x_1^2+x_2^2+x_3^2$$

[15] Vectors have an invariant meaning, their coordinates are contravariant tensors. Similarly, covectors have an invariant meaning, their coordinates are covariant tensors.

and let φ be the natural inclusion

$$\varphi\colon S^2 \longrightarrow \mathbb{R}^3,\ x \mapsto x.$$



Then for any smooth function $f \in \mathcal{O}_{\mathbb{R}^3,p}$ defined on $\mathbb{R}^3$ near p, one has

$$d_p\varphi(X)(f) = X(f\circ\varphi) = X(f|_{S^2})$$

Let f := F. Then $f|_{S^2}$ is identically zero on S^2. Since tangent vectors are linear operators,

$$d_p\varphi(X)(F) = 0 \text{ for every tangent vector } X \in T_p(X).$$

For every vector

$$Y = (Y_1, Y_2, Y_3) \in \mathrm{Im}(d_p\varphi(X))$$

one has

$$0 = Y(F) = \sum_{i=1}^{3} \frac{\partial F(p)}{\partial x_i} X_i = 2p_1X_i + 2p_2X_2 + 2p_3X_3$$

hence

$$0 = p_1X_i + p_2X_2 + p_3X_3.$$

This is just the usual equation of the tangent space to S^2 at the point p (where the point p is considered to be the origin). In the coordinates of $\mathbb{R}^2$ (with the origin at (0, ... ,0)) we obtain

$$0 = p_1(X_i-p_1) + p_2(X_2-p_2) + p_3(X_3-p_3).$$

A similar argument also works for submanifolds in $\mathbb{R}^n$ defined by more than one equation. More precisely, if

$$M \subseteq \mathbb{R}^n$$

is defined by the equations

$$f_1(x_1, \dots, x_n) = \dots = f_m(x_1, \dots, x_n) = 0$$

that the tangent space $T_p(M) \subseteq T_p(\mathbb{R}^n) = \mathbb{R}^n$ of M at $p \in M$ is defined by the linear equations

$$d_pf_1(X_1, \dots, X_n) = \dots = d_pf_m(X_1, \dots, X_n) = 0$$

Example 4

Let $I = (a, b) \subseteq \mathbb{R}$ a none empty open intervall and

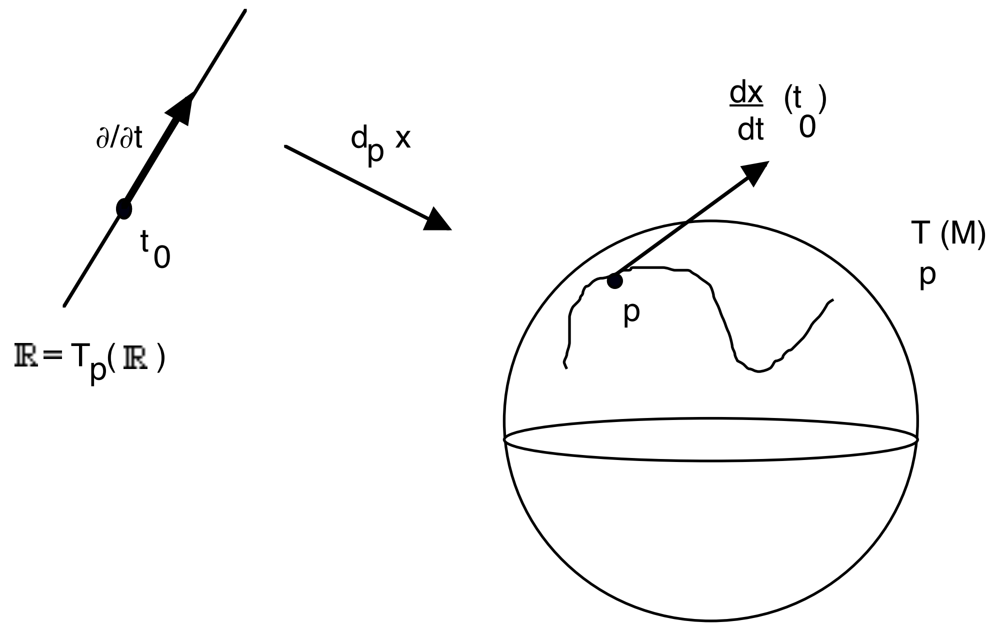
$$x: I \longrightarrow M, t \mapsto x(t),$$

a C^r map of C^r In that what follows, such a map will be also called a C^r curve. For every t_0 we write

$$\frac{dx}{dt}(t_0) = (d_{t_0} x)(\frac{\partial}{\partial t}|_{t_0}) \in T_pM, \quad p := x(t_0),$$

for short

$$\frac{dx}{dt} = dx\,(\frac{\partial}{\partial t}).$$



This tangent vector is also called the derivative of the curve x at t_0 or the tangent vector of the curve x at $x(t_0)$. Note that $d_{t_0} x$ is a linear map

$$d_{t_0} x: T_{t_0}(I) \longrightarrow T_{x(t_0)}M$$

and $\frac{\partial}{\partial t}|_{t_0}$ is a tangent vector at t_0,

$$\frac{\partial}{\partial t}|_{t_0} \in T_{t_0}(I).$$

In case, M is an open set in euclidian space, say

$$M \subseteq \mathbb{R}^n \text{ open},$$

we obtain for every C^r function $\varphi \in \mathcal{O}_{M,x(t_0)}$ defined near $x(t_0)$:

$$(d_{t_0} x)(\frac{\partial}{\partial t}|_{t_0})(\varphi) = \frac{\partial}{\partial t}|_{t_0}(\varphi \circ x) = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial x_i}(x(t_0)) \bullet \frac{\partial x_i}{\partial t}(t_0) = (\sum_{i=1}^{n} \frac{\partial x_i}{\partial t}(t_0) \bullet \frac{\partial}{\partial x_i}|_{x(t_0)})(\varphi).$$

This holds for every φ, hence

$$\frac{dx}{dt}(t_0) = (d_{t_0}x)(\frac{\partial}{\partial t}|_{t_0}) = \sum_{i=1}^{n} \frac{\partial x_i}{\partial t}(t_0)\bullet\frac{\partial}{\partial x_i}|_{x(t_0)}$$

If we identify $T_{x(t_0)}M$ in the usual way with $\mathbb{R}^n$ (idenitying $\frac{\partial}{\partial x_i}|_{x(t_0)}$ with the i-th standard unit vector), we obtain

$$\frac{dx}{dt}(t_0) = \sum_{i=1}^{n} \frac{\partial x_i}{\partial t}(t_0)\bullet e_i = \begin{pmatrix} dx_1/dt \\ \dots \\ dx_n/dt \end{pmatrix}(t_0),$$

We have proved, the above definition von $\frac{dx}{dt}$ coincides with the usual one, if M is an open subset in euclidial space.

Example 5

Consider the differential equation

$$\frac{dx}{dt} = X(x, t),\ x(t) \in \mathbb{R},\ X(x,t) \in \mathbb{R}.$$

and let

$$\{ (t, x(t))\}$$

be an integral curve through a given point

$$(t_0, x_0) = (t_0 , x(t_0)).$$

The differential equation tells us that the tangent line to the given integral curve in the given point has slople $X(x_0,t_0)$, i.e., is given by the equation

$$x - x_0 = X(x_0,t_0)(t - t_0).$$

We see that for every $p = (x_0,t_0)$ the vector

$$\begin{pmatrix} 1 \\ X(p) \end{pmatrix} = \frac{\partial}{\partial t}|_p + X(p)\bullet\frac{\partial}{\partial x}|_p$$

is a tangent vector at (x,t) to the integral curves going throug p. Now consider the coordinate functions

$$t: \mathbb{R}^2 \longrightarrow \mathbb{R},\ \begin{pmatrix} u \\ v \end{pmatrix} \mapsto u$$

$$x: \mathbb{R}^2 \longrightarrow \mathbb{R},\ \begin{pmatrix} u \\ v \end{pmatrix} \mapsto v$$

The values of their differentials at this tangent vector are

$$d_p t \begin{pmatrix} 1 \\ X(p) \end{pmatrix} = 1$$

$$d_p x \begin{pmatrix} 1 \\ X(p) \end{pmatrix} = X(p)$$

hence

$$d_p x = X(p) d_p t \text{ for ever point } p$$

or

$$dx = X(p)\bullet dt.$$

In other words, the differential equation translates into a relation between differential forms, the left hand side of the differential equation can be considered as the quotient of the two functions ‘dx’ and ‘dt’.

For example, the differential equation

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

translates into the relation

$$\frac{1}{f(x)}\bullet dx = g(t)\bullet dt.$$

To understand why the calculations used to solve this differential equation, we have to learn how to integrate differential forms.

1.4.4 Inverse function theorem[16]

Let

$$\varphi: M \longrightarrow M'$$

be a C^r map of C^r-manifolds ($r \geq 2$) and $p \in M$ a point such that the linear map

$$d_p\varphi: T_p(M) \longrightarrow T_{\varphi(p)}(M')$$

is bijective. Then there are neighbourhoods $U \subseteq M$ and $U' \subseteq M'$ of p and $\varphi(p)$ such that

1. $\varphi(U) = U'$
2. $\varphi|_U: U \longrightarrow U'$ is a diffeomorphism.

In particular, the inverse $\varphi|_U^{-1}: U' \longrightarrow U$ exists and is a C^r map.

Example 1

Consider the exponential map

$$\varphi: \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}, t \mapsto e^t ,$$

Since its derivative $\frac{d\varphi}{dt}(t) = e^t$ is always > 0, this function is strongly monotone, hence has an inverse,

$$\varphi^{-1}: \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}, t \mapsto \log(t).$$

Claim: this inverse is a C^r function for every r. To see this, consider the differential of φ at a given point $t = u$. It is given by

$$d_u\varphi: \mathbb{R} \longrightarrow \mathbb{R}, t \mapsto (e^u)\bullet t.$$

Since $e^u \neq 0$, this is an isomorphisms of vector spaces. Therefore, the exponential map is a local diffeomorphism at every point, i.e. its inverse is a C^r function for every r.

Example 2

Consider the map

[16] This theorem will be soon formulated and proved in the analysis course running parallel to these lectures. Below we present a translation into the language of manifolds.

$$\varphi: \mathbb{R} \longrightarrow (-1, +1), t \mapsto \sin t ,$$

Its derivate satisfies

$$\frac{d\varphi}{dt}(t) = \cos t > 0 \text{ for } -\pi/2 < t < +\pi/2,$$

hence is strongly monotone in this interval, i.e. there is an inverse

$$\varphi^{-1}: (-1, +1) \longrightarrow (-\pi/2, +\pi/2), t \mapsto \text{arc cos } t.$$

Claim: this inverse is a C^r function for every r. To see this, consider the differential of φ at a given point $t = u \in (-\pi/2, +\pi/2)$. Is is given by

$$d_u\varphi: \mathbb{R} \longrightarrow \mathbb{R}, t \mapsto (\cos u)\bullet t..$$

(1)

Since $\cos u \neq 0$, this is an isomorphism of vector spaces. Therefore, sin is a local diffeomorphism at every point of the intervall $(-\pi/2, +\pi/2)$, i.e. the inverse (1) is a C^r function for every r.

1.4.5 Implicite function theorem[17]

Let

$$\varphi: M \longrightarrow M'$$

be a C^r map of C^r manifolds ($r \geq 2$) and

$$p \in M, p' \in M', \varphi(p) = p',$$

be points such that

$$d_p\varphi: T_p(M) \longrightarrow T_{p'}(M')$$

is surjective. Then there is an open neightbourhood $U \subseteq M$ such that

$$U \cap \varphi^{-1}(p')$$

is a <u>submanifold</u> of U (i.e. is locally given by a system of linear equations for approbriately chosen coordinate systems). In particular, this intersection is a C^r manifold.

Example

Let

$$M = \mathbb{R}^3$$
$$M' = \mathbb{R}$$
$$\varphi: \mathbb{R}^3 \longrightarrow \mathbb{R}, (x, y, z) \mapsto x^2 - y^2 - z^2 - 1.$$

Then

$$S^1 := \varphi^{-1}(0) = \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 - y^2 - z^2 = 1\}$$

is the unit sphere. For $p = (u, v, w) \in \mathbb{R}^3$ the differential of φ at p is given by

$$d_p\varphi: \mathbb{R}^3 \longrightarrow \mathbb{R}, (X, Y, Z) \mapsto 2u\bullet X + 2v\bullet Y + 2w\bullet Z, \qquad (1)$$

For every $p \in N$ at least one coodinate is none-zero, i.e., the differential (1) is not the zero map, and hence is surjective for every $p \in S^1$. By the implicite function theorem, for every point $p \in S^1$ there is an open set U,

$$p \in U \subseteq \mathbb{R}^3$$

[17] ibid.

such that $S^1 \cap U$ is submanifold of U. Since this is true for every p, the unit sphere is a C^r manifold (for every r).

Problem: how to find the charts of this manifold.

To illustrate how to find charts of S^1 let us construct a charts which identifies the northern hemissphere with the open unit disc in the plane. To be precise, consider the map

$$\psi: U \longrightarrow V, (x, y, z) \mapsto (x, y),$$

where

$$U := \{ (x,, y, z) \in S^1 \mid z > 0\}$$

is the upper hemisphere and

$$V := \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$$

is the open unit disc in the plane.

We claim that this is a chart. The map is easily seen to be bijective. We have to prove that is diffeomorphisms (i.e. the map and its invers are C^r maps). By the inverse function theorem it is sufficient to show that the differential

$$d_p\psi: \mathbb{R}^2 = T_p(U) \longrightarrow T_p(V) = \mathbb{R}^2$$

is an isomorphism for every $p \in U$. To prove this we may replace U with S^1, V with $\mathbb{R}^2$ and factor ψ over $\mathbb{R}^3$:

$$\psi: S^1 \overset{i}{\rightarrowtail} \mathbb{R}^3 \overset{pr}{\twoheadrightarrow} \mathbb{R}^2, (x, y, z) \mapsto (x, y, z) \mapsto (x, y),$$

i.e. i is the natural inclusion and pr is the orthogonal projection of the 3-space to the plane. The maps ψ, i and pr are obviously linear, hence coincide with their respective linearizations:

$$d_p \psi: \mathbb{R}^2 = T_p(S^1) \overset{d_p(i)}{\longrightarrow} \mathbb{R}^3 \overset{d_p(pr)}{\longrightarrow} \mathbb{R}^2, (X, Y, Z) \mapsto (X, Y, Z) \mapsto (X, Y).$$

To prove that this is an isomorphism, it is sufficient to prove injectivity.[18] Thus it will be sufficient to show that the kernel of this map is trivial. Note that

$$\mathrm{Ker}(d_p \psi) = \{ (X, Y, Z) \in T_p(S^1) \mid X = Y = 0\}.$$

Since S^1 is defined by the equation $\varphi(x, y, z) = 0$, the tangent space at p is defined by

$$T_p(S^1): 0 = d_p\varphi (X, Y, Z) = 2u\bullet X + 2v\bullet Y + 2w\bullet Z.$$

Therefore

$$\mathrm{Ker}(d_p \psi) = \{ (X, Y, Z) \in \mathbb{R}^3 \mid 0 = X = Y = 2u\bullet X + 2v\bullet Y + 2w\bullet Z \}$$
$$= \{ (X, Y, Z) \in \mathbb{R}^3 \mid 0 = X = Y = 2w\bullet Z \}$$

For $p = (u, v, w) \in U$ in the upper hemissphere one has $w > 0$, i.e.

$$\mathrm{Ker}(d_p \psi) = \{ (X, Y, Z) \in \mathbb{R}^3 \mid 0 = X = Y = Z \} = \{(0,0,0)\}.$$

The kernel is trivial. We have proved ψ is a chart identifying the upper hemis sphere with the unit disc.

1.4.6 Vector fields and differential equations

Let M be a C^r manifold ($r \geq 2$) and write

---

[18] Since this is a linear map of vector space with equal dimension 2.

$$T(M) := \vee_{p \in M} T_p(M)$$

for the disjoint union of all tangent spaces to M. This disjoint union is called tangent bundle[19] of M. A vector field on M is a map

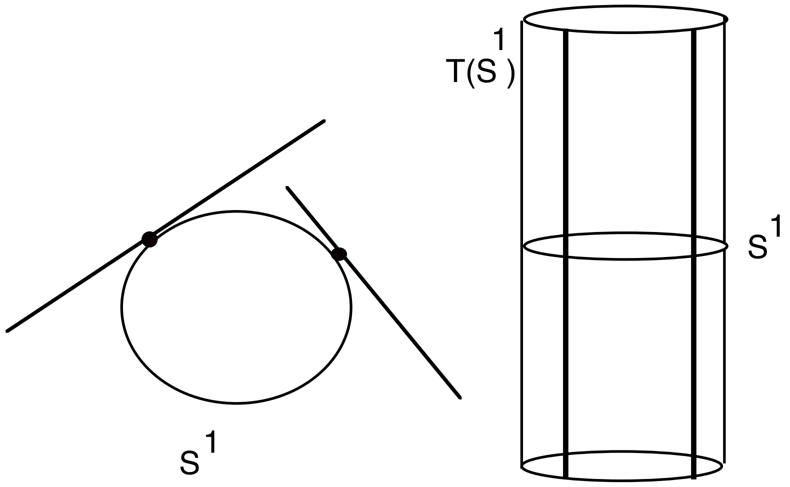
$$X: M \longrightarrow T(M), p \mapsto X_p ,$$

such that $X_p \in T_p(X)$, i.e. to every point p of M one associates a tangent vector X_p of M at p.

Example 1

If $M = S^1$, the tangent bundle can be identified with the cylinder over S^1,

$$TM = S^1 \times \mathbb{R} .$$



Similarly, if $M = S^1 \times S^1$ is a torus, the tangent bundle can be identified with

$$TM = M \times \mathbb{R}^n.$$

Note that the unit circle

$$S^1 = \{ z \in \mathbb{C} \mid |z| = 1 \}$$

is a group with respect to complex multiplication. Hence the torus is a group, too. In general, if the manifold is a Lie group with neutral element e∈M, the tangent bundle can be identified with the direct product

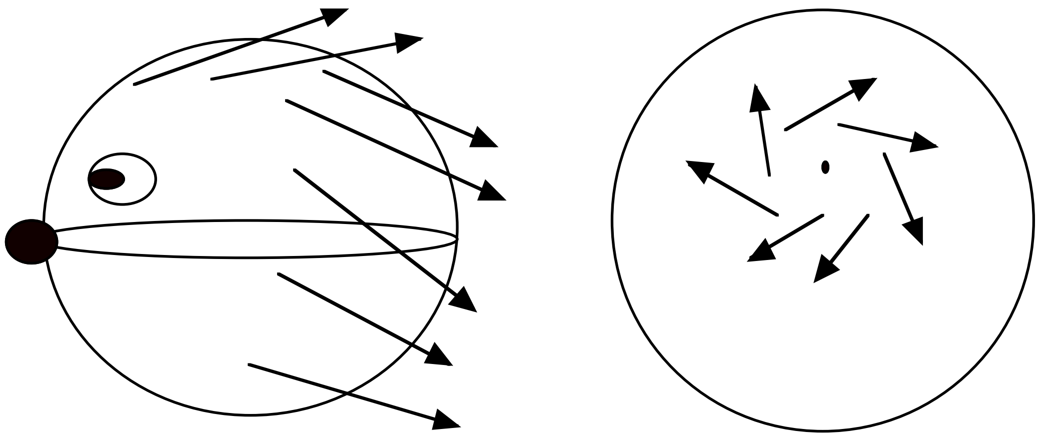
$$TM = M \times T_e(M).$$

If $M = S^2$, the tangent bundle cannot be identified with $S^2 \times \mathbb{R}^2$,

$$M \neq S^2 \times \mathbb{R}^2.$$

This is a corollary of a deep theorem, the hedgehog theorem: every continuous vector field on the 2-sphere has at least one zero. “There is no haircut for a hedgehog: there will always remain a whirl”.

[19] One can prove that T(M) is a C^{r-1} manifold.



Example 2

Let $M = \mathbb{R}^n$ and

$$x_i: M \longrightarrow \mathbb{R},, \begin{pmatrix} u_1 \\ \dots \\ u_n \end{pmatrix} \mapsto u_i ,$$

the i-th coordinate function. Then

$$\frac{\partial}{\partial x_i}: M \longrightarrow T(M), p \mapsto \frac{\partial}{\partial x_i}\Big|_p .$$

is a vector field associating to every point $p \in M$ the i-th standard unit vector at. It is called the i-the standard unit vektor field.

Example 3

Let

$$x: U \longrightarrow V \subseteq \mathbb{R}^n, u \mapsto \begin{pmatrix} x_1(u) \\ \dots \\ x_n(u) \end{pmatrix}$$

be a chart of the C^r manifold M ($r \geq 2$). Then, for $i = 1, \dots, x$, one have a vector field on U,

$$\frac{\partial}{\partial x_i}: U \longrightarrow T(U), u \mapsto \frac{\partial}{\partial x_i}\Big|_p ,$$

associating with each point $p \in U$ the unit vector at p in the direction of the i-th coordinate axis. It is called the i-th standard unit vector field of this coordinate system.

Remarks

(i) Let

$$X: M \longrightarrow T(M)$$

be a vector field on a C^r manifold and

$$x: U \longrightarrow V \subseteq \mathbb{R}^n, u \mapsto \begin{pmatrix} x_1(u) \\ \dots \\ x_n(u) \end{pmatrix}$$

be a chart of M. For every $p \in U$,

$$X_p \in T_p(X) = \mathbb{R}\cdot \frac{\partial}{\partial x_1}\Big|_p + \dots + \mathbb{R}\cdot \frac{\partial}{\partial x_n}\Big|_p .$$

Since the vectors $\frac{\partial}{\partial x_i}\Big|_p$ form a basis of $T_p(X)$,

$$X_p = f_1(p)\cdot \frac{\partial}{\partial x_1}\Big|_p + \dots + f_n(p)\cdot \frac{\partial}{\partial x_n}\Big|_p$$

with uniquely determined real numbers $f_i(p) \in \mathbb{R}$. Thus the restriction of the vector field X to U can be written

$$X|_U = f_1 \cdot \frac{\partial}{\partial x_1} + \dots + f_n \cdot \frac{\partial}{\partial x_n}$$

with uniquely determined functions

$$f_1, \dots, f_n : U \longrightarrow \mathbb{R},$$

which are called the coordinate functions of the vector field X with respect to the given chart x: U ⟶ V. Note that, since the covectors $d_p x_j$ form a dual basis,

$$d_p x_j(X_p) = f_i(p) = X_p(x_i)$$

(ii) The vector field X is called a C^s vektor field at p ∈ U (for s < r), if its coordinate functions f_i are s times continuously differentialble at p (analytic at p in case s = r = ω), i.e. if

$$U \longrightarrow \mathbb{R}, u \mapsto X_u(x_i),$$

is a C^s function at p for i = 1, ... , n. This is equivalent to the condition that[20]

$$U \longrightarrow \mathbb{R}, u \mapsto X_u(f),$$

is a C^s function at p for every C^{s+1} function f: U ⟶ ℝ (the latter condition being independent upon the choise of the chart x: U ⟶ V. A C^s vector field on M is a vector field on M which is C^s at every point of M. In case s = 0 one obtains the notion of continuous vector field, in case r = ∞ one obtains the one of smooth vector field and in case r = ω the notion of analytic vector field.

(iii) We are now ready to define the notion of differential equation in the context of manifolds. Let

$$X: M \longrightarrow T(M)$$

be a C^s vector field on the C^r manifold M (s < r). A solution of the differential equation

$$\frac{dx}{dt} = X \qquad (1)$$

is a C^{s+1} curve

$$x: I \longrightarrow M, t \mapsto x(t), \qquad (2)$$

[20] The condition is obviously sufficient: if it holds for every f, it also holds for f = x_1 , ... , x_n. Lets prove that it is necessary. For every C^{s+1}-function f: U ⟶ ℝ one has

$$X_u(f) = \Big(\sum_{i=1}^{n} f_i(u)\cdot \frac{\partial}{\partial x_i}\Big|_u\Big)(f) = \sum_{i=1}^{n} f_i(u)\cdot \frac{\partial f}{\partial x_i}(u)$$

But this is a C^s function of u , provided so is $f_i(u) = X_u(x_i)$ for every i.

such that

$$\frac{dx}{dt}(t) = X(x(t)).$$

for every $t \in I$. If $x(t_0) = x_0$ for some $t_0 \in I$ and some $x_0 \in M$, the map (2) is also called a solution of the initial value problem

$$\frac{dx}{dt} = X, \qquad x(t_0) = x_0 \tag{3}$$

Differential equations of type (1), where the vector field X does not depent upen t, are called autonomic.

(iv) The vector C^s field

$$X: M \longrightarrow T(M), x \mapsto X(x),$$

on the phase space defines a vector field

$$\tilde{X}: M \times \mathbb{R} \longrightarrow T(M \times \mathbb{R}) = T(M) \times \mathbb{R}, (x, t) \longrightarrow (X(x), 1),$$

on the extended phase space which is called direction field associated with X. Every solution

$$x: I \longrightarrow M,, t, \mapsto x(t),$$

of the differential equation (1), defines a curve in the extended phase space,

$$\tilde{x}: I \longrightarrow M \times \mathbb{R}, t \mapsto (x(t), t),$$

(whose image is the associated integral curve). The differential equation (1) is equivalent to the differential equation

$$\frac{d\tilde{x}}{dt} = (\frac{dx}{dt}, 1) = (X(x), 1) = \tilde{X}.$$

The initial value problem (3) is equivalent to the initial value problem

$$\frac{d\tilde{x}}{dt} = \tilde{X}, \tilde{x}(t_0) = (x(t_0), 1).$$

(v) Directional fields have the advandage that they have no zeroes.

(vi) The use of the extended phase space has the advandage that one can easily replace the vector field $X: M \longrightarrow TM$ by a time dependent vector field

$$X: M \times I \longrightarrow T(M), (x, t) \mapsto X(x,t),$$

($I \subseteq \mathbb{R}$ open) without changing anything in the above calculations. This way, none-autonomic differential equations on the phase space become autonomic differential equation on the extended phase space.

1.4.7 Differential forms

Let M be a C^r manifold ($r \geq 2$) and write

$$T(M)^* := \vee_{p \in M} T_p(M)^*$$

for the disjoint union of all dual tangent spaces to M (the cotangent spaces). This disjoint union is called cotangent bundle[21] of M. A differential form on M is a map

$$\omega: M \longrightarrow T(M), p \mapsto \omega_p ,$$

[21] One can prove that T(M) is a C^{r-1} manifold.

such that $\omega_p \in T_p(X)^*$, i.e. to every point p of M one associates a covector ω_p of M at p.

Remarks

(i) Let

$$\omega: M \longrightarrow T(M)^*, p \mapsto \omega_p ,$$

be a differential on a C^r manifold and

$$x: U \longrightarrow V \subseteq \mathbb{R}^n, u \mapsto \begin{pmatrix} x_1(u) \\ ... \\ x_n(u) \end{pmatrix}$$

be a chart of M. For every $p \in U$,

$$\omega_p \in T_p(X)^* = \mathbb{R}\bullet d_p x_1 + ... + \mathbb{R}\bullet d_p x_n .$$

Since the covectors $d_p x_i$ form a basis of the cotangential space,

$$\omega_p = f_1(p)\bullet d_p x_1 + ... + f_n(p)\bullet d_p x_n$$

with uniquely determined real numbers $f_i(p) \in \mathbb{R}$. Thus the restriction of the differential form ω to U can be written

$$\omega |_U = f_1 \bullet dx_1 + ... + f_n \bullet dx_n$$

with uniquely determined functions

$$f_1 , ... , f_n : U \longrightarrow \mathbb{R},$$

which are called the <u>coordinate functions</u> of the differential form ω with respect to the given chart x: U ⟶ V. Note that, since the vectors $\frac{\partial}{\partial x_j}|_p$ form basis of the tangent space at p which is dual to the basis $\{d_p x_i\}_{i=1,...,n}$,

$$\omega_p(\frac{\partial}{\partial x_j}|_p) = f_i(p).$$

(ii) The differential form ω is called a C^s <u>differential form</u> at $p \in U$ (for $s < r$), if its coordinate functions f_i are s times continuously differentialble at p (analytic at p in case $s = r = \omega$), i.e. if

$$U \longrightarrow \mathbb{R}, u \mapsto \omega_p(\frac{\partial}{\partial x_j}|_p),$$

is a C^s function at p for $i = 1, ... , n$. This is equivalent to the condition that[22]

---

[22] The condition is obviously sufficient: if it holds for every X, it also holds for

$$X = \frac{\partial}{\partial x_i}|_p \quad i = 1, ... , n.$$

Lets prove that it is necessary. Let

$$X = \sum_{i=1}^{n} f_i \bullet \frac{\partial}{\partial x_i}$$

be a vector field on U, which is of class C^s at p, i.e., each f_i is a C^s function at p. Then

$$U \longrightarrow \mathbb{R}, u \mapsto \omega_p(X_p),$$

is a C^s function at p for every vector field X on U, which is of class C^s at p. This latter condition is independent upon the choise of the chart x, i.e. the notion of C^s differential form of class C^s has an invariant meaning. In casse $s = 0$, one obtains the notion of continuous differential form. In case $s = r = \infty$, the differential form is called smooth and in case $r = s = \omega$ it is called analytic.

(iii) Let

$$\varphi: M \longrightarrow \mathbb{R}$$

be a C^s function on M. Then $d_p\varphi \in T_p(M)^*$, hence as above

$$d_p\varphi = f_1(p)\bullet d_px_1 + ... + f_n(p)\bullet d_px_n$$

for ever $p\in U$ and

$$d\varphi = f_1(p)\bullet d_px_1 + ... + f_n(p)\bullet d_px_n$$

with uniquely determined coordinate functions f_i: U $\longrightarrow$ R. By Example 2 of 1.4.3 (Special case $n = 1$) we see

$$\frac{\partial\varphi}{\partial x_i}(p) = d_p\varphi\left(\frac{\partial}{\partial x_i}|_p\right) = f_i(p),$$

hence

$$d\varphi = \frac{\partial\varphi}{\partial x_i}\bullet d_px_1 + ... + \frac{\partial\varphi}{\partial x_i}\bullet d_px_n$$

in particular, the map

$$d\varphi: M \longrightarrow T(M)^*, p \mapsto d_p\varphi,$$

is a differential form of class C^{s-1}. Differential forms ω of this type, i.e. such that

$$\omega = d\varphi \text{ for some } \varphi$$

are called exact.

1.4.8 Integration

Let

$$\omega: M \longrightarrow T(M)^*$$

be a continuous differential form on a C^r manifold ($r \geq 2$) and

$$\gamma: I \longrightarrow M, I = [a, b] \subseteq R$$

a piecewise continuously differentiable curve. Then in terms of a local coordinate system

$$x: U \longrightarrow V \subseteq \mathbb{R}^n$$

one can write

$$\omega_p(X_p) = \sum_{i=1}^{n} f_i(p)\bullet \omega_p\left(\frac{\partial}{\partial x_i}|_p\right).$$

and, as a function of p, the left hand side is of class C^s, provided so is $\omega_p\left(\frac{\partial}{\partial x_i}|_p\right)$ for every i.

$$\gamma(t) = \begin{pmatrix} \gamma_1(t) \\ \dots \\ \gamma_n(t) \end{pmatrix}$$

and

$$\frac{d\gamma}{dt} = d\gamma\,(\frac{\partial}{\partial t}) = \begin{pmatrix} d\gamma_1(t)/dt \\ \dots \\ d\gamma_n(t)/dt \end{pmatrix} = \sum_{i=1}^{n} \frac{d\gamma_i}{dt}(t) \bullet \frac{\partial}{\partial x_i}|_{\gamma(t)} \in T_{\gamma(t)}M$$

In particular, the coordinate functions of $\frac{d\gamma}{dt}$ are piecewise continuous functions of t. The coordinate functions of ω,

$$\omega = \sum_{j=1}^{n} f_j(p)\, dx_j \in T_p(M)^*$$

are continuous functions of p. Therefore,

$$\begin{aligned} \omega(\frac{d\gamma}{dt}(t)) &=^{23} (\sum_{j=1}^{n} f_j(\gamma(t))\, dx_j)\, (\frac{d\gamma}{dt}(t)) \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} f_j(\gamma(t))\, \frac{d\gamma_i}{dt}(t)\, dx_j(\frac{\partial}{\partial x_i}|_{\gamma(t)}) \\ &= \sum_{i=1}^{n} f_i(\gamma(t))\, \frac{d\gamma_i}{dt}(t) \end{aligned}$$

is piecewise continuous as a function of t,

$$I \longrightarrow \mathbb{R},\ t \mapsto \omega(\frac{d\gamma}{dt}|_t).$$

Hence the integral

$$\int_\gamma \omega := \int_a^b \omega(\frac{d\gamma}{dt}|_t)\, dt$$

is well defined. It is called <u>integral over the differential form</u> ω along the curve γ.

Example

Let $M = \mathbb{R}$ and

$$\omega = f(x)dx,$$

where f is a continuous function

$$f\colon M \longrightarrow \mathbb{R}.$$

Moveover, let

$$\gamma\colon [a, b] \longrightarrow M,\ t \mapsto t,$$

be the identity map. Then

$$\omega(\frac{d\gamma}{dt}|_t) = f(\gamma(t))\, dx\, (d\gamma\, (\frac{\partial}{\partial t})) =^{24} f(\gamma(t))\, dx\, (\frac{\partial}{\partial x}) = f(t)$$

---

[23] $\frac{d\gamma}{dt} = \frac{d\gamma}{dt}(t) \in T_{\gamma(t)}(M)$.

[24] For every function $\varphi = \varphi(x)$ defined near $\gamma(t)$ one has

$$d\gamma\, (\frac{\partial}{\partial t}|_t)(\varphi) = \frac{\partial}{\partial t}|_t(\varphi\circ\gamma) = \frac{\partial\varphi}{\partial x}|_{\gamma(t)} \bullet \frac{\partial\gamma}{\partial t}(t) = \frac{\partial\varphi}{\partial x} = \frac{\partial}{\partial x}|_{\gamma(t)}\,(\varphi),$$

hence

$$d\gamma\, (\frac{\partial}{\partial t}|_t) = \frac{\partial}{\partial x}|_{\gamma(t)}$$

Hence, by definition

$$\int_\gamma \omega = \int_a^b f(t)dt$$

Remark
Let

$$\varphi: [c, d] \longrightarrow [a, b]$$

be a condinous map, which induces a diffeomorphisms $(c, d) \longrightarrow (a, b)$. Then

$$\int_\gamma \omega = \int_{\gamma\circ\varphi} \omega,$$

i.e., the integral does not depent upon the parametrization of the curve γ. To prove this, we may assume that γ is of class C^1 everywhere on the open intervall (a, b).[25] Then

$$\int_{\gamma\circ\varphi} \omega, = \int_c^d \omega(\frac{d\gamma\circ\varphi}{dt}|_t)\ dt$$

By definition, for every C^1 function α defined near $\gamma(\varphi(t))$ one has

$$\begin{aligned}
\frac{d\gamma\circ\varphi}{dt}(\alpha) &= d(\gamma\circ\varphi)(\frac{\partial}{\partial t})(\alpha) && \text{(definition of } \tfrac{d}{dt}\text{)}\\
&= \frac{\partial}{\partial t}(\alpha\circ\gamma\circ\varphi) && \text{(definition of } d(\gamma\circ\varphi)\text{)}\\
&= \frac{\partial\varphi}{\partial t}\bullet\frac{\partial}{\partial\varphi}(\alpha\circ\gamma) && \text{(chain rule)}^{26}\\
&= \frac{\partial\varphi}{\partial t}\bullet d\gamma(\frac{\partial}{\partial\varphi})(\alpha) && \text{(definition of } d\gamma\text{)}\\
&= \frac{\partial\varphi}{\partial t}\bullet\frac{d\gamma}{d\varphi}(\alpha) && \text{(definition of } \tfrac{d}{d\varphi}\text{)}
\end{aligned}$$

Therefore,

$$\frac{d\gamma\circ\varphi}{dt} = \frac{\partial\varphi}{\partial t}\bullet\frac{d\gamma}{d\varphi}. \qquad (\frac{\partial\varphi}{\partial t}\in\mathbb{R}, \frac{d\gamma}{d\varphi}\in T_{\gamma(t)}(M))$$

Since ω is a linear function at every point,

$$\omega(\frac{d\gamma\circ\varphi}{dt}|_t) = \frac{\partial\varphi}{\partial t}(t)\bullet\omega(\frac{d\gamma}{d\varphi}|_{\varphi(t)}),$$

hence

$$\int_{\gamma\circ\varphi} \omega, = \int_c^d \omega(\frac{d\gamma}{d\varphi}|_{\varphi(t)})\bullet\frac{d\varphi}{dt}|_t\bullet dt = \int_a^b \omega(\frac{d\gamma}{d\varphi}|_\varphi)\ d\varphi = \int_\gamma \omega.$$

---

[25] Otherwise we could divide the intervall into finitely many pieces where this condition is satisfied.
[26] Note that the functions φ and $\alpha\circ\gamma$ both are maps between open subsets of the real line. We consider γ as a function of the variable φ.

2. The fundamental theorems

2.1 Statement of the theorems

2.1.1 Recifiable vector fields

Let M be a C^r manifold ($r \geq 2$),

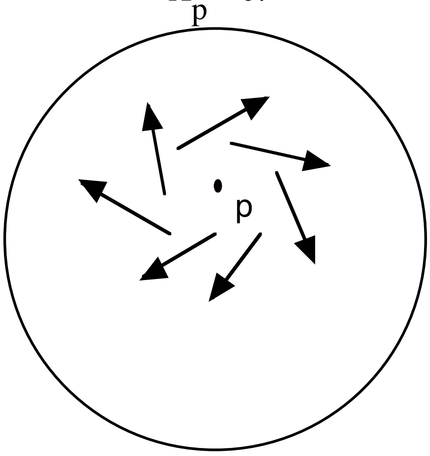
$$X: M \longrightarrow TM$$

a C^s vector field on M ($0 \leq s < r$)[27] and

$$p \in M$$

be a point. The point p is called a singular point of X and X is called singular at p if the vector of X at p is the zero vector,

$$X_p = 0.$$



A singular point of vector field

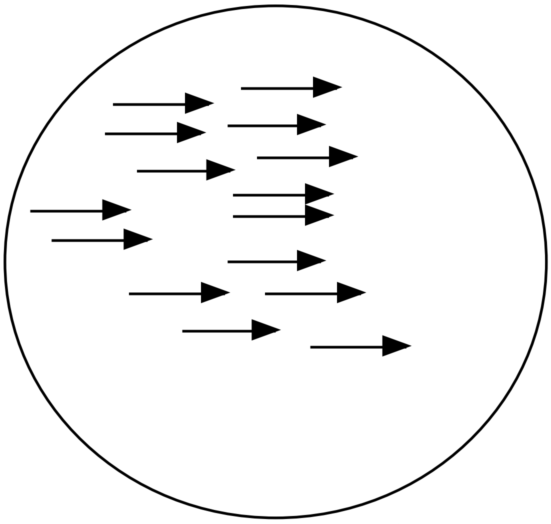
The vector field X is called (locally) rectifiable at p, if there is a chart

$$x: U \longrightarrow V \subseteq \mathbb{R}^n, u \mapsto \begin{pmatrix} x_1(u) \\ \dots \\ x_n(u) \end{pmatrix}$$

such that $p \in U$ such that all coordinate functions of $X|_U$ with respect to x are constant,

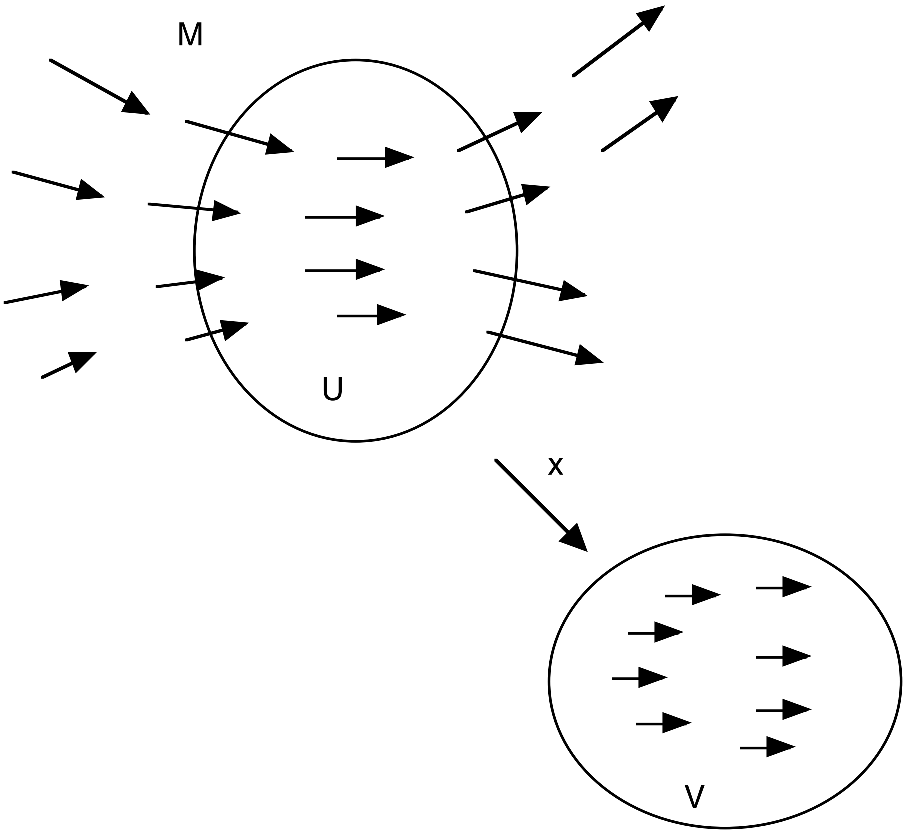
$$X(x_1) = c_1 , ... , X(x_n) = c_n \ , c_1 , ... , c_n \in \mathbb{R}.$$

[27] In case $r = \infty$ or $r = \omega$ we allow $s = r$.



A constant vector field

With other words, $X|_U$ is a constant vector field in the given coordinate system: all vector have the same direction and the same length. In this situation, x is called a rectifying chart for X at p and one says the chart x rectifies X at p.



A rectifiable vector field

The vector field X is called (globally) rectifiable, if there is a diffeomorphismus

$$\varphi: M \longrightarrow V \ (\subseteq \mathbb{R}^n)$$

with an open subset V of euclidian space such that $d\varphi(X)$ is a constant vector field on V.

Remark

(i) Every constant vector field on $\mathbb{R}^n$ is rectifiable. In particular, the standard vector fields

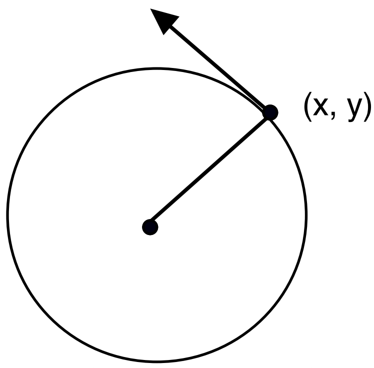
$$\frac{\partial}{\partial x_i} : \mathbb{R}^n \longrightarrow T\,\mathbb{R}^n$$

are rectifiable.

(ii) The rotational vector field

$$\mathbb{R}^2 \longrightarrow T\,\mathbb{R}^2, p = \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x\bullet\frac{\partial}{\partial y} - y\bullet\frac{\partial}{\partial x}|_p$$

is singular at the origin. It is not rectifiable at p : its vector at p is zero in every coordinate system, all other vectors are none-zero (in every coordinate system).



(iii) If the vector field is rectifiable and none singular at p, the coordinate system

$$x: U \longrightarrow V$$

can be always chosen such that $X|_U$ becomes the i-th standard vektor field,

$$X|_U = \frac{\partial}{\partial x_i}$$

(for every given i). Just compose the given chart with a linear transformation

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n$$

mapping the constant none-zero vektor

$$\begin{pmatrix} c_1 \\ \dots \\ c_n \end{pmatrix}$$

to the i-th standard vektor e_i.

(iv) If X is rectifiable at p and $\varphi: U \longrightarrow V$ is a rectifying coordinate system at p, then

$$d\varphi\,(X)$$

is a constant vetor field on V, i.e. $X|_U$ is globally rectifiabel.

(v) The direction field of a differential equation (in the extended phase space) is none-singular at every point: ist t-coordinate is equal to 1 at every point.

2.1.2 The rectification theorem for vector fields

Let

$$X : M \longrightarrow TM$$

be a C^s vector field with $s \geq 1$ on the smooth manifold M which is none-singular at a point

$$p \in M.$$

Then there is a C^s chart of M near p which rectifies X at p.[28]

Warning
The assertion is in general wrong in case s = 0. There are continuous vektor fields which cannot be rectified (we will see this later).

2.1.3 The rectification theorem for direction fields
Every direction field of a differential equation

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

on a smooth manifold M with X smooth is (loaclly) rectifyable at every point.[29]

2.1.4 Simplification theorem
Every differential equation

$$\frac{dx}{dt} = X(x, t) \qquad (1)$$

associated to a C^s vector field X with $s \geq 1$ is locally equivalent to a differential equation of type

$$\frac{dx}{dt} = e_1 \qquad (2)$$

in euclidian space[30], i.e. to a system

$$\begin{aligned} \frac{dx_1}{dt} &= 1 \\ \frac{dx_2}{dt} &= 0 \\ &\dots \\ \frac{dx_n}{dt} &= 0 \end{aligned} \qquad (3)$$

Remark
Note that the initial value problem

$$x(t_0) = x_0$$

of the differential equation (2) has the unique solution

$$x(t) = x_0 + (t - t_0)\cdot e_1$$

(which is defined for every t).

2.1.5 Local existence and uniqueness theorem
Consider the differential equation

$$\frac{dx}{dt} = X(t, x).$$

---

[28] If M is a C^r manifold and X is a C^s vector field which is none-singular at p (with r and s as usual, i.e. $r \geq 2$ and $0 < s < r$ where $s = r$ is allowed in the cases $r = \infty$ and $r = \omega$), then there is a C^s chart of M at p which is retifying for X.

[29] If M a a C^r manifold and X is a C^s vector field with r and s as usual(in particular $s \geq 1$, then for every point p of the extended phase space there is a C^s chart at p which is rectifying for the direction field at p.

[30] i.e. there are diffeomorphisms of class C^s transforming the direction field of this differential equation locally into this simplified form.

where X is a C^S vector field with $s \geq$ smooth Manifold M. Then for every point $x_0 \in M$ and every $t_0 \in \mathbb{R}$ the solution of the initial value problem

$$x(t_0) = x_0$$

exists and is unique in the sense that any two such solutions coincide on the intersection on their intervalls of definition.[31]

Proof: This is trivial for constant vector fields X(t, x), hence follows from the simplification theorem.

QED.

2.1.6 Dependency upon the initial values

Consider the initial value problem

$$\frac{dx}{dt} = X(t, x), x(t_0) = x_0 \qquad (1)$$

where X is a C^S vector field with $s \geq 1$ on a smooth Manifold M.[32] Then there is an open intervall $I \subseteq \mathbb{R}$ containing t_0 and a neighbourhood $U \subseteq M$ containing x_0 such that the map

$$I \times I \times M \longrightarrow M, (t, t_0, x_0) \mapsto x(t, t_0, x_0),$$

is well defined and of class C^S, where

$$x = x(t, t_0, x_0): I \longrightarrow M$$

denotes the solution of the initial value problem (1).

In other words, the solution $x = x(t, t_0, x_0)$ of the initial value problem (1) is of class C^S as a function of t, t_0 and x_0.

Proof: The claim is trivial for constant vector fields X(t, x), hence follows from the simplification theorem.

QED.

2.1.7 Dependency upon parameters

Consider the initial value problem

$$\frac{dx}{dt} = X(t, x, \lambda), x(t_0) = x_0 \qquad (1)$$

where X is a C^S vector field with $s \geq 1$ on a smooth Manifold M depending upon a parameter $\lambda \in V$ varying in a open set $V \subseteq \mathbb{R}$.[33]

---

[31] A solution is considered to be a function $x: I \longrightarrow M$, which is defined on an open intervall $I = (a, b)$ of the real line.

[32] i.e. there is an open interval $I \subseteq \mathbb{R}$ containing $t_0 \in \mathbb{R}$ and an open neighbourhood $U \subseteq M$ of the point $x_0 \in M$ such that the associated vector field in the extented phase space

$I \times U \longrightarrow T(I \times U) = \mathbb{R} \times TU, (t, u) \mapsto (1, X(t,u)),$

is of class C^S.

[33] i.e. for every $\lambda_0 \in V$ there is a neighbourhood $V' \subseteq V$ of λ_0, an open interval $I \subseteq \mathbb{R}$ containing $t_0 \in \mathbb{R}$ and an open neighbourhood $U \subseteq M$ of the point $x_0 \in M$ such that the associated vector field

$I \times U \times V' \longrightarrow T(I \times U \times V') = \mathbb{R} \times TU \times \mathbb{R}^m, (t, u, \lambda) \mapsto (1, X(t,u, \lambda), \lambda),$

Then for every $\lambda_0 \in V$ there is a neighbourhood $V' \subseteq V$ of λ_0, an open intervall $I \subseteq \mathbb{R}$ containing t_0 and a neighbourhood $U \subseteq M$ containing x_0 such that the map

$$I \times I \times M \times V' \longrightarrow M, (t, t_0, x_0, \lambda) \mapsto x(t, t_0, x_0, \lambda),$$

is well defined and of class C^s, where

$$x = x(t, t_0, x_0, \lambda): I \longrightarrow M$$

denotes the solution of the initial value problem (1).

In other words, the solution

$$x = x(t, t_0, x_0, \lambda)$$

of the initial value problem (1) is of class C^s as a function of t, the intial values t_0, x_0 and the parameters λ.

Proof:. Consider the system of differential equations

$$\frac{dx}{dt} = X(t, x, \lambda)$$

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

on the manifold $M \times V$. A solution of this equation is a pair

$$(x(t), \mu_0)$$

where x(t) is a solution of the original equation (1) and μ_0 is any point of U. The initial value problem

$$x(t_0) = x_0$$

$$\lambda(t_0) = \lambda_0$$

has a solution, which is a C^s function of t, t_0, x_0 and λ_0.

QED.

2.1.8 Local phase flow theorem

Let

$$X: M \longrightarrow TM$$

be a C^s vector field with $s \geq 1$ on a smooth manifold M and $p \in M$ be a point. Then there is a neighbourhood U containing p,

$$p \in U \subseteq M$$

an open Intervall $I = (-\varepsilon, +\varepsilon) \subseteq \mathbb{R}$ containing 0,

$$0 \in I \subseteq \mathbb{R}$$

and a C^s map

$$g: I \times U \longrightarrow M, (t, x) \mapsto g^t x,$$

such that the following conditions are satisfied.

(i) For every $x_0 \in U$ the map

$$I \longrightarrow M, t \mapsto g^t x_0$$

---

defined by X is of class C^s.

is a solution of the (autonomous) initial value problem

$$\frac{dx}{dt} = X(x), x(0) = x_0 .$$

(ii) For arbitraryf $x \in U$, $s,t \in I$ such that $s+t \in I$ one has

$$g^t g^s x = g^{t+s} x.$$

Proof. By the rectification theorem we may assume that M is an open subset of $\mathbb{R}^n$ and X is constant. The map

$$\varphi: \mathbb{R}\times\mathbb{R}^n \longrightarrow \mathbb{R}^n, (t, x) \mapsto x + X\bullet t$$

is continuous, hence $\varphi^{-1}(M)$ is an open subset of $\mathbb{R}\times\mathbb{R}^n$ containing the point

$$(0, p) \in \mathbb{R}\times M.$$

Thus, there is some reel $\varepsilon > 0$ and an open neighbourhood $U \subseteq M$ of p such that

$$I \times M \subseteq \varphi^{-1}(M).$$

Then the restriction

$$I \times M \longrightarrow M, (t, x) \mapsto g^t x := x + X\bullet t ,$$

of φ satisfies the requirements of the theorem. Note that

$$g^t g^s x = g^t(x + X\bullet s) = (x + X\bullet s) + X\bullet t = x + X\bullet(s+t) = g^{s+t} x.$$

QED.

Remark

All the theorems above we deduced from the rectification theorem, except for the implication

2.1.6 ⇒ 2.1.7

(depency upon initial values implies the depency upon parameters)

Roughly speaking, all the theorems above can be derived from each other. For example, one has obviously[34] the implication

2.1.6 (+ 2.1.5) ⇒ 2.1.8

(depency upon initial values implies the local phase flow theorem)

Since this will be important for the final proof of these theorems, we will give below the proofs for two further implications:

2.1.8 ⇒ 2.1.3

2.1.8 ⇒ 2.1.2

This will reduce the proofs of all the theorems above to prove of existence and uniquenes and of the C^s dependency upon the initial values.

2.1.9 The local phase flow theorem implies the rectification theorem for direction fields

Passing to the extended phase space we may assume the differential equation in 2.1.3 is autonomous,

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

---

[34] If $g^t x_0$ denotes the solution of the initial value problem $x(0) = x_0$ then

$$g^t g^s x_0 \text{ and } g^{t+s} x_0$$

both denote solutions of the intial value problem $x(0) = g^s x_0$, hence must be equal.

i.e., $X: M \longrightarrow TM$ is a C^s vector field with $s \geq 1$ on a smooth manifold M. Since the claim is of local nature, we may assume that the manifold M is an open subset of euclidian space, say

$$M \subseteq \mathbb{R}^n.$$

For a given point $p \in M$ consider the phase flow as in 1.5.8, say

$$g: I \times U \longrightarrow M, (t, x) \mapsto g^t x,$$

with $U \subseteq M$ open with $p \in U$ and $I = (-\varepsilon, +\varepsilon) \subseteq \mathbb{R}$. Consider the induced map in the extended phase space

$$\tilde{g}: I \times U \longrightarrow I \times M, (t, x) \mapsto (t, g^t x).$$

This is a C^s map (by 1.5.8). It will be sufficient to prove that the following.

1. Locally at (0, p) the map $\tilde{g}$ is a diffeomorphism (hence its inverse can be used as a chart of $I \times M$ near $\tilde{g}(0, p) = (0, p)$.
2. The standard vector field $\frac{\partial}{\partial t}$ on $I \times U$ corresponds to the vector field

$$\tilde{X}: I \times M \longrightarrow T(I \times M) = \mathbb{R} \times T(M), (t, x) \mapsto (1, X(x)).$$

<u>Proof of 1</u>. By the inverse function theorem, it will be sufficient to show that the Jacobiam matrix $J(\tilde{g})$ of $\tilde{g}$ at (0, p) is invertible. One has

$$J(\tilde{g}) = \frac{\partial(t, g^t x)}{\partial(t,x)} = \begin{pmatrix} \frac{\partial t}{\partial t} & \frac{\partial t}{\partial x_1} & \cdots & \frac{\partial t}{\partial x_n} \\ \frac{\partial (g^t x)_1}{\partial t} & \frac{\partial (g^t x)_1}{\partial x_1} & \cdots & \frac{\partial (g^t x)_1}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial (g^t x)_n}{\partial t} & \frac{\partial (g^t x)_n}{\partial x_1} & \cdots & \frac{\partial (g^t x)_n}{\partial x_n} \end{pmatrix} (0,p)$$

Here we denote by $(g^t x)_i$ the i-th coodinate of the vector $g^t x \in M \subseteq \mathbb{R}^n$ .Now, by definition $g^0 x = x$, hence the submatrix formed by the n bottom and right rows and columns is the unit matrix, i.e.,

$$J(\tilde{g}) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ * & 0 & \cdots & 1 \end{pmatrix}$$

is invertible.

<u>Proof of 2</u>. It will be sufficient to show that

$$(d_{(t,x)}\tilde{g})(\frac{\partial}{\partial t}) = \tilde{X}(\tilde{g}(t,x))$$

The vector on the left hand side is the product of the Jacobian matrix of $\tilde{g}$ at (t, x) with the first standard unit vector $e_1 = \frac{\partial}{\partial t}$, i.e. it is the first column of the Jacobian matrix,

$$(d_{(t,x)}\tilde{g})(\frac{\partial}{\partial t}) = \begin{pmatrix} 1 \\ \frac{\partial(g^t x)_1}{\partial t} \\ \dots \\ \frac{\partial(g^t x)_n}{\partial t} \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{\partial(g^t x)}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ X(g^t x) \end{pmatrix} = \tilde{X}(\tilde{g}(t,x))$$

QED.

2.1.10 The local phase flow theorem implies the rectification for vector fields

Let

$$X: M \longrightarrow TM, u \mapsto X_u$$

be a C^s vector field with $s \geq 1$ on the smooth manifold M, which is none-singular at the point $p \in M$, i.e.,

$$X_p \neq 0.$$

We have to prove that X is rectifiable at p (by a C^s chart of M).

Since the assertion to be proved is local in nature, we may assume that M is an open subset in euclidian space, say

$$M \subseteq \mathbb{R}^n \text{ open}.$$

Moreover, using a translation of the euclidian space, we may assume that the given point p is the origin,

$$p = o = \begin{pmatrix} 0 \\ \dots \\ 0 \end{pmatrix} \in M \subseteq \mathbb{R}^n.$$

By assumption, the vector X_p is none-zero. Using a linear (invertible) transformation[35], we may assume the X_p is the last standard unit vector,

$$X_p = e_n = \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \end{pmatrix} \in T_p(M) = T_p(\mathbb{R}^n) = \mathbb{R}^n$$

Consider the C^s map

$$g: I \times U \longrightarrow M, (t, x) \mapsto g^t x,$$

of the local phase flow theorem, where U is a neighbourhood of $p \in M$,

$$p \in U \subseteq M \subseteq \mathbb{R}^n.$$

and $I \subseteq \mathbb{R}$ is an open intervall containing 0,

$$0 \in I \subseteq \mathbb{R}.$$

[35] For example, a rotation followed by a multiplication with a none-zero constant.

In particular, the map $I \longrightarrow M, t \mapsto g^t x_0$, is a solution of the initial value problem

$$\frac{dx}{dt} = X(x),\ x(0) = x_0 . \tag{1}$$

for every $x_0 \in U$.

Identify $\mathbb{R}^{n-1}$ with the hyperplane in $\mathbb{R}^n$ which is orthogonal to e_n ,

$$\mathbb{R}^{n-1} \subseteq \mathbb{R}^n ,\ \begin{pmatrix} u_1 \\ \cdots \\ u_{n-1} \end{pmatrix} \mapsto \begin{pmatrix} u_1 \\ \cdots \\ u_{n-1} \\ 0 \end{pmatrix},$$

an write

$$U' := U \cap \mathbb{R}^{n-1}.$$

It will be sufficient to prove that restriction of g to $I\times U'$, i.e., the map

$$\tilde{g}: I\times U' \longrightarrow M,\ (t, x) \mapsto g^t x,$$

satisfies the following conditions.

1. $\tilde{g}$ is a local diffeomorphism at (0, p) and maps (0, p) to p (hence its inverse can be used as a chart of M in a neighbourhood of p).
2. The first standard vector field e_1 on
$$I\times U' \subseteq \mathbb{R}\times\mathbb{R}^{n-1} = \mathbb{R}^n$$
corresponds to the vector field X on U (in a neighbourhood of p).

Proof of 1. Obviously

$$\tilde{g}(0,p) = g^0 p = p.$$

By the inverse function theorem it will be sufficient to show that the Jacobian matrix

$$J(\tilde{g})$$

of $\tilde{g}$ at (0, p) is invertible. By definition

$$J(\tilde{g}) = \begin{pmatrix} \frac{\partial(g^t x)_1}{\partial t} & \frac{\partial(g^t x)_1}{\partial x_1} & \cdots & \frac{\partial(g^t x)_1}{\partial x_{n-1}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial(g^t x)_n}{\partial t} & \frac{\partial(g^t x)_n}{\partial x_1} & \cdots & \frac{\partial(g^t x)_n}{\partial x_{n-1}} \end{pmatrix} (0, p).$$

Since $t \mapsto g^t x_0$ is the solution of the initial value problem (1), the first colum of this matrix is equal to $X_p = e_n$. Since $x \mapsto g^0 x$ is the identity map,

$$\frac{\partial(g^0 x)_i}{\partial x_j} = \delta_{ij}$$

i.e., the submatrix of $J(\tilde{g})$ formed by the first n-1 rows and last n-1 columns, is the identity matrix. Therefore,

$$J(\tilde{g}) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 1 & * & * & \dots & * \end{pmatrix}$$

is invertible.

Proof of 2. It will be sufficient to show that

$$(d_{(t,x)}\tilde{g})(\frac{\partial}{\partial t}) = X_{\tilde{g}(t,x)}$$

for every $t \in I$ and every $x \in U'$. The vector on the left hand side is just the first column of the Jacobian matrix of $\tilde{g}$ at $\begin{pmatrix} t \\ x \end{pmatrix}$, which is the value of the vector field X at $\tilde{g}(t, x)$.

QED.

Remark

The proofs of the theorems formulated in this section is now reduced to the proof of the theorems about local existence and uniquenes and about the depency upon the initial values (theorems 1.5.5. and 1.5.6).

2.1.11 Rectification examples

A. Rectification of direction fields

Consider the vector field

$$X: \mathbb{R}^2 \longrightarrow T\mathbb{R}^2$$

with

$$X(t, x) = \begin{pmatrix} 1 \\ \cos^2 x \end{pmatrix} = \frac{\partial}{\partial t} + \cos^2(x) \cdot \frac{\partial}{\partial x}$$

This is the direction field of the differential equation

$$\frac{dx}{dt} = \cos^2 x$$

The solution of the initial value problem $x(0) = x_0$ is given by

$$\int_{x_0}^{x} \frac{dx}{\cos^2 x} = \int_0^t dt$$

$$[\tan(x)]_{x_0}^{x} = t$$

$$\tan x - \tan x_0 = t$$

$$\tan x = \tan x_0 + t$$

$$x = \arctan(t + \tan x_0)$$

Thus the phase flow is the map

$$\mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R},\ (t, x) \mapsto g^t x = \arctan(t + \tan x)$$

Denote the new rectifying coordinates by t' and x'. Then the rectifying transformation is

$$\mathbb{R}^2 \longrightarrow \mathbb{R}^2, \begin{pmatrix} t' \\ x' \end{pmatrix} \mapsto \begin{pmatrix} t' \\ \arctan(t' + \tan x') \end{pmatrix}$$

i.e.

$$\boxed{\begin{array}{ll} t & = t' \\ x & = \arctan(t' + \tan x') \end{array}}$$

To get the rectifying coordinate system, we have to inverte this map:

$$\begin{array}{ll} \tan x & = t' + \tan x' \\ \tan x' & = \tan x - t' \\ x' & = \arctan(\tan x - t') \end{array}$$

Thus, the rectifying chart is given by.

$$\boxed{\begin{array}{ll} t' & = t \\ x' & = \arctan(\tan(x) - t) \end{array}}$$

i.e. it is the map

$$\mathbb{R}^2 \longrightarrow \mathbb{R}^2, \begin{pmatrix} t \\ x \end{pmatrix} \mapsto \begin{pmatrix} t \\ \arctan(\tan(x) - t) \end{pmatrix}$$

Let check whether this chart is really rectifying. From the above coordinate changes we get (by the chain rule)

$$\begin{aligned} \frac{\partial}{\partial t'} &= \frac{\partial t}{\partial t'} \bullet \frac{\partial}{\partial t} + \frac{\partial x}{\partial t'} \bullet \frac{\partial}{\partial x} \\ &= \frac{\partial}{\partial t} + \frac{\partial x}{\partial t'} \bullet \frac{\partial}{\partial x} \end{aligned}$$

Note that

$$\begin{aligned} \arctan \tan x &= x \\ \frac{d \arctan}{dx}(\tan x) \bullet \frac{1}{\cos^2 x} &= 1 \\ \frac{d \arctan}{dx}(\tan x) &= \cos^2 x =^{36} \frac{1}{1+\tan^2 x} \\ \frac{d \arctan(x)}{dx} &= \frac{1}{1+x^2} \\ \frac{\partial x}{\partial t'} &= \frac{1}{1+(t'+\tan x')^2} \\ &= \frac{1}{1+(\tan x)^2} \\ &= \cos^2 x \end{aligned}$$

Therefore

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \cos^2 x \bullet \frac{\partial}{\partial x} \qquad = X(t,x)$$

The given vector field X , expressen in the new coordinates t', x', is just the constant vector field in the direction of the t'-axis.

B. Rectification of none-singular vector fields

Problem: find a recifying coordinate system at the origin $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for the vector field

---

[36] $\tan^2 x \bullet \cos^2 x = \sin^2 x = 1 - \cos^2 x$, hence $(\tan^2 x + 1) \bullet \cos^2 x = 1$.

$$X: \mathbb{R}^2 \longrightarrow T\mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto X\begin{pmatrix} x \\ y \end{pmatrix} := \cos^2(x)\bullet\frac{\partial}{\partial x} + \cos^2(y)\bullet\frac{\partial}{\partial y}$$

The vector field is none-singular at the origin (and of class C^1),

$$X\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq 0,$$

hence a rectifying coordinate systeam exists.
The differential equation defined by the vector field is the following system.

$$\frac{dx}{dt} = \cos^2(x)$$

$$\frac{dy}{dt} = \cos^2(y)$$

Its general solution is given by

$$\int \frac{dx}{\cos^2 x} = \int dt$$

$$\int \frac{dy}{\cos^2 y} = \int dt$$

i.e.

$$\tan x = t + \text{const}$$
$$\tan y = t + \text{const}$$

i.e.

$$x = \text{arc}\tan(t + C_1) + \pi,m\,, \qquad m \in \mathbb{Z}$$

$$y = \text{arc}\tan(t + C_2) + \pi n, \qquad n \in \mathbb{Z}$$

To calculate the phase flow, we consider the initial value problem

$$x(0) = x_0\,,\; y(0) = y_0$$

For the constants C_1 and C_2 we get

$$x_0 = \text{arc}\tan C_1 + \pi,m$$
$$y_0 = \text{arc}\tan C_2 + \pi,n$$

i.e.,

$$C_1 = \tan x_0$$
$$C_2 = \tan y_0$$

Hence the phase flow is given by

$$\boxed{g^t\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \text{arc}\tan(t + \tan(x)) + \pi m(x) \\ \text{arc}\tan(t + \tan(y)) + \pi n(y) \end{pmatrix}}$$

where the integers

$$m = m(x)$$
$$n = n(y)$$

are such that

$$x = \text{arc}\tan(\tan(x)) + \pi m$$
$$y = \text{arc}\tan(\tan(y)) + \pi n$$

The above formula assumes that x, and y are in the domain of definition of ‘tan’, i.e.

$x \neq (2k+1)\bullet\frac{\pi}{2}$

$y \neq (2k+1)\bullet\frac{\pi}{2}$

If x is an odd multiple of $\frac{\pi}{2}$ we define the expression

$$\arctan(t + \tan(x)) + \pi m(x) := x$$

to be x independently upon the value of t. An similarly

$$\arctan(t + \tan(y)) + \pi m(y) := y$$

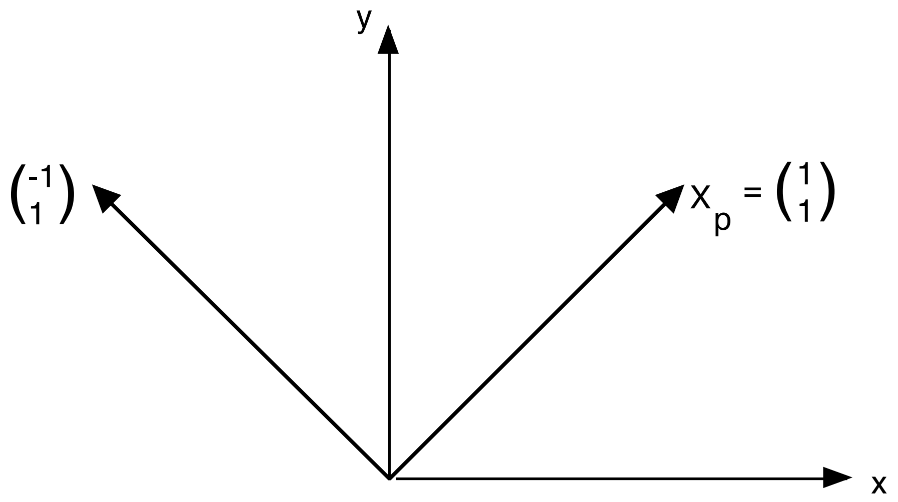
Then the above formula gives the phase flow for every point in $\mathbb{R}^2$.

$$g: \mathbb{R}\times\mathbb{R}^2 \longrightarrow \mathbb{R}^2, (t, \begin{pmatrix} x \\ y \end{pmatrix}) \mapsto g^t\begin{pmatrix} x \\ y \end{pmatrix}.$$

To get the rectifying transformation at $p := \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, we have to restrict g to a subspace

$$\mathbb{R}\times H \subseteq \mathbb{R}\times\mathbb{R}^2,$$

where $H \subseteq \mathbb{R}^2$ is a hyperplane through p, which is transversal[37] to $X_p = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.



Choose, for example, the straight line through the origine in the direction of $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Let s denote the coordinate on this line. Then the rectifying transformation is given by

$$\psi:\mathbb{R}^2 \longrightarrow \mathbb{R}^2, \begin{pmatrix} t \\ s \end{pmatrix} \mapsto g^t(s\bullet\begin{pmatrix} -1 \\ 1 \end{pmatrix}) = g^t\begin{pmatrix} -s \\ s \end{pmatrix},$$

i.e.

$$\psi\begin{pmatrix} t \\ s \end{pmatrix} = \begin{pmatrix} \arctan(t - \tan(s))+\pi m(-s) \\ \arctan(t + \tan(s))+\pi n(s) \end{pmatrix}$$

i.e.

$$\begin{aligned} x &= \arctan(t - \tan(s)) + \pi m(-s) \\ y &= \arctan(t + \tan(s)) + \pi n(s) \end{aligned}$$

---

[37] i.e. the straigh line through p in the direction von X_p is not contained in H.

The coordinate system, we are looking for, is locally at the origine the inverse

$$\varphi = \psi^{-1}$$

of ψ. Let calculate this inverse.

$$\begin{aligned}
\tan(x) &= t - \tan(s)\\
\tan(y) &= t + \tan(s)\\
\tan(y) + \tan(x) &= 2t\\
\tan(y) - \tan(x) &= \tan(s)\\
\\
t &= \frac{1}{2}(\tan(y) + \tan(x))\\
s &= \arctan(\tan(y) - \tan(x)) + \pi\ell(s)
\end{aligned}$$

where the integer ℓ is such that s - $\pi\ell(s)$ is in the domain of definition of the arc tan function. Near the origin we can take ℓs) = 0. Thus our chart is given by

$$\boxed{\begin{aligned} t &= \frac{1}{2}(\tan(y) + \tan(x))\\ s &= \arctan(\tan(y) - \tan(x)) \end{aligned}}$$

ore

$$\boxed{\varphi\begin{pmatrix} x\\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(\tan(y) + \tan(x))\\ \arctan(\tan(y) - \tan(x)) \end{pmatrix}}$$

Let's check whether we got what we want:

$$\begin{aligned}
\frac{\partial}{\partial t} &= \frac{\partial x}{\partial t}\bullet\frac{\partial}{\partial x} + \frac{\partial y}{\partial t}\bullet\frac{\partial}{\partial y}\\
\frac{\partial x}{\partial t} &= \frac{\partial}{\partial t}(\ \arctan\ (t-\tan(s))\\
&= \frac{1}{1 + (t - \tan(s))^2}\\
&= \frac{1}{1 + \tan^2 x}\\
&= \cos^2 x\\
\frac{\partial y}{\partial t} &= \frac{\partial}{\partial t}(\arctan(t + \tan(s)))\\
&= \frac{1}{1 + (t + \tan(s))^2}\\
&= \frac{1}{1 + \tan^2 y}\\
&= \cos^2 y
\end{aligned}$$

Together we get

$$\frac{\partial}{\partial t} = \cos^2(x)\bullet\frac{\partial}{\partial x} + \cos^2(y)\bullet\frac{\partial}{\partial y} = X\begin{pmatrix} x\\ y \end{pmatrix}.$$

In the coordinate system with the new coordinates s,t, the given vector field X is just the standard vector field $\frac{\partial}{\partial t}$ in the direction of the t-axis (near the origin).

We conclude this section with a slightly different version of the local phase flow theorem: the small interval of 0 in the above formulation can be replaced with a small interval containing a precribed value $t_0 \in \mathbb{R}$ and the vector field can be replaced with a time dependent vector field. The price for this generalization: we have to skip the local 1-parameter group property.

2.1.12 Local phase flow theorem II

Let $I = (a,b) \subseteq \mathbb{R}$ be an interval and

$$X: I \times M \longrightarrow TM$$

be a time dependent C^s vector field[38] with $s \geq 1$ on a smooth manifold M. Further let

$$t_0 \in \mathbb{R} \text{ and } p_0 \in M.$$

Then there are an open interval containing 0 and a neighbourhood containing p_0, say

$$t_0 \in I = (t_0 - \varepsilon, t_0 + \varepsilon) \text{ and } p_0 \in U \subseteq M,$$

and a C^s map

$$g: I \times U \longrightarrow M, (t, x) \mapsto g^t x,$$

such that the map

$$I \longrightarrow M, t \mapsto g^t x_0$$

is a solution of the initial value problem

$$\frac{dx}{dt} = X(t, x), x(t_0) = x_0 ,$$

for every $x_0 \in U$.

Proof. We pass to the extended phase space and replace the given differential equation by the equivalent equation

$$\frac{d}{dt}\begin{pmatrix} s \\ x \end{pmatrix} = \begin{pmatrix} 1 \\ X(s,x) \end{pmatrix}$$

i.e.,

$$\frac{ds}{dt} = 1 \text{ and } \frac{dx}{dt} = X(s, x).$$

This way we may assume the differential equation is autonomous.:

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

with a C^s vector field

$$X: M \longrightarrow TM$$

which does not depend upon the time. From the local phase flow theorem 2.1.8 we know that there is a C^s map

$$\tilde{g}: \tilde{I} \times U \longrightarrow M, (t, x) \mapsto \tilde{g}^t x,$$

where $\tilde{I} = (-\varepsilon , +\varepsilon)$ is an open interval containing 0 and $U \subseteq M$ is an open subset containing p_0 such that

$$\tilde{I} \longrightarrow M, t \mapsto \tilde{g}^t x_0 ,$$

---

[38] i.e. the locally given coordinate functions of X are C^s functions of $t \in I$ and $x \in M$.

is a solution of the intial value problem

$$\frac{dx}{dt} = X(x), x(0) = x_0.$$

Let

$$I := (t_0 - \varepsilon, t_0 + \varepsilon)$$

and consider the C^s map

$$g: I \times U \longrightarrow M, (t, x) \mapsto g^t x := \tilde{g}^{t-t_0} x.$$

For every $x_0 \in U$ one has

$$g^{t_0} x_0 = \tilde{g}^0 x_0 = x_0.$$

Moreover,

$$\frac{d}{dt}(g^t x) = \frac{d}{dt} \tilde{g}^{t-t_0} x = X(\tilde{g}^{t-t_0} x) = X(g^t x),$$

i.e.

$$I \longrightarrow M, t \mapsto g^t x_0,$$

is a solution of the initial value problem $x(t_0) = x_0$.

QED.

2.2 Extension theorems

We have yet to prove theorems 2.1.5 and 2.1.6 (existence and uniquenes and dependency upon initial values). But before giving these proofs, we want to turn to the question how to derive from the above theorems assertions of none-local nature.

Up to now all we know is that the solutions exist some small intervals, and we do not know how small the intervalls must be chosen. The aim of this section is to fill this gap (at least to some extent). The proofs of this section are based on the theorems formulated in 2.1.

To illustrate what we have in mind, we start with an example.

2.2.1 Linear systems

A. Existence and uniqueness for homogeneous systems

Let $I = (a, b) \subseteq \mathbb{R}$ be an open intervall,

$$A: I \longrightarrow \mathbb{R}^{n\times n}, t \mapsto A(t),$$

a C^s map with $s \geq 1$ associating with each $t \in I$ an $n\times n$-matrix $A(t)$ of real numbers. Consider the differential equation

$$\frac{dx}{dt} = A(t)\bullet x$$

on the manifold $M = \mathbb{R}^n$. Then there is a C^s map

$$g: I\times I\times M \longrightarrow M, (t, t_0, x_0) \mapsto g(t, t_0, x_0),$$

such that

$$I \longrightarrow M, t \mapsto g(t, t_0, x_0),$$

is a solution of the intial value problem $x(t_0) = x_0$. The solution of the initial value problem

$$\frac{dx}{dt} = A(t)\bullet x,\ x(t_0) = x_0\ ,$$

is unique (for every $t_0 \in I$ and every $x_0 \in \mathbb{R}^n$) in the sence that any two solutions coincide on the intersection of their domains of definition.

Remark
Note that this assertion is a global theorem. It does not claim that there is an interval, where a solution is defined without saying how large or small this intervall is. Quite the contrary, is asserts much more: it gives explicitly the domain of definition of all solutions.
Proof. Uniquess. Let $x = x(t)$ and $y = y(t)$ be two solutions of the intitial value problem

$$\frac{dx}{dt} = A(t)\bullet x,\ x(t_0) = x_0\ ,$$

(defined on certain intervalls containing t_0). We have to prove, they are equal on the intersection of their domains of definition. Replacing these domains by their intersecion, we may assume their domains of defintion are equal, say equal to an intervall $J \subseteq \mathbb{R}$ containing t_0 ,

$$x: J \longrightarrow \mathbb{R}^n,\ y: J \longrightarrow \mathbb{R}^n.$$

By the local uniqueness theorem 2.1.5 the solutions must be equal on a small intervall containing t_0, say

$$x|_{(t_0 - \varepsilon,\ t_0 + \varepsilon)} = y|_{(t_0 - \varepsilon,\ t_0 + \varepsilon)} \text{ for some } \varepsilon > 0.$$

Assume there is some $a \in J$ such that

$$x(a) \neq y(a).$$

Then either

$$a \le t_0 - \varepsilon \text{ or } t_0 + \varepsilon \le a.$$

In what follows we will assume

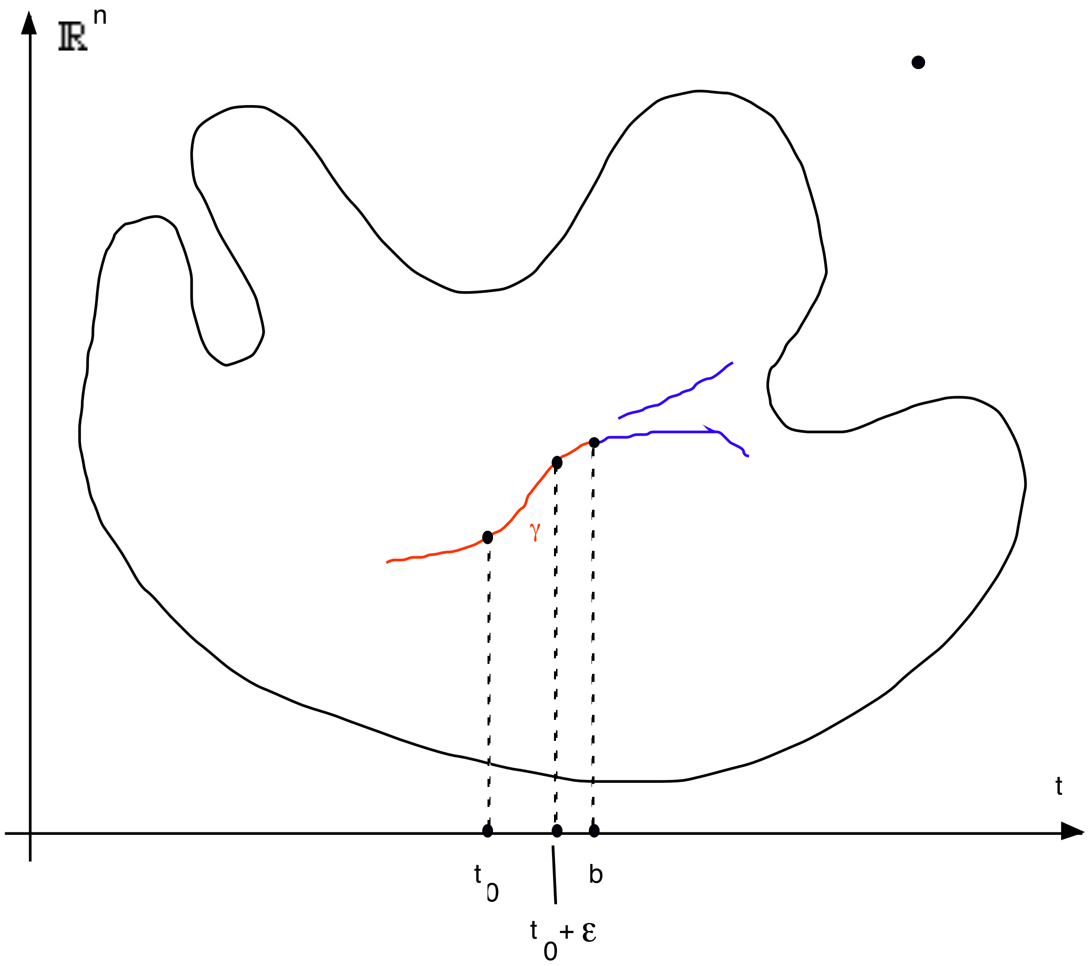
$$t_0 + \varepsilon \le a$$

(the case $a \le t_0 - \varepsilon$ is treated in the same way). Consider the set of t to the right of t_0 where x and y are different and take the largest lower bound for this set, say

$$b := \inf \{t \in J \mid t_0 + \varepsilon \le t,\ x(t) \neq y(t)\}.$$

Since the set in consideration contains a (hence is none-empty), we get

$$t_0 + \varepsilon \le b$$



By the choise of b we have

$$x(t) = y(t) \text{ for every } t \in (t_0 - \varepsilon, b).$$

Since x and y are continuous,

$$x(b) = y(b).$$

By the local uniqueness theorem 2.1.5 the two solutions must coincide on the small intervall containing b, i.e. there is some real $\delta > 0$ such that

$$x(t) = y(t) \text{ for every } t \in (t_0 - \varepsilon, b + \delta).$$

However, this contradicts the choise of b as a largest lower bound. This contradiction shows that there is now $a \in J$ to the right of t_0 where x and y are different. The case that a is a value to the left of t_0 is treated in the same way.

Existence. The idea of the proof is to glue together the local solutions and to use for this the fact that the unit sphere in n-space around the origin is compact. Denote by

$$S := S^{n-1} \subseteq \mathbb{R}^n$$

this unit sphere, i.e.,

$$S := \{x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 = 1\}.$$

This set is bounded and closed, hence it is compact. For every $t_0 \in I$ and every $x_0 \in S$ consider the map

$$g_{t_0,x_0} : I_{t_0,x_0} \times U_{t_0,x_0} \longrightarrow M, (t,x) \mapsto g^t_{t_0,x_0} x,$$

where $I_{t_0,x_0} \subseteq \mathbb{R}$ is an interval containg t_0 and $U_{t_0,x_0} \subseteq M$ is an open set containing x_0 such that

$$I_{t_0,x_0} \longrightarrow M, t \mapsto g^t_{t_0,x_0} y, \qquad (1)$$

is a solution of the initial value problem $x(t_0) = y$ for every $y \in U_{t_0,x_0}$.

If the point x_0 varies on S, the open sets U_{t_0,x_0} cover the unit sphere S. Since S is compact, a finite number of these open sets do, say

$$S \subseteq U_{t_0,x_1} \cup \ldots \cup U_{t_0,x_n}, x_1, \ldots, x_n \in S.$$

Consiter the intersection of time intervals associated to the points $x_1, \ldots, x_n$, say

$$I_{t_0} := I_{t_0,x_1} \cap \ldots \cap I_{t_0,x_n}.$$

Then for $x_0 \in \{x_1, \ldots, x_n\}$ the maps (1) can be restricted to I_{t_0} and give solutions of the initial value problem $x(t_0) = y$. From the uniqueness assertion proved above we see that

$$g^t_{t_0,x_i} y = g^t_{t_0,x_j} y \text{ for } t \in I_{t_0} \; y \in U_{t_0,x_i} \cap U_{t_0,x_j}$$

(solutions of the same initial value problem coincide on the intersection of their domains of definition). With other words, any two of the maps

$$g_{t_0,x_i} : I_{t_0} \times U_{t_0,x_i} \longrightarrow M, (t,x) \mapsto g^t_{t_0,x_i} x,$$

coincide on the intersections of their domains of defintion. Hence they glue together to a C^s map

$$\tilde{g}_{t_0} : I_{t_0} \times S \longrightarrow M, (t,x) \mapsto \tilde{g}^t_{t_0} x,$$

such that

$$I_{t_0} \longrightarrow M, t \mapsto \tilde{g}^t_{t_0} y,$$

is a solution of the intial value problem $x(t_0) = y$ for every $y \in S$. To get rid of the restriction to the unit sphere, consider the map

$$g_{t_0} : I_{t_0} \times M \longrightarrow M, (t,x) \mapsto g^t_{t_0} x,$$

with

$$g^t_{t_0} x := \begin{cases} 0 & \text{if } x = 0 \\ |x| \bullet g^t_{t_0}(\frac{1}{|x|} \bullet x) & \text{otherwise} \end{cases}$$

Here

$$|x| := \sqrt{x_1^2 + ... + x_n^2} \text{ (if } x = \begin{pmatrix} x_1 \\ ... \\ x_n \end{pmatrix})$$

denotes the euclician norm of the vector x. Note that

$$g^{t_0}_{t_0} x = |x| \bullet \tilde{g}^{t_0}_{t_0}(\frac{1}{|x|} \bullet x) = |x| \bullet \frac{1}{|x|} \bullet x = x$$

Moreover,

$$\begin{aligned} \frac{d}{dt}(g^t_{t_0} x) &= \frac{d}{dt}(|x| \bullet \tilde{g}^t_{t_0}(\frac{1}{|x|} \bullet x)) \\ &= |x| \bullet \frac{d}{dt} \tilde{g}^t_{t_0}(\frac{1}{|x|} \bullet x)) \\ &= |x| \bullet A(t) \bullet (\frac{1}{|x|} \bullet x) \\ &= A(t) \bullet x \end{aligned}$$

We see that the map

$$I_{t_0} \longrightarrow M, t \mapsto g^t_{t_0} y, \qquad (2)$$

is a solution of the intial value problem $x(t_0) = y$ (for every $y \in M$). To prove the claim of the theorem, we have to show that in the definition of the map

$$g_{t_0} : I_{t_0} \times M \longrightarrow M, (t, x) \mapsto g^t_{t_0} x, \quad (3)$$

the interval $I_{t_0} \subseteq I$ can be replaced with I. To see this consider all maps of type (3) such that (2) gives a solution of the initial value problem $x(t_0) = y$ (with all possible intervalls I_{t_0} - we know at least one such map exists). From the uniqueness assertion proved above we know that two such maps coicide on the intersection of their domains of definition. Hence these maps glue together to a map

$$g_{t_0} : I^{max}_{t_0} \times M \longrightarrow M, (t, x) \mapsto g^t_{t_0} x, \qquad (4)$$

where the interval $I^{max}_{t_0}$ is the union of all possible intervals I_{t_0}, i.e. it is the largest possible one. It will be sufficient to show

$$I^{max}_{t_0} = I.$$

Assume the converse, i.e.

$$I^{max}_{t_0} = (a', b') \subset I = (a, b)$$

is strictly contained in I, i.e.

$$a < a' \text{ or } b' < b.$$

In case

$$b' < b$$

one has $b' \in I$ and by the above there is a local phase flow map

$$g_{b'}: I_{b'}^{max} \times M \longrightarrow M, (t, x) \mapsto g_{b'}^{t}x, (4)$$

where $I_{b'}^{max} \subseteq \mathbb{R}$ is a open interval containing b'. In particular it has a non-empty intersection with $I_{t_0}^{max}$,

$$\Delta := I_{t_0}^{max} \cap I_{b'}^{max} \neq \varnothing.$$

For every $(s, y) \in \Delta \times M$ the two integral curves through (s,y) defined by the two local phase flows g_{t_0} and $g_{b'}$ must be equal (the uniqueness assertion proved above), i.e.,

$$\{(t, g_{t_0}^{t} y) \mid t \in \Delta\} = \{(t, g_{b'}^{t} y) \mid t \in \Delta\},$$

i.e.

$$(t, g_{t_0}^{t} y) = (t, g_{b'}^{t} y) \text{ for } t \in \Delta$$

i.e.

$$g_{t_0}^{t} y = g_{b'}^{t} y \text{ for } t \in \Delta .$$

We see that g_{t_0} and $g_{b'}$ coincide a every point where both maps are defined. Hence they clue together to give a local phase flow map

$$(I_{t_0}^{max} \cup I_{b'}^{max}) \times M \longrightarrow M.$$

But this contradicts the maximality of the interval $I_{t_0}^{max}$. This contradiction shows

$$b' = b.$$

The case $a < a'$ is treated in the same way.

QED.

B. The structure of the solution set

As above let $I = (a, b) \subseteq \mathbb{R}$ be an open intervall,

$$A: I \longrightarrow \mathbb{R}^{n\times n}, t \mapsto A(t),$$

a C^s map with $s \geq 1$ associating with each $t \in I$ an $n\times n$-matrix A(t) of real numbers. Further let

$$S = S(I, A)$$

denote the set of solutions

$$x: I \longrightarrow \mathbb{R}^n$$

(defined on I) of the differential equation

$$\frac{dx}{dt} = A(t) \cdot x$$

Then:

(i) S is a real vector space of dimension n.

(ii) For every $t_0 \in I$ the map

$$\varphi : S \longrightarrow \mathbb{R}^n, x \mapsto x(t_0),$$

is an isomorphisms of vector spaces.

(iii) For every $t_0 \in I$ the solutions $x_1, \ldots, x_m \in S$ are linear independent if and only if the vectors $x_1(t_0), \ldots, x_m(t_0) \in \mathbb{R}^n$ are independent.

Proof. S is a vector space. Let $x, y \in S$ be two solutions and $c, d \in \mathbb{R}$ be constants. We have to prove that

$$c \cdot x + d \cdot y \in S.$$

One has

$$\begin{aligned} \frac{d}{dt}(c \cdot x(t) + d \cdot y(t)) &= c \cdot \frac{d}{dt} x(t) + d \cdot \frac{d}{dt} y(t) \\ &= c \cdot A(t) \cdot x(t) + d \cdot A(t) \cdot y(t) \quad \text{(for, x,, y} \in S) \\ &= A(t)(c \cdot x(t) + d \cdot y(t)). \end{aligned}$$

We have proved, $c \cdot x + d \cdot y$ is a solution. To complete the proof of the theorem, it is sufficient to prove assertion (ii).

The map φ of assertion (ii) is linear. For $x, y \in S$ and $c, d \in \mathbb{R}$ one has

$$\begin{aligned} \varphi(c \cdot x + d \cdot y) &= (c \cdot x + d \cdot y)(t_0) \\ &= c \cdot x(t_0) + d \cdot y(t_0) \\ &= c \cdot \varphi(x) + d \cdot \varphi(y). \end{aligned}$$

The map φ of assertion (ii) is surjectiv. This follows from the existence assertion of the previous theorem.

The map φ of assertion (ii) is injectiv. This follows from the uniqueness assertion of the previous theorem.

QED.

C. Fundamental systems

As above let $I = (a, b) \subseteq \mathbb{R}$ be an open intervall,

$$A: I \longrightarrow \mathbb{R}^{n \times n}, t \mapsto A(t),$$

a C^s map with $s \geq 1$ associating with each $t \in I$ an $n \times n$-matrix $A(t)$ of real numbers. A fundamental system of the differential equation

$$\frac{dx}{dt} = A(t) \cdot x \qquad (1)$$

is a matrix

$$X(t) = (x_1(t), \ldots, x_n(t))$$

such that the columns $x_1, \ldots, x_n$ form a basis of the space $S = S(I, A)$ of solutions of the differential equation (1).

Properties of fundamental systems

(i) $\frac{d}{dt}X(t) = A(t)\bullet X(t)$.

(ii) For every invertible matrix $C \in \mathbb{R}^{n\times n}$, the production $X(t)\bullet C$ is again a fundamental system.

(iii) Every fundamental system of (1) is obtained as in (ii) from a single one.

Proof. Assertion (i). One has

$$\begin{aligned}\frac{d}{dt}X(t) &= \frac{d}{dt}(x_1(t), \ldots, x_n(t))\\ &= (\frac{d}{dt}x_1(t), \ldots, \frac{d}{dt}x_n(t))\\ &= (A(t)\bullet x_1(t), \ldots, A(t)\bullet x_n(t))\\ &= A(t)\bullet(x_1(t), \ldots, x_n(t))\\ &= A(t)\bullet X(t).\end{aligned}$$

Assertion (ii). Let

$$Y(t) := X(t)\bullet C$$

and denote by $y_i(t)$ the i-th column of Y(t). Further let c_{ij} the entry of C in Position (i, j). Then

$$(x_1(t), \ldots, x_n(t)) = (x_1(t), \ldots, x_n(t))\bullet C$$

hence

$$y_i(t) = \sum_{j=1}^{n} x_j(t)\bullet c_{ji} \in S$$

In a similar way one sees from $Y(t)\bullet C^{-1} = X(t)$ that the x_i's can be written as a linear combination of the y_i's. Since the former are a basis of S, the same is true for the latter. Therefore $Y(t) = (y_1(t), \ldots, y_n(t))$ is a fundamental system.

Assertion (iii). Let be

$$Y(t) = (y_1(t), \ldots, y_n(t))$$

be a matrix whose columns are in S. Since $X(t) = (x_1(t), \ldots, x_n(t))$ is a fundamental system, the y_i's are linear combination of the x_i's, say

$$y_i(t) = \sum_{j=1}^{n} x_j(t)\bullet c_{ji} \in S$$

with uniquely determined $c_{ji} \in \mathbb{R}$.

But than,

$$Y(t) = X(t)\bullet C$$

where C is the matrix whose entry at position (i, j) is C. If, moreover, Y(t) is a fundamental system, then the same reasing as above gives a matrix $D \in \mathbb{R}^{n\times n}$ with

$$X(t) = Y(t)\bullet D.$$

In particular

$$X(t) = X(t)CD.$$

Since the colums of X(t) are linear independent (for evey $t \in I$), the matrix X(t) is invertible, i.e. CD is the unit matrix and

$$1 = \det(CD) = \det C \cdot \det D.$$

Therefore, det (C) is none-zero and C is invertible.

QED.

D. Wronski determinants

As above let $I = (a, b) \subseteq \mathbb{R}$ be an open intervall,

$$A: I \longrightarrow \mathbb{R}^{n\times n}, t \mapsto A(t),$$

a C^s map with $s \geq 1$ associating with each $t \in I$ an $n\times n$-matrix $A(t)$ of real numbers. Moreover, let

$$x_1, \dots, x_n : I \longrightarrow \mathbb{R}^n \qquad (1)$$

be solutions of the differential equation

$$\frac{dx}{dt} = A(t)\cdot x \qquad (2)$$

The Wronski determinant of the solutions $x_1, \dots, x_n$ is by definition the determinant

$$W(x_1.\text{---}.x_n) := \det(x_1.\text{---}.x_n)$$

of the matrix, whose columns are the solutions (1).

Properties of Wronski determinants

(i) For $x_1, \dots, x_n \in S(I, A)$ the following assertions are equivalent.

1. $X(t) = (x_1(t), \dots, x_n(t))$ is a fundamental system.
2. $x_1, \dots, x_n$ are linear independent elements of $S = S(I,A)$.
3. $W(x_1, \dots, x_n)(t_0) \neq \det X(t_0) = 0$ for every $t_0 \in I$.
4. $W(x_1, \dots, x_n)(t_0) \neq \det X(t_0) = 0$ for a single $t_0 \in I$

(ii) Every Wronski determinant $W(t)$ of the system (1) satisfies

$$\frac{d}{dt}W(t) = tr(A(t))\cdot W(t)$$

where $tr(M)$ denotes the trace of the square matrix M, i.e. the sum of the entries on the main diagonal.

Proof. Assertion (i). By the definition of the notion of fundamental system, the first two assertions are equivalent. Since the map

$$S \longrightarrow \mathbb{R}^n, x \mapsto x(t_0),$$

is an isomorphism, one has

$x_1, \dots, x_n$ linear independent

$\Leftrightarrow x_1(t_0), \dots, x_n(t_0)$ linear independent

$\Leftrightarrow \det X(t_0) \neq 0$

This shows that 2. is equivalent ot 3. and to 4.

Assertion (ii). Fix some

$$t_0 \in I$$

and consider the fundamental system

$$Y(t) = (y_1(t), \dots, y_n(t))$$

such that

$$Y(t_0) = Id.$$

is the unit matrix Id. Then

$$\frac{d}{dt} Y(t) = A(t)Y(t) \tag{3}$$

and

$$\frac{d}{dt} \det Y(t) = \sum_{i=1}^{n} \det (y_1(t), \dots, y_{i-1}(t), \frac{d}{dt} y_i(t), y_{i+1}(t), \dots, y_n(t))$$

$$= \sum_{i=1}^{n} \det (y_1(t), \dots, y_{i-1}(t), A(t)y_i(t), y_{i+1}(t), \dots, y_n(t))$$

In particular, for $t = t_0$ we obtain

$$\frac{d}{dt} \det Y(t)|_{t=t_0} = \sum_{i=1}^{n} \det (e_1, \dots, e_{i-1}(t), A(t_0)e_i(t), e_{i+1}(t), \dots, e_n(t))$$
$$= tr(A(t_0))$$

i.e.

$$\frac{d}{dt} \det Y(t)|_{t=t_0} = tr(A(t_0)) \tag{4}$$

Now let X(t) be an arbitrary square matrix whose columns are solution of the system (2) and write

$$X(t) = Y(t) \cdot C \qquad \text{with } C \in \mathbb{R}^{n \times n}$$
$$W(t) := \det X(t).$$

Then

$$X(t_0) = Y(t_0) \cdot C = Id \cdot C = C$$

hence

$$X(t) = Y(t) \cdot X(t_0)$$

and

$$W(t) = \det Y(t) \cdot W(t_0).$$

Therefore

$$\frac{d}{dt} W(t) = \frac{d}{dt} Y(t) \cdot W(t_0)$$

From (4) we see that

$$\frac{d}{dt} W(t)|_{t=t_0} = \frac{d}{dt} Y(t)_{t=t_0} \cdot W(t_0)$$
$$= tr(A(t_0)) \cdot W(t_0)$$

i.e.

$$\frac{d}{dt} W(t)|_{t=t_0} = tr(A(t_0)) \cdot W(t_0).$$

The last identity is true for evey t_0, i.e., assertion (ii) is true.

QED.

Remark

If $A \in \mathbb{R}$ is a real constant, the differential equation

$$\frac{dx}{dt} = A \cdot x$$

for functions x = x(t) with values in the 1-dimensional euclidian space allows to separate variables, and we know that the solutions are given by

$$\int \frac{dx}{x} = \int A dt$$

i.e.

$$x(t) = C \cdot e^{A \cdot t} \text{ with } C \in \mathbb{R}.$$

It turns out that the same method of solution works is a much more general context. Recall that the exponential function can be written as a power series

$$e^x = \sum_{i=0}^{\infty} \frac{1}{i!} x^i$$

which converges (absolutely) for every $x \in \mathbb{R}$.

E. The exponential function of a square matrix

For every n•n matrix A with real or complex entries one writes

$$e^A := \sum_{i=0}^{\infty} \frac{1}{i!} A^i$$

Example 1.
For $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ one has

$$\begin{aligned} e^A &= \sum_{i=0}^{\infty} \frac{1}{i!} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}^i \\ &= \sum_{i=0}^{\infty} \begin{pmatrix} \frac{1}{i!} \cdot a^i & 0 \\ 0 & \frac{1}{i!} \cdot b^i \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=0}^{\infty} \frac{1}{i!} \cdot a^i & 0 \\ 0 & \sum_{i=0}^{\infty} \frac{1}{i!} \cdot b^i \end{pmatrix} \\ &= \begin{pmatrix} e^a & 0 \\ 0 & e^b \end{pmatrix} \end{aligned}$$

Example 2.
For

$$A := \begin{pmatrix} \frac{2}{3} & 0 & \frac{-2}{3} \\ 0 & \frac{-2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{-2}{3} & 0 \end{pmatrix}$$

By direct calculation one sees

$$A^2 = \begin{pmatrix} \frac{2}{9} & \frac{4}{9} & \frac{-4}{9} \\ \frac{1}{9} & \frac{2}{9} & \frac{-2}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{-4}{9} \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

hence

$$e^{t\bullet A} = Id + t\bullet A + \frac{1}{2}\bullet t^2 A^2$$

hence

$$\begin{aligned} \frac{d}{dt} e^{t\bullet A} &= A + t\bullet A^2 \\ &= A\bullet(Id + t\bullet A) \\ &= A\bullet(Id + t\bullet A + \frac{1}{2}\bullet t^2 A^2) \qquad (\text{since } X^3 = 0) \\ &= A\bullet e^{t\bullet A}, \end{aligned}$$

Thus

$$x(t) := e^{t\bullet A}$$

is a solution of the differential equation

$$\frac{dx}{dt} = A\bullet x.$$

Properties of the exponential function

(i) $e^X := \sum_{i=0}^{\infty} \frac{1}{i!} X^i$ converges absolutely for every $X \in \mathbb{C}^{n\times n}$.[39]

(ii) $e^{X+Y} = e^X\bullet e^Y$ for every two commuting $n\times n$ matrices X and Y.[40]

(iii) $e^{X\bullet Y\bullet X^{-1}} = X\bullet e^Y\bullet X^{-1}$ for arbitrary square matrixes X and Y with X invertible.

(iii) $e^{(s+t)\bullet X} = e^{s\bullet X}\bullet e^{t\bullet X}$ for arbitrary square matrices X and arbitrary complex number s and t.

(iv) $\frac{d}{dt}(e^{t\bullet X}) = X\bullet e^{t\bullet X}$ for arbitrary square matrices X and arbitrary real variables t.j

Proof. The proofs are essentially the same like for the usual exponential function. As for (iii) note that

$$\begin{aligned} (X\bullet Y\bullet X^{-1})^i &= (X\bullet Y\bullet X^{-1})\bullet(X\bullet Y\bullet X^{-1})\bullet\ldots\bullet(X\bullet Y\bullet X^{-1}) \\ &= X\bullet Y^i\bullet X^{-1} \end{aligned}$$

hence

$$\sum_{i=0}^{n} \frac{1}{i!}\bullet(X\bullet Y\bullet X^{-1})^i = X\bullet\left(\sum_{i=0}^{n} \frac{1}{i!}\bullet Y^i\right)\bullet X^{-1}.$$

---

[39] with respect to the norm

$$||X|| = \sqrt{\sum_{i,j=1}^{n} x_{ij}^2}$$

where x_{ij} denotes the entry of the matrix X at position (i,j).

[40] We say that X and Y commute if X•Y = Y•X. Without this assumption the above identity does not hold in general.

Passing to the limit $n \longrightarrow \infty$ gives assertion (iii).
QED.

F. Linear homogeneours systems with constant coefficients

Definition. A linear homogeneous differential equation with constant coefficients is a differential equation of type

$$\frac{dx}{dt} = X \cdot x,$$

where

$$X \in \mathbb{C}^{n \times n}$$

is a matrix whose entries are constant (i.e. real or complex numbers).

The most important fact for the solution of such equations is the following lemma.

Lemma.
Let A, B, C be $n \times n$ matrices with C invertible and

$$B = C \cdot A \cdot C^{-1}$$

Then for

$$x: I \longrightarrow \mathbb{R}^n$$

the following assertions are equivalent.

(i) x is a solution of $\frac{dx}{dt} = A \cdot x$.

(ii) $y(t) = C \cdot x(t)$ is a solution of $\frac{dy}{dt} = B \cdot y$.

Proof. Consider the change of coordinates defined by the invertible matrix C,

$$y = Cx.$$

Then

$$x = C^{-1}y,$$

and substituting into the differential equation

$$\frac{dx}{dt} = A \cdot x.$$

gives

$$\frac{d}{dt}(C^{-1}y) = A \cdot C^{-1}y$$

$$C^{-1}\frac{d}{dt}(y) = A \cdot C^{-1}y$$

$$\frac{dy}{dt} = (CAC^{-1}) \cdot y.$$

$$\frac{dy}{dt} = B \cdot y.$$

This proves (i) $\Rightarrow$ (ii). The converse implication follows by symmetry.
QED.
Example.

Consider the system of differential equations

$$\frac{dx_1}{dt} = x_1 + 2x_2$$

$$\frac{dx_2}{dt} = 2x_1 + x_2$$

i.e.

$$\frac{dx}{dt} = A \cdot x \quad \text{with } A := \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Since A is symmetric, the matrix can be diagonalized. From

$$\det(A - T \cdot Id) = \det\begin{pmatrix} 1 - T & 2 \\ 2 & 1 - T \end{pmatrix} = (1 - T)^2 - 4 = T^2 - 2T - 3 = (T + 1)(T - 3).$$

we see that A has eigen values 3 and -1.
The eigen vectors for the eigen value 3 are given by

$$\begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

The eigen space is generated by

$$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

The eigen vectors for the eigen value -1 are given by

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

The eigen space is generated by

$$v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Using v_1, v_2 as a new coordinate sysstem,

$$v_1 = \frac{1}{\sqrt{2}} \cdot e_1 + \frac{1}{\sqrt{2}} \cdot e_2 \qquad (1)$$

$$v_2 = \frac{1}{\sqrt{2}} \cdot e_1 - \frac{1}{\sqrt{2}} \cdot e_2$$

the corresponding coordinate transformation is given by

$$x_1 e_1 + x_2 e_2 = y_1 v_1 + y_2 v_2,$$

i.e.,

$$x_1 = \frac{1}{\sqrt{2}} v_1 + \frac{1}{\sqrt{2}} v_2$$

$$x_2 = \frac{1}{\sqrt{2}} v_1 - \frac{1}{\sqrt{2}} v_2$$

i.e.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

i.e.,

$$x = C \cdot y$$

with

$$C = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Substitution of $x = Cy$ into the given differential equation $\frac{dx}{dt} = A \cdot x$ yields

$$\frac{d}{dt}(Cy) = A \cdot Cy$$

$$C \cdot \frac{dy}{dt} = (AC)y$$

$$\frac{dy}{dt} = (C^{-1}AC)y$$

$$\frac{dy}{dt} = B \cdot y$$

with

$$B = C^{-1}AC = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

We are reduced to solve the equation

$$\frac{dy_1}{dt} = 3y_1$$

$$\frac{dy_2}{dt} = -y_2$$

The general solution of this system is

$$y_1(t) = C_1 \cdot e^{3t}$$

$$y_2(t) = C_2 \cdot e^{-t}$$

A fundamental system for this equation is, for example,

$$Y(t) = \begin{pmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{pmatrix}$$

(the determinant is none-zero). The corresponding fundamental system of the original equation is (as seen from the above relation between x and y).

$$X(t) = C \cdot Y(t) = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} e^{3t} & 0 \\ 0 & e^{-t} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{3t} & e^{-1} \\ e^{3t} & -e^{-1} \end{pmatrix}$$

Lets check wether the columns are solutions:

$$\frac{dX(t)}{dt} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3e^{3t} & -e^{-1} \\ 3e^{3t} & e^{-1} \end{pmatrix}$$

$$A \cdot X(t) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} e^{3t} & e^{-1} \\ e^{3t} & -e^{-1} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3e^{3t} & -e^{-1} \\ 3e^{3t} & e^{-1} \end{pmatrix} = \frac{dX(t)}{dt}$$

Note that X(t) is a fundamental system since

$$\det X(0) = \det \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = 2 \cdot (1 \cdot (-1) - 1 \cdot 1) = -4 \neq 0.$$

Remarks

(i) To solve a differential equation of the type

$$\frac{dx}{dt} = A \cdot x \qquad (1)$$

one should try to diagonalize the matrix A and solve the system for the resulting diagonal matrix. The solutions of the original equation (1) are obtained from those of the new system by change of coordinates.

(ii) Unfortunately, there are matrices which cannot be diagonalized. The best one can get by conjugation is the Jordan normal form of the matrix X, i.e. one of the type

$$C \cdot A \cdot C^{-1} = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \dots & \dots & \dots & .. \\ 0 & 0 & \dots & J_\ell \end{pmatrix}$$

where each J_i is a Jordan block, i.e.

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix}, \lambda_i \in \mathbb{C}.$$

This reduces the solution of (1) to the case that A is a Jorden block, say

$$A = A(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}$$

In this case one has

$$e^{t\bullet A} = \begin{pmatrix} e^{\lambda t} & t\bullet e^{\lambda t} & \frac{t^2}{2!}\bullet e^{\lambda t} & \dots & \frac{t^{n-2}}{(n-2)!}\bullet e^{\lambda t} & \frac{t^{n-1}}{(n-1)!}\bullet e^{\lambda t} \\ 0 & e^{\lambda t} & t\bullet e^{\lambda t} & \dots & \frac{t^{n-3}}{(n-3)!}\bullet e^{\lambda t} & \frac{t^{n-2}}{(n-2)!}\bullet e^{\lambda t} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & e^{\lambda t} & t\bullet e^{\lambda t} \\ 0 & 0 & 0 & \dots & 0 & e^{\lambda t} \end{pmatrix}$$

Proof. Let f (t) denote the matrix on the right hand side. Then

$$\begin{aligned} e_i \bullet \frac{d}{dt} f(t) &= \text{i-th row of } f(t) \\ &= \lambda\bullet e_i \bullet f_n(t) + e_{i+1}\bullet f(t) \\ &= (\lambda\bullet e_i + e_{i+1})\bullet f(t) \\ &= (\text{i-th row of } A(\lambda))\bullet f(t) \\ &= e_i\bullet A(\lambda)\bullet f(t). \end{aligned}$$

This is true for every i. So

$$\frac{d}{dt} f(t) = A(\lambda)\bullet f(t),$$

i.e. the columns of f(t) are solutions of $\frac{dx}{dt} = X(\lambda)\bullet x$. Moreover, det $f(t) \neq 0$ (for example at t = 0), i.e., f(t) is a fundamental system.
QED.

G. Inhomogeneours systems

Let $I = (a, b) \subseteq \mathbb{R}$ be an open intervall and

$$A: I \longrightarrow \mathbb{R}^{n\times n}, t \mapsto A(t),$$

$$B: I \longrightarrow \mathbb{R}^{n}, t \mapsto B(t),$$

be C^s maps with $s \geq 1$. Moreover, let

$$X: I \longrightarrow \mathbb{R}^{n\times n}$$

be a fundamental system of the differential equation

$$\frac{dx}{dt} = A(t)\bullet x.$$

Then the solutions of the inhomogeneous system

$$\frac{dx}{dt} = A(t)\bullet x + B(t)$$

can be obtained via variation of constants:

$$Y(t) = X(t)(\int X(t)^{-1}B(t)\,dt + C_0), C_0 \in \mathbb{R}^n.$$

Bemerkung
Man beachte, für je zwei Lösungen u(t) und v(t) des inhomogenen Systems

$$\frac{dx}{dt} = A(t)\bullet x + B(t)$$

ist u(t) - v(t) eine Lösung des inhomogenen Systems.

Proof. Given a solution y(t) of the inhomogeneous system, we assume that it can be written

$$y(t) = X(t)c(t), c(t) \in \mathbb{R}^n. \qquad (2)$$

Then

$$A(t)y(t) + B(t) = \frac{d}{dt}y(t) = (\frac{d}{dt}X(t))\bullet c(t) + X(t)\bullet \frac{d}{dt}c(t)$$

i.e.

$$A(t)X(t)c(t) + B(t) = A(t)X(t)c(t) + X(t)\bullet \frac{d}{dt}c(t)$$

i.e.

$$B(t) = X(t)\bullet \frac{d}{dt}c(t)$$

$$B(t)\bullet Y(t)^{-1} = \frac{d}{dt}c(t)$$

$$c(t) = \int_{t=t_0}^{t} B(t)\bullet X(t)^{-1}dt + C_0$$

Substitituting into (2) we get the claim.
QED.

2.2.2 Extendable integral curves

Let M be a smooth manifold,

$$X: I\times M \longrightarrow TM, (t, x) \mapsto X(t,x),$$

a time dependent C^s vector field with $s \geq 1$ and

$$t_0 \in I, x_0 \in M.$$

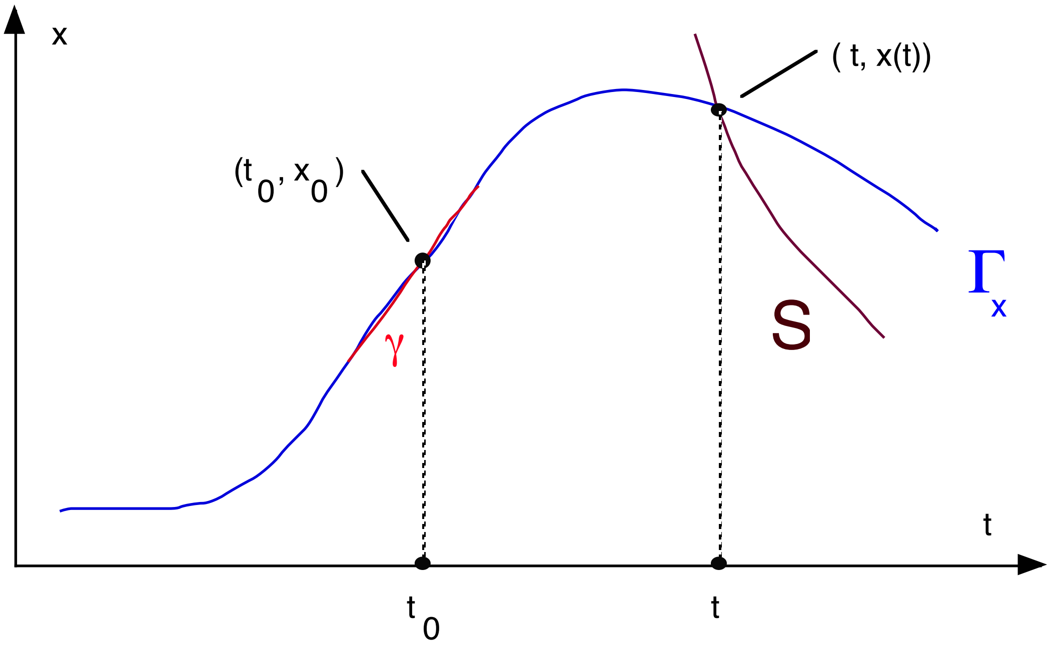
Further let

$$S \subseteq \mathbb{R}\times M$$

be a set of the extented phase space. An integral curve γ through (t_0,x_0) of the differential equation

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

is <u>extendable forward</u> to S (resp. <u>extendable backward</u> to S), if there is an integral curve of this differential equation containing γ which intersects S for some time value $t \geq t_0$ (resp. $t \leq t_0$). Such an integral curve is called <u>extension</u> of γ to S.



A forward extendable integral curve γ

Equivalently, there is a solution

$$x: J \longrightarrow M, t \mapsto x(t),$$

of the differential equation whose graph contains γ,

$$(t_0,x_0) \in \gamma \subseteq \Gamma_x := \{ (t, x(t)) \mid t \in J\},$$

such that

$$(t, x(t)) \in S$$

for some $t \geq t_0$ (resp. $t \leq t_0$).

Let us recall two notions used in the next theorem.

A subset K of a topological space is called compact, if for every covering of K by open sets one can find a finite subcovering. In euclician space a set is compact, if and only if it is closed and bounded.
A point

$$p \in X$$

of a topological space X is called a boundary point of a subset

$$S \subseteq X,$$

if every neighbourhood of p contains points of S and the complement X - S. The set of boundary points of S is denoted

$$\partial S.$$

Examples
The boundary of the open or closed unit disc in the plane is the unit circle.
The boundary of the euclidian space is the empty set.
The boundary of every set is closed.
The boundary of a boundary is the whole set, $\partial\partial S = \partial S$.
The boundary of a compact set is compact (since it is a closed subset).

2.2.3 Extension theorem for integral curves

Let M be a smooth manifold,

$$X: I\times M \longrightarrow TM, (t, x) \mapsto X(t,x),$$

a time dependent C^s vector field with $s \geq 1$ and

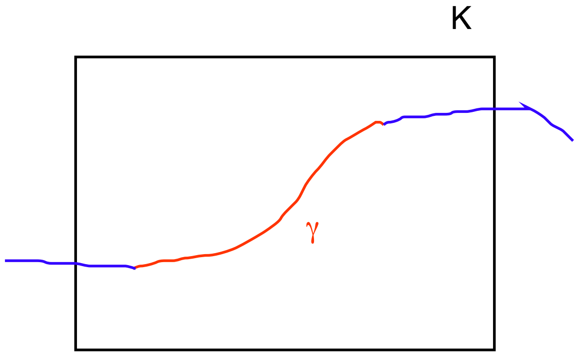
$$p_0 = (t_0, x_0) \in K \subseteq I\times M$$

be a point of a compact set K. Then every integral curve through p_0 of the differential equation

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

is extendable forward and backward to the boundary ∂K of K.

The extensions are unique in the sense the corresponding solutions of the differential equations coincide on the intersections of their domains of definition.



Proof. We restict the question of forward extentability. The backward case is treated in the same way.

1. <u>Uniqueness of the integral curve through</u> p_0. Consider two integral curves through p_0. Let

$$x: I \longrightarrow M \text{ and } y: J \longrightarrow M$$

be the solutions of the differential equations defining these integral curves. By assumption

$$t_0 \in I \cap J \text{ and } x(t_0) = x_0 = y(t_0).$$

By the local uniqueness theorem 2.1.5 these solutions coincide in on small intervall containing t_0, say

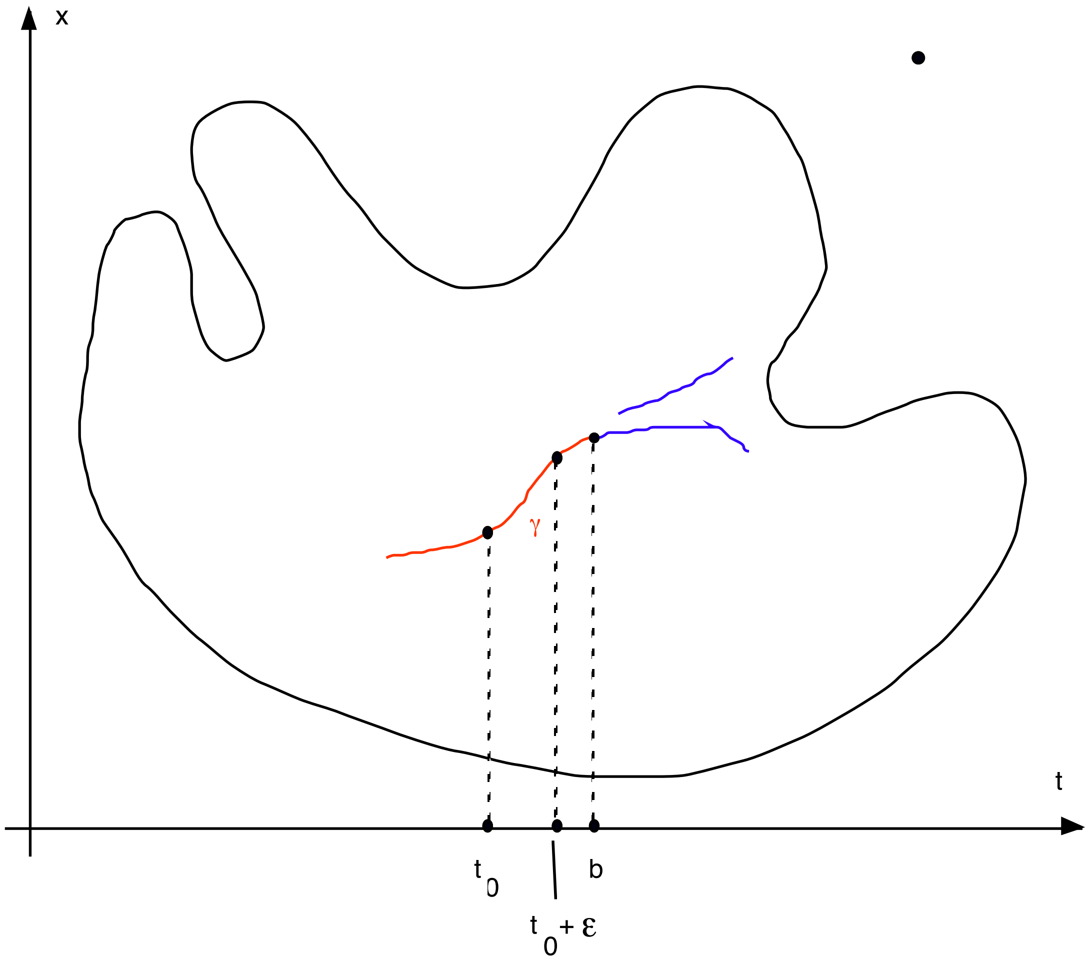
$$x(t) = y(t) \text{ for } t \in (t_0 - \varepsilon, t_0 + \varepsilon) \subseteq I \cap J.$$

Assume that there is some

$$t \in I \cap J, t_0 \leq t,$$

with $x(t) \neq y(t)$. Consider the set of all these values of t and take its infimum, say

$$b := \inf \{t \in I\cap J \mid x(t) \neq y(t), t_0 \leq t\}$$



Then

$$t_0 < t_0 + \varepsilon \leq b$$

and

$$x(t) = y(t) \text{ for every } t \in (t_0 - \varepsilon, b).$$

Since x(t) and y(t) are continuous, they must be equal also at b,

$$x(b) = y(b).$$

But then, x(t) and y(t) must be equal in a small open interval containing b (by the local uniqueness theorem 1.4.5): there is some $\delta > 0$ with

$$x(t) = y(t) \text{ for every } t \in (t_0 - \varepsilon, b + \delta).$$

But this contractict the definiton of b. Hence there is no $t \in I \cap J$ the right of t_0 where x(t) and y(t) are different.

2. Existence of the extension

For every point

$$p = (s, y) \in K$$

apply the local phase flow theorem 2.1.8 to the differential equatin

$$\frac{ds}{dt} = 1$$

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

to get a map

$$g_p : I_p \times U_p \longrightarrow M, (t, (s,x)) \mapsto {}_pg^t(s,x) \,,$$

where $I_p \subseteq \mathbb{R}$ is an interval containing 0 and $U_p \subseteq I \times M$ is an open set containing p such that

$$I_p \longrightarrow M, t \mapsto {}_pg^t(s,x),$$

is a solution of the intial value problem

$$\frac{dx}{dt} = X(t,x), x(s) = y, \tag{1}$$

for every $(s,y) \in U_p$. The open set U_p cover K, hence finitely many of them do, say

$$K \subseteq U_{p_1} \bigcup ... \bigcup U_{p_n} := U$$

Choose $\varepsilon > 0$ so small that

$$I := (-2\varepsilon, +2\varepsilon) \subseteq I_{p_1} \bigcap ... \bigcap I_{p_n}$$

and replace each I_{p_i} with I (i = 1, ... , n). Then any two of the maps

$$g_{p_i} : I \times U_{p_i} \longrightarrow M, (t, (s,x)) \mapsto {}_{p_i}g^t(s,x) ,$$

coincide at every point where both are defined, for they provide solution for the same initial value problems at that point. Hence the maps g_{p_i} glue together to give a map

$$g: I \times U \longrightarrow M, (t, (s,x)) \mapsto g^t(s,x),$$

such that

$$I \longrightarrow M, t \mapsto g^t(s,x),$$

solves the initial value problem (1).
We use this map g to construct a sequence of solutions

$$\xi_i: I \longrightarrow M, t \mapsto x_i(t), i = 0, 1, 2, ...$$

Let x_0 be the map

$$\xi_0: I = (-2\varepsilon, +2\varepsilon) \longrightarrow M, t \mapsto g^t(t_0,x_0),$$

and define

$$(t_1,x_1) = g^\varepsilon(t_0,x_0).$$

If (t_1,x_1) is in $K \subseteq U$, let ξ_1 be the map

$$\xi_1: I = (-2\varepsilon, +2\varepsilon) \longrightarrow M, t \mapsto g^t(t_1,x_1).$$

Assume we have already constructed $\xi_0, ... , \xi_{n-1}$. Define

$$(t_n,x_n) = g^\varepsilon(t_{n-1},x_{n-1}).$$

If (t_n,x_n) is in $K \subseteq U$, let ξ_n be the map

$$\xi_n: I = (-2\varepsilon, +2\varepsilon) \longrightarrow M, t \mapsto g^t(t_n,x_n).$$

By construction

$$\xi_n (0) = g^0(t_n,x_n) = (t_n,x_n) = g^\varepsilon(t_{n-1},x_{n-1}) = \xi_{n-1}(\varepsilon)$$

From the uniquenes assertion of the first step we see that

$$\xi_n(t) = \xi_{n-1}(\varepsilon+t) \text{ for } t \in (-2\varepsilon, +\varepsilon).$$

Hence the solutions ξ_i glue together to give a solution

ξ: $(-2\varepsilon, n\varepsilon) \longrightarrow M$

with

$\xi(t) = \xi_i(t - i\varepsilon)$ for $t \in ((i-2)\varepsilon, +(i+2)\varepsilon)$, $i = 0, \dots, n$.

This construction continues as long as the point (t_n, x_n) is in K. If this point is not in K, the integral curve

$$\{(t, \xi(t)) \mid t \in (-2\varepsilon, (n+1)\varepsilon)\} \quad (2)$$

contains points both in K and not in K (t = 0 gives (t_0,x_0) and t = n-1 gives (t_n,x_n)). Hence it contains boundary point of K, and the proof of the theorem is complete.

Assume that the process does not stop. Then K contains the integral curves (2) for arbitrary natural numbers n. This implies image of K under the projection

$$\varphi: K \subseteq \mathbb{R}\times M \longrightarrow \mathbb{R}, (t, x) \mapsto t,$$

contains the intervalls

$$(-2\varepsilon, (n+1)\varepsilon) \subseteq \varphi(K).$$

But the image $\varphi(K)$ of the compact set K is compact, hence bounded in $\mathbb{R}$. Thus this inclusion cannot hold for every n.

QED.

2.2.4 Extendable solutions

Let M be a smooth manifold,

$$X: M \longrightarrow TM, x \mapsto X(x),$$

a C^s vector field with $s \geq 1$ and

$$t_0 \in \mathbb{R}, x_0 \in M.$$

Further let

$$S \subseteq M$$

be a set of the phase space. A solution

$$x: I \longrightarrow M$$

of the initial value problem

$$\frac{dx}{dt} = X(t,x), x(t_0) = x_0 \quad (1)$$

is <u>extendable forward</u> to S (resp. <u>extendable backward</u> to S), if there is a solution

$$y: J \longrightarrow M$$

of the initial value problem (1) such that

$$I \subseteq J, y|_I = x$$

and such that

$$y(t) \in S$$

for some $t \geq t_0$ (resp. $t \leq t_0$) in the interval J. Such a solution is called <u>extension</u> of x to S. The solution x is called <u>infinitely extendable forward</u> (resp. <u>infinitely extendable backward</u>), if there is a solution

$$y: J = (a, \infty) \longrightarrow M \text{ (resp. } y: J = (-\infty, b) \longrightarrow M)$$

of the initial value problem (1) such that

$$I \subseteq J, y|_I = x.$$

In this situation, the solution y is called an extension of the the solution x.

2.2.5 Extension theorem for solutions

Let M be a smooth manifold,

$$X: M \longrightarrow TM, x \mapsto X(x),$$

a C^s vector field with $s \geq 1$ and

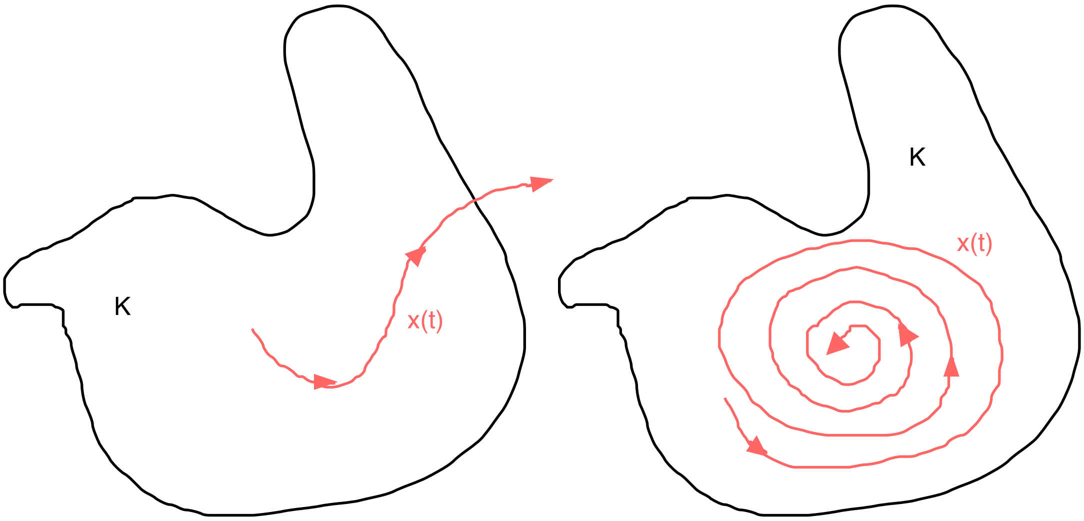
$$x_0 \in K \subseteq M$$

be a point of a compact set K. Then every solution of the initial value problem

$$\frac{dx}{dt} = X(t, x), x(t_0) = x_0$$

is extendable forward (resp. backward) either infinitely or to the boundary ∂K of K.

The extensions are unique in the sense that any two extensions of a given solution coincide on the intersection of their domains of definition.



Proof. Let

$$x: I \longrightarrow M$$

be a solution of the intial value problem

$$x(t_0) = x_0.$$

Consider the associated integral curve

$$\{(t, x(t)) \mid t \in I\} \tag{1}$$

For arbitrary real numbers

$$a, b \in \mathbb{R}$$

such that

$$a < t_0 < b$$

the set

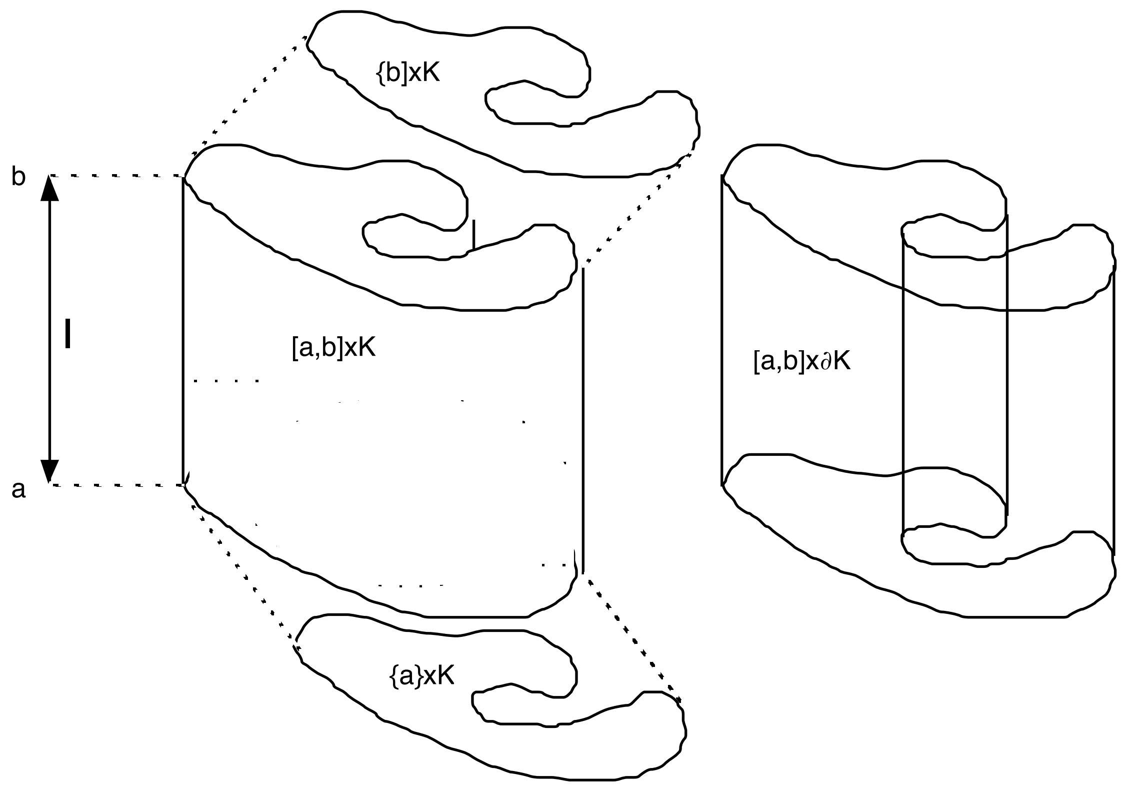
$$[a, b] \times K \subseteq M$$

is a compact. By 2.2.3 every integral curve through

$$(t_0, x_0) \in I \times K$$

is uniquely extendable forward to the boundary

$$\partial([a,b]\times K) = (\partial[a,b])\times K \cup [a,b]\times(\partial K) = \{a\}\times K \cup \{b\}\times K \cup [a,b]\times \partial K$$



In particular, there is a unique forward extension

$$\{(t, y(t)) \mid t \in J\} \qquad (2)$$

of the integral curve (1), which contains a point of the above boundary, say

$$(t, y(t)) \in \partial([a,b]\times K. \text{ for some } t \in J \text{ such that } t_0 < t.$$

Since $a < t_0 < t$, this point cannot be in $\{a\}\times K$, i.e.

$$(t, y(t)) \in \{b\}\times K \cup [a,b]\times \partial K$$

There are two possibilities.

1. Case. For every $b > t_0$ the integral curve intersects $\{b\}\times K$.

Then, for evey $b > t_0$, there is an extension $y: J \longrightarrow M$ of x such that $b \in J$. The solution x is infinitely extentable backward.

2. Case. There is some $b > t_0$ such that the integral curve does not intersect $\{b\}\times K$.

Then the forward extension y of x to the boundary of $[a,b]\times K$ satisfies

$$(t, y(t)) \in [a,b]\times \partial K$$

for some t, hence

$$y(t) \in \partial K.$$

The solution x is extendable to the boundary of K.

We have proved the claim of the theorem concerning forward extendability. The case of backward extendability is treated in the same way.

The uniqueness assertions follows from the uniqueness assertions for the integral curves in 2.2.3.
QED.

2.2.6 The global phase flow theorem

Let

$$X: M \longrightarrow TM$$

be a C^s vector field on the smooth manifold M with $s \geq 1$. Suppose the X is a vector field with <u>compact support</u>.[41]

Then there is a unique C^s map

$$\mathbb{R}\times M \longrightarrow M, (t, x) \mapsto g^t x,$$

such that the following is satisfied.

(i) For every $x_0 \in M$ the map

$$\mathbb{R} \longrightarrow M, t \mapsto g^t x_0 ,$$

is a solution of the initial value problem $\frac{dx}{dt} = X(x), x(0) = x_0$.

(ii) $g^s g^t x = g^{s+t} x$ for arbitrary $x \in M$ and $s,t \in \mathbb{R}$.

Proof. The proof is divided into severy steps.
<u>3. Step</u>. Uniqueness of the map g.

It is sufficient to show that any two solutions of the initial value problem

$$\frac{dx}{dt} = X(x), x(0) = x_0 \tag{1}$$

coincide on the intersection of their domains of definition. But this is proved as uniquesness in the proof of 2.2.3: consider two solutions of the initial value problem. By local uniqueness they must coincide on a small time interval containing 0. Consider the largest lower bound of all positive time values where the two solutions have different values. As in 2.2.3 on sees the two solutions must have the same value at this largest lower bound, hence must coincide on a small interval around this largest lower bound. This gives the disired contradiction.

<u>1. Step</u>. Every solution of $\frac{dx}{dt} = X(x)$, $x(0) = x_0$ is extentable infinitely forward and backward.

Let K denote the support of the vector field X,

$$K = \text{supp } X.$$

In case x_0 is outside K, the vector field X is zero at x_0. Hence the constant function

---

[41] This means that there is a compact subset $K \subseteq M$ such that X is zero outside K,

$$X_p = 0 \text{ for every } p \in M - K.$$

Equivalently on can require that the <u>support</u> supp X of X is compact, where supp X is defined to be the closure of the set

$$\{x \in M \mid X(x) \neq 0\}.$$

The condition is satisfied for every vector field on a compact manifold.

$$x(t) = x_0$$

is a solution of the initial value problem, which is trivially extentable infinitely (and by local uniqueness all solutions are of this type).

In case x_0 is on the boundary of K, the vector field X is again zero at x_0 (by continuity). Hence the situation is as above: the only solution is the constant function, which is infinitely extentable.

Finally let x_0 be in the interior of K. From the previous theorem 2.2.5 we see that the initial value problem (1) is (backward and forward) extentable either to the boundary of K or infinitely. Moreover, from the above we know that a solution taking a value on the boundary must be constant. Hence a solution of (1) with x_0 from the interior of K cannot be extented to the boundary. But then, the solution must be infinitely extentable. This proves the claim of the second step.

2. Step. Proof of (i).

For every $x_0 \in M$ let

$$g_{x_0} : \mathbb{R} \longrightarrow M, t \mapsto g_{x_0}(t),$$

denote the solution of the initial value problem (1) (which exists on a small interval by the local existence theorem and on the whole of the real line by the second step), and write

$$g^t x := g_x(t).$$

This gives a well-defined map

$$g: \mathbb{R} \times M \longrightarrow M$$

such that the condition of assertion (i) is satisfied. By uniqueness (as proved in the first step), this map coicides locally with the local phase flow (as in 2.1.12), hence is a C^s map.

4. Step. Proof of (ii).

By assertion (i),

$$\frac{d}{dt}(g^t x) = X_{g^t x}$$

for every $t \in \mathbb{R}$ and every $x \in M$. The definition of $\frac{d}{dt}$ gives

$$(d\, g^t x)(\frac{\partial}{\partial t}) = X_{g^t x}$$

Identify vectors with their directional derivatives. Then for every test function φ one has

$$\begin{aligned} X_{g^t x}(\varphi) &= (d\, g^t x)(\frac{\partial}{\partial t})(\varphi) \\ &= \frac{\partial}{\partial t}(\varphi \circ g^t x) \quad \text{(Definition of the differential of a map).} \end{aligned}$$

Replacing t by t+s we get

$$\begin{aligned} X_{g^{t+s} x}(\varphi) &= \frac{\partial}{\partial t}(\varphi \circ g^{t+s} x) \\ &= (d\, g^{t+s} x)(\frac{\partial}{\partial t})(\varphi) \\ &= \frac{d}{dt}(g^{t+s} x)(\varphi) \end{aligned}$$

hence

$$\frac{d}{dt}(g^{t+s}x) \quad = X_{g^{t+s}x}.$$

The latter identity says that the map

$$\mathbb{R} \longrightarrow M, t \mapsto g^{t+s}x,$$

is a solution of the initial value problem

$$\frac{dx}{dt} = X, x(0) = g^s x.$$

Uniqueness as proved in the first step implies that

$$g^{t+s}x \quad = g^t(g^s x)$$

(as claimed).
QED.

2.3 The contraction mapping theorem

2.3.1 Modivation: Differential and integral equations

The aim of this section is the proof of the local existence and uniqueness theorems for differential equations and of the theorem about the dependency upon the initial contitions. Since these theorems are of a local nature, we may assume throughout that the manifold in consideration is an open set of euclidian space, i.e. we will consider differential equations

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

where $X: M \longrightarrow TM$ is a C^s vector field on an open subset

$$M \subseteq \mathbb{R}^n$$

with $s \geq 1$.
Given a solution

$$x: I \longrightarrow M, t \mapsto x(t)$$

of the initial value problem

$$\frac{dx}{dt} = X(t, x), x(t_0) = x_0 \qquad (1)$$

with $t_0 \in I$, $x_0 \in M$, one has

$$\int_{t_0}^{t} X(t, x(t))\, dt = \int_{t_0}^{t} \frac{dx(t)}{dt} dt = x(t) - x(t_0) = x(t) - x_0$$

hence

$$x(t) \quad = x_0 + \int_{t_0}^{t} X(t, x(t))\, dt \qquad (2)$$

Conversely, every C^1 function $x: I \longrightarrow$ satisfying condition (2) is a solution of the initial value problem (1). This modivates the introduction of the integral operator

$$A: C(I, \mathbb{R}) \longrightarrow C(I, \mathbb{R}), \varphi \mapsto x_0 + \int_{t_0}^{t} X(t, \varphi(t))\, dt,$$

mapping real valued continuous functions $\varphi: I \longrightarrow \mathbb{R}$ to real valued continuous function on I. From the above we see

$$\varphi \text{ is a solution of (1)} \Leftrightarrow A(\varphi) = \varphi.$$

This way the question about the existence ans uniqueness of an initial value problem is translated into the question whether a certain integral operator has a fixed point.

Below we will see that

1. A slight modification of the domain of definition will make this domain into a complete metric space.
2. If one chooses the intervall I small enough, the operator A will becomes a so-called contracting operator (an operator that shortens distances).
3. Contracting operators on complete metric spaces have precisely on fixed point.

We will proceede as follows now. First we will define the notions used above. Then we will formulate the contraction operator theorem (assertion 3). An finally we will construct the domain of definition for A, which allows to apply the contraction operator theorem to ordinary differential equations.

2.3.2 Metric spaces

A metric space is a set M which is equipped with a distance function

$$d: M \times M \longrightarrow \mathbb{R}$$

also called metric, i.e. a function satifying the following conditions.

(i) $d(x, y) \geq 0$ for arbitrary $x, y \in M$.

(ii) $d(x, y) = 0 \Leftrightarrow x = y$.

(iii) $d(x, y) = d(y, x)$ for arbitrary $x, y \in M$.

(iv) triangle identity.

$d(x, y) + d(y, z) \geq d(x, z)$ for arbitrary $x, y, z \in M$.

For every point $x \in M$ and every positive real $\varepsilon \in \mathbb{R}$ the set of points

$$U_\varepsilon(x) := \{ x' \in M \mid d(x',x) < \varepsilon\}$$

having distance $< \varepsilon$ from x is calle ε-neighbourhood of x. A subset

$$U \subseteq M$$

is called open, if it is a union of (possibly infinitely many) ε-neighbourhoods. This way every metric space becomes a topological space (prove this!). Its topology is called metric topology. Metric spaces will be always equipped with their metric topology.

A sequence $\{x_i\}_{i = 1,2, \ldots}$ of points $x_i \in M$ in the metric space is called convergent, if there is some point $x \in M$ such that every open set $U \subseteq M$ containing x contains x_i for allmost every i (i.e. every i with the possible exception of finitely many of them). The point x is unique[42] in this case and is called limit of $\{x_i\}_{i = 1,2, \ldots}$. One writes

$$x_i \longrightarrow x \text{ and } \lim_{i \longrightarrow \infty} x_i = x.$$

The sequence is called Cauchy sequence, if for every given $\varepsilon > 0$ there is a natural number

$$N(\varepsilon) \in \mathbb{N}$$

such that

---

[42] Let x' be a second limit. Then for every $\varepsilon > 0$ one has

$$x_i \in U_{\varepsilon/2}(x) \text{ and } x_i \in U_{\varepsilon/2}(x')$$

for at least one i (even for almost every i). Therefore

$$0 \leq d(x, x') \leq d(x, x_i) + d(x_i , x') \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrarily choosen, we see $d(x,x') = 0$, hence $x = x'$.

$$d(x_i , x_j) < \varepsilon \text{ for arbitray } i, j \geq N(\varepsilon).$$

Convergent sequences are easily seen to be Cauchy sequences.[43] The metric space M is called complete, if every Cauchy sequence is convergent.

Example 1

The real line is a complete metric space with respect to the distance function

$$d(x , y) := | x - y |.$$

Example 2

The euclidian space $\mathbb{R}^n$ is a complete metric space with respect to the distance function

$$d(x, y) := \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

This kind of metric is called euclidian metric.

Example 3

Let M be a metric space and K a compact metric space. Then every continuous function

$$f: K \longrightarrow M$$

is bounded, i.e. for every $p \in M$ there is a constant $C \in \mathbb{R}$ with

$$d(f(x), p) < C$$

for every $x \in K$.[44] Consequently, for evey two continuous functions

$$f, g: K \longrightarrow M$$

the real value d(f(x), g(x)) is bounded, i.e.

$$d(f, g) := \sup \{d(f(x), g(x)) \mid K\}$$

is a well-defined none-negative real number. From the properties of the distance function on M one sees that this defines a distane function on the set

$$C(K, M)$$

---

[43] Let x be the limit. Then $x_i \in U_{\varepsilon/2}$ for almost every i, say for every $i \geq N(\varepsilon)$. For $i , j \geq N(\varepsilon)$ this implies

$$d(x_i, x_j) \leq d(x_i , x) + d(x, x_j) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

[44] For every $x \in K$ take some $C_x \in \mathbb{R}$ such that

$$d(f(x), p) < C_x ,$$

i.e. $f(x) \in U_{C_x}(p)$. Since f is continuous, there is some $\varepsilon(x) > 0$ such that

$$f(U_{\varepsilon(x)}(x)) \subseteq U_{C_x}(p).$$

The open sets $U_{\varepsilon(x)}(x)$ cover K, hence finitely many of them do, say

$$K = U_{\varepsilon(x_1)}(x_1) \cup ... \cup U_{\varepsilon(x_n)}(x_n)$$

Let $\varepsilon = \min\{\varepsilon_1,...,\varepsilon_n\}$ and $C = \max \{C_{x_1},...,C_{x_n}\}$. Then

$$f(U_\varepsilon(x_i)) \subseteq U_{\varepsilon(x_i)}(x_i) \subseteq U_C(p),$$

i.e.

$$d(p, f(x)) < C$$

for $x \in U_{\varepsilon(x_i)}(x_i)$. Since the $U_{\varepsilon(x_i)}(x_i)$ cover K, the inequality holds for every $x \in K$.

of continuous functions $K \longrightarrow M$ and defines on C(K M) the structure of a metric space. To see that this space is complete, we need the notions of uniform continuity and uniform convergence.

2.3.3 Uniform continuity

Let M and M' be metric spaces. A function

$$f: M \longrightarrow M'$$

is called uniformly continuous, if for every $\varepsilon > 0$ there is some reel $\delta > 0$ such that

$$d(f(x), f(y)) < \varepsilon \text{ for arbitrary points such that } d(x, y) < \delta.$$

Remark

(i) Uniformly continuous functions are easily seen to be continuous.[45]

(ii) If M is compact, every continuous function $M \longrightarrow M'$ is uniformly continuous.[46]

2.3.4 Uniform convergence

Let M and M' be metric spaces with M' complete. A sequence $\{f_i\}_{i=1,2,...}$ of functions

$$f_i: M \longrightarrow M'$$

is called uniformly convergent, if for every $\varepsilon > 0$ there is some natural number $N(\varepsilon)$ such that

$$d(f_i(x), f_j(x)) < \varepsilon$$

for arbitrary $x \in M$ and arbitrary $i, j \geq N(\varepsilon)$.

---

[45] For, $U_\delta(x) \subseteq f^{-1}(U_\varepsilon(f(x)))$, is f is continuous in x.

[46] Fix any $\varepsilon > 0$. For every $x \in M$ chose $\delta = \delta(x) > 0$ such that

$$U_\delta(x) \subseteq f^{-1}(U_{\varepsilon/2}(f(x))),$$

which is possible since f is continuous. Since M is compact finitely many $U_{\delta(x)/2}(x)$ cover M, say

$$M = U_{\delta_1/2}(x_1) \cup ... \cup U_{\delta_n/2}(x_n).$$

Let

$$\delta := \min\{\delta_1/2, ... , \delta_n/2\}.$$

Then

$$d(f(x), f(y)) < \varepsilon \text{ if } d(x, y) < \delta.$$

To see this, let i be such that

$$x \in U_{\delta_i/2}(x_i). \qquad (*)$$

The inequality $d(x, y) < \delta \leq \delta_i/2$ implies

$$y \in U_{\delta_i}(x_i) \subseteq f^{-1}(U_{\varepsilon/2}(f(x_i))),$$

i.e.,

$$d(f(x_i), f(y)) < \varepsilon/2.$$

Similarly, from (2) we see

$$d(f(x_i), f(x)) < \varepsilon/2.$$

Thus $d(f(x), f(y)) < \varepsilon$, as claimed.

Remarks

(i) If $\{f_i\}_{i=1,2,...}$ is uniformly convergent, then it is convergent, i.e. $\{f_i(x)\}_{i=1,2,...}$ is convergent for every $x \in M$.[47]

(ii) If $\{f_i\}_{i=1,2,...}$ is uniformly convergent and every f_i is continuous at a fixed point p $\in M$, then the limit function is continuous at p.[48]

(iii) As a consequence of the above remarks we see that the space C(K, M) of 2.2.2 Example 3 is complete: a Cauchy sequense of C(K, M) is by definition a sequence $\{f_i\}_{i=1,2,...}$ of continuous functions $K \longrightarrow M$, which is uniformly convergent. Define

$$f(x) := \lim_{i\longrightarrow\infty} f_i(x)$$

vor every $x \in K$. By remark (ii), this defines a continuous function

$$f: K \longrightarrow M,$$

i.e. an elment of C(K,M). By assumption,

$$d(f_i(x), f_j(x)) < \varepsilon/2 \text{ for arbitrary } x \in M \text{ and arbitrary } i, j \geq N(\varepsilon/2).$$

For $j \longrightarrow \infty$ we get

$$d(f_i(x), f(x)) \leq \varepsilon/2 < \varepsilon \text{ for arbitrary } x \in M \text{ and arbitrary } i \geq N(\varepsilon/2),$$

hence

$$d(f_i, f) < \varepsilon \text{ for } i \geq N(\varepsilon/2),$$

i.e. $f_i \longrightarrow f$ in C(K, M). The sequence is convergent in C(K, M).

2.3.5 Contrations

A <u>contraction</u> is a map

$$A: M \longrightarrow M, x \mapsto Ax,$$

of a metric space M into itself such that there is real number $\lambda < 1$ satisfying

$$d(Ax, Ay) \leq \lambda \bullet d(x, y)$$

for arbitrary $x, y \in M$. The real number λ is called in this case a <u>contraction factor</u> of A.

---

[47] Since $\{f_i(x)\}_{i=1,2,...}$ is a Cauchy sequence in M'.

[48] Write f(x) for the limit function and fix $\varepsilon > 0$. Since $\{f_i\}_{i=1,2,...}$ is uniformly convergent, there is some $N \in \mathbb{N}$ such that

$$d(f_i(x), f_j(x)) \leq \varepsilon/3$$

for $i, j \geq N$ and arbitrary $x \in M$. For $j \longrightarrow \infty$ we get

$$d(f_i(x), f(x)) \leq \varepsilon/3$$

for $i \geq N$ and arbitrary $x \in M$. Fix i, say $i = N$. Since f_N is continuous at p there is some $\delta > 0$ with,

$$d(f_N(p), f_N(x)) < \varepsilon/3 \text{ for every } x \in U_\delta(p).$$

Therefore

$$d(f(p), f(x)) \leq d(f(p), f_N(p)) + d(f_N(p), f_N(x)) + d(f_N(x), f(x)) < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

for every $x \in U_\delta(p)$. Thus f is continuous at p.

Example 1

A C^1 function A: $\mathbb{R} \longrightarrow \mathbb{R}$ such that $|\frac{dA}{dt}(t)| \le \lambda < 1$ for every $t \in \mathbb{R}$ is a contraction, for

$$| A(t_1) - A(t_2) | \le | \int_{t_1}^{t_2} |\frac{dA}{dt}(t) | dt | \le \max |\frac{dA}{dt}(t)| \bullet |t_2 - t_1 | \le \lambda \bullet |t_2 - t_1 |$$

Example 2

Let A: $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear operator such that all eigenvalues of A are in the open unit disc (of the complex plane). Then there is an euclidian metrix on $\mathbb{R}^n$ such that A becomes a contraction. This is easily seen in for diagonal matrices, say

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \max \{ |\lambda_1| , \dots , |\lambda_n| \} =: \lambda < 1.$$

For evey $x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ one has with respect to the euclidian metric:

$$d(0, Ax) = (|\lambda_1 x_1|^2 + \dots + | \lambda_n x_n|^2)^{1/2} \le \lambda(|x_1|^2 + \dots + | x_n|^2)^{1/2} = d(0, x)$$

hence $d(Ax, Ay) \le \lambda \bullet d(x, y)$. In the general case one uses the Jordan normal form over the complex numbers.

2.3.6 Contraction mapping theorem

Let M be a complete metric space and A: M $\longrightarrow$ M a contraction. Then A has precisely on fixed point, i.e. there is one and only one point $p \in M$ satisfying $Ap = p$.

Moreover, for evey point $x \in M$ the sequence

$$x, Ax, A^2x, A^3x, \dots$$

converges to p. The distance of its n term from the fixed point can be estimated as follows.

$$d(p, A^nx) \le \frac{\lambda^n}{1-\lambda} \bullet d(Ax, x).$$

Proof. Let $\lambda \in (0, 1)$ as in the definition of the contraction and denote

$$d := d(x, Ax).$$

Then

$$d(A^nx, A^{n+1}) \le \lambda \bullet d(A^{n-1}x, A^nx) \le \lambda^2 \bullet d(A^{n-2}x, A^{n-1}x) \le ,,,$$

hence

$$d(A^nx, A^{n+1}x) \le \lambda^n \bullet d.$$

For $n, m \ge N$ one obtains

$$d(A^nx, A^mx) \le \sum_{i \text{ between } n \text{ and } m} d(A^ix, A^{i+1}x) \le d \bullet \sum_{i=N}^{\infty} \lambda^i \qquad (1)$$

Since the sum $\sum_{i=0}^{\infty} \lambda^i$ converges, the partial sums can be made arbitrarily small, i.e.

$$\{ A^n x \}_{n=1,2,...}$$

is a Cauchy sequence in M. Since M is complete, this sequence converges. Let p be its limit,

$$p := \lim_{n \longrightarrow \infty} A^n x .$$

From

$$d(Ap, A^n x) \le \lambda \cdot d(p, A^{n-1} x) \longrightarrow 0$$

we see

$$Ap = \lim_{n \longrightarrow \infty} A^n x = p.$$

Thus p is a fixed point of A. If p' is another fixed point, we get

$$d(p, p') = d(Ap, Ap') \le \lambda \cdot d(p, p'),$$

hence

$$(1-\lambda) \cdot d(p, p') = 0,$$

hence $d(p, p') = 0$, i.e., $p = p'$. The fixed point of A is unique. We have yet to estimate the distance $d(p, A^n x)$. From (1) we get, passing to the limit $n \longrightarrow \infty$,

$$d(p, A^m) \le d \cdot \sum_{i=m}^{\infty} \lambda^i = d \cdot \lambda^m \cdot \sum_{i=0}^{\infty} \lambda^i = d \cdot \lambda^m \cdot \frac{1}{1-\lambda}$$

QED.

A generatlization

Let M be a complete metric space, $\lambda < 1$ a real number and $\{A_n : M \longrightarrow M\}_{n=1,2,...}$ a sequence of maps such that

$$d(A_n x, , A_n y) \le \lambda \cdot d(x, y)$$

for arbitrary $x, y \in M$. In other words, the A_n form a sequence of contractions with a common contraction factor.

(i) Then for evey $x \in M$ the sequence

$$\{ x_n \}_{n=1,2,...}$$

such that

1. $x_1 = x$
2. $x_{n+1} = A_n x_n$

in converges to some point $p \in M$,

$$\lim_{n \longrightarrow \infty} x_n = p.$$

(ii) If, moreover, there is a map $A: M \longrightarrow M$ such that

$$Ax = \lim_{n \longrightarrow \infty} A_n x \text{ for every } x \in M,$$

then A is a contraction and the point p above is the only fixed point of A.

Proof. Assertion (i). Consider a second point, say $y \in M$, and write

$y_1 = y$

$y_{n+1} = A_n y_n$

Then

$$d(x_{n+1}, y_{n+1}) = d(A_n x_n, A_n y_n) \le \lambda \bullet d(x_n, y_n)$$

hence

$$d(x_{n+1}, y_{n+1}) \le \lambda^n \bullet d(x_1, y_1),$$

i.e.

$$d(x_n, y_n) \quad \le \lambda^{n-1} \bullet d(x_1, y_1),$$

In particular, for $y = x_2$ we get

$$d(x_n, x_{n+1}) \quad \le \lambda^{n-1} \bullet d(x_1, x_2).$$

For n, m > N one obtains

$$d(x_n, x_m) \le \sum_{\text{i between n and m}} d(x_i, x_{i+1}) \le d(x_1, x_2) \bullet \sum_{i=N}^{\infty} \lambda^i$$

Since the sum $\sum_{i=0}^{\infty} \lambda^i$ converges, the partial sums can be made arbitrarily small, i.e.

$$\{ x_n \}_{n=1,2,...}$$

is a Cauchy sequence in M. Since M is complete, this sequence converges. Let p be its limit,

$$p := \lim_{n \longrightarrow \infty} x_n .$$

Assertion (ii). From

$$d(A_n x, , A_n y) \le \lambda \bullet d(x, y)$$

we obtain, passing to the limit $n \longrightarrow \infty$,

$$d(Ax, , Ay) \le \lambda \bullet d(x, y),$$

i.e. A is a contraction with contraction factor λ. Moreover,

$$\begin{aligned} d(Ap, x_n) &\le d(Ap, A_n p) + d(A_n p, A_n x_n) + d(A_n x_n, x_n) \\ &\le d(Ap, A_n p) + \lambda \bullet d(p, x_n) + d(x_{n+1}, x_n) \end{aligned}$$

QED.

2.3.8 Lipschitz conditions

Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ be an open set and

$$X: U \times V \longrightarrow \mathbb{R}^n, (x, y) \mapsto X(x,y)$$

a function. One says that X satisfies with respect to x the Lipschitz condition with Lipschitz constant

$$C \in \mathbb{R},$$

if

$$|X(p, y) - X(q,y)| \le | p - q |$$

for arbitrary points $p, q \in U$ and arbitrary $y \in V$. Here $|x|$ denotes the euclidian norm, i.e.

$$|x| = \sqrt{\sum_{i=1}^{n} x_i^2} \text{ if } x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$$

Remarks

(i) If X is a C^1 function with bounded derivatives on U and U is convex[49],

$$\left| \frac{\partial X}{dx_i} \right| \le C,$$

then X satisfies a Lipschitz condition (with respect to x with Lipschitz constant n•C). To see this, consider the restriction of X to the segement connecting p and q,

$$\tilde{X}(t) = X(tp + (1-t)q).$$

Then

$$X(p,y) - X(q,y) = \tilde{X}(1,y) - \tilde{X}(0,y) = \int_0^1 \frac{d\tilde{X}}{dt}(t,y)\, dt$$

hence

$$| X(p,y) - X(q,y)| = \left| \int_0^1 \frac{d\tilde{X}(t, y)}{dt} dt \right|$$

$$\le \int_0^1 \left| \frac{d\tilde{X}(t, y)}{dt} \right| dt$$

$$= \int_0^1 \left| \sum_{i=1}^{n} \frac{\partial X(tp+(1-t)q\ , y)}{\partial x_i} \bullet (p_i - q_i) \right| dt \qquad \text{(chain rule)}$$

$$\le \sum_{i=1}^{n} |p_i - q_i| \int_0^1 \left| \frac{\partial X(tp+(1-t)q\ , y)}{\partial x_i} \right| dt$$

$$\le \sum_{i=1}^{n} |p_i - q_i| \int_0^1 C \bullet dt$$

$$\le \sum_{i=1}^{n} |p_i - q_i|$$

$$\le n \bullet C \bullet | p - q |$$

(ii) The first remark implies that the local existence an uniqueness theorem 2.1.5 follows from the Picard-Lindelöf theorem below.

2.3.9 Picard-Lindelöf theorem

Let

$$X: \mathbb{R} \times \mathbb{R}^n \longrightarrow T\mathbb{R}^n, (t, x) \mapsto X(t,x),$$

be a time dependent continuous vector field satisfying a Lipschitz condition with respect to x. Then every initial value problem

$$\frac{dx}{dt} = X(t, x), x(t_0) = x_0$$

---

[49] i.e., for every pair of points p, q ∈ U the segement connecting p and q is completely contained in U.

with $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^n$ has a solution which is unique in the sense that every two solutions coincide on the intersection of their domain of definitions.

This solution is a continuous function of t and the initial value x_0. Moreover, in case the vector field depends continuously upon certain parameters, the solution is also a continuous function of these parameters.

Remark
Note that the fact that this theorem is formulated for vector fields X which are defined everywhere on $\mathbb{R}^n$ does not matter. If X is defined only on a open subset of euclician space, one can restrict the field to a small open ε-neighbourhood of the given point x_0 and then use an appropriate map to identify this ε-neighbourhood with $\mathbb{R}^n$, for example in case ε = 1 one can use[50]

$$\{x \in \mathbb{R}^n \mid |x| < 1\} \longrightarrow \mathbb{R}^n, x \mapsto \frac{x}{1 - |x|}$$

(the inverse being $y \mapsto \frac{y}{1+|y|}$).

Alternatively, one may use a partitions of unity to exentend a given vector field (more precisely, its restrictions to compact subsets) to the whole euklidian space.

Proof. Step 1: Existence
Let C be the Lifschitz constant for the vector field X,

$$|X(t, x) - X(t, x)| \le C\cdot| x - x'|,$$

and let K be the compact set

$$K := \{(t, x) \in \mathbb{R}\times\mathbb{R}^n \mid | t_0 - t| \le \varepsilon, d(x_0, x) \le a \}$$

with positive real numbers ε and a. From 2.2.2 Example 3 we know, the set

$$M = C(K, \mathbb{R}^n)$$

of continuous functions

$$h: K \longrightarrow \mathbb{R}^n$$

is a complete metric space.

Consider the map

$$A: C(K, \mathbb{R}^n) \longrightarrow C(K, \mathbb{R}^n), \varphi \mapsto A\varphi,$$

such that

$$(A\varphi)(t, x) := x + \int_{t_0}^{t} X(t, \varphi(t,x))\, dt$$

For any two functions $\varphi, \psi \in C(K, \mathbb{R}^n)$ we get

$$(A\varphi - A\psi)(t, x) = \int_{t_0}^{t} (X(t, \varphi(t,x)) - X(t, \psi(t, x))\, dt$$

[50] This map may not preserve the Lipschitz condition, but is preserves differentiability.

$$|(A\varphi - A\psi)(t, x)| \le |\int_{t_0}^{t} |X(t, \varphi(t,x)) - X(t, \psi(t, x)| dt |$$

$$\le |\int_{t_0}^{t} C\cdot|\varphi(t, x) - \psi(t, x)| dt |$$

$$\le |\int_{t_0}^{t} C\cdot d(\varphi, \psi) dt |$$

$$= C\cdot d(\varphi, \varphi)\cdot | t - t_0 |$$

$$\le C\cdot\varepsilon\cdot d(\varphi, \psi),$$

hence

$$d(A\varphi, A\psi) \le C\cdot\varepsilon\cdot d(\varphi, \psi).$$

We see that for ε sufficiently small (i.e., for $C\cdot\varepsilon < 1$), the operation A is a contraction. But then A has precisely one fixed point. In other word, the initial value problem

$$\frac{dx}{dt} = X(x),\ x(t_0) = x_0 \qquad (1)$$

has a unique solution on the interval $(t_0 - \varepsilon, t_0 + \varepsilon)$.

Step 2: Uniqueness.
The uniqueness assertion is proved as in 2.2.3.

Step 3: Dependency upon the initial value.
By Step 1 the local solution is an Element of $C(K, \mathbb{R}^n)$, hence a continuous function of both, t and x_0.

Step 4: Dependency upon parameters.
This follows from the already proved part of the theorem applied to the initial value problem

$$\frac{dx}{dt} = X(t, x, \mu), \qquad x(t_0) = x_0$$

$$\frac{d\mu}{dt} = 0, \qquad \mu(t_0) = \lambda$$

(see the theorem of 2.1.7 and its proof). Alternatively, one can look at the Picard approximates of the solution of (1) and use their continuous dependency upon λ.

QED.

Remark

(i) Using in the above proof a modified distance, which depends in addition to the values of the functions also upon their derivatives, on can prove that the solution is a C^s function of t and x_0 , if the vector field is of class C^s. Below we will give an alternative proof for this fact.

(ii) The functions $A^n\varphi$ occuring in the above proof and approximating the solutions of the initial value problem are called Picard approximates of (1). The contraction mapping theorem gives us an extimate for their distance from the exact solution x(t),

$$d(x, A^n\varphi) \le \frac{\lambda^n}{1-\lambda}\cdot d(A\varphi, \varphi), \text{ with } \lambda = C\cdot\varepsilon.$$

2.3.10 Peano theorem

Let

$$X: \mathbb{R}\times\mathbb{R}^n \longrightarrow T\mathbb{R}^n$$

be a continuos vector field. Then every initial value problem

$$\frac{dx}{dt} = X(t, x), x(t_0) = x_0$$

with $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^n$ has a solution (which is a C^1 function defined in an interval containing t_0).

Remarks

(i) From the remark to 2.3.9 we see that the theorem is also true if X is a vector field which is only defined (and continuous) on an open subset of $\mathbb{R}^n$.

(ii) The continuous vector field X is boundet on compact subsets of $\mathbb{R}\times\mathbb{R}^n$, for example on the closure of the 1-neighbourhood of x_0. Restricting to this neighbourhood and identifying the latter with $\mathbb{R}^n$ we reduce to the case that the vector field is bounded, say

$$|X(t, x)| < C$$

for every $x \in \mathbb{R}^n$ and every $t \in \mathbb{R}$.

(iii) From the proof below one can deduce that the solution is (for bounded vector fields defined on the whole of $\mathbb{R}^n$) infinitely extentable.

Proof. Let $I = [a, b] \subseteq \mathbb{R}$ be any closed interval containting t_0 and denote by

$$C(I)$$

the set of continuous function $I \longrightarrow \mathbb{R}^n$. We have to find an element $x(t) \in C(I)$ which is a solution of the initial value problem. For this it will be sufficient to find functions $y(t), z(t) \in C(t)$ such that[51]

1. $y(t)$ satisfies the differential equation for $t_0 \le t$ and equals x_0 for $t \le t_0$.
2. $z(t)$ satisfies the differential equation for $t \le t_0$ and equals x_0 for $t_0 \le t$.

We restrict to the construction of y, the construction of z being similar. Thus we are looking for an element of C(t) satisfying

$$y(t) = x_0 + \int_{t_0}^{t} X(t, y(t))\, dt \text{ for } t \in I \text{ and } t_0 \le t. \qquad (1)$$

This function will be constructed as a limit of a sequence of approximating functions

$$y_\varepsilon(t) \in C(t).$$

Define

$$y_\varepsilon(t) = \begin{cases} x_0 & \text{for } t \le t_0 \\ x_0 + \int_{t_0}^{t} X(t, y_\varepsilon(t-\varepsilon))dt & \text{for } t_0 \le t \end{cases}$$

---

[51] The function we are looking for can than be obtained as

$$x(t) := \begin{cases} y(t) \text{ for } t_0 \le t \\ z(t) \text{ for } t \le t_0 \end{cases}$$

This gives a correctly defined differentiable function on the whole of the interval I: from the definition of y_ε for $t \le t_0$ one obtaines the definition for $t \le t_0 + \varepsilon$, then the definition for $t \le t_0 + 2\varepsilon$, and so on.

From $|X(t, x)| < C$ we see that

$$| y_\varepsilon(t') - y_\varepsilon(t'') | \le C\cdot |t' - t''|, \tag{2}$$

i.e. the functions $y_\varepsilon(t)$ satisfy a Lipschitz condition.

Now consider the sequence of functions

$$\{y_{1/n}(t)\}_{1,2,3,...}$$

By the theorem of Acoli-Arzelà there is a converging subsequence, say $\{y_{\varepsilon_n}(t)\}$, which converges uniformly and hence has a continuous limit, say y(t). Write

$$y^{(n)}(t) := y_{\varepsilon_n}(t).$$

Then

$$y^{(n)}(t) = x_0 + \int_{t_0}^{t} X(t, y^{(n)}(t - \varepsilon_n))\, dt \text{ for } t \in I \text{ and } t_0 \le t \tag{3}$$

Note that

$$| y^{(n)}(t - \varepsilon_n) - y(t)| \le | y^{(n)}(t - \varepsilon_n) - y^{(n)}(t)| + |y^{(n)}(t) - y(t)|$$
$$\le C\cdot\varepsilon_n + |y^{(n)}(t) - y(t)|$$

The sequence $\{ y^{(n)}(t - \varepsilon_n) \}_{1,2,3,...}$ also converges uniformly to y(t). Form (3) we obtain, passing to the limit

$$y (t) = x_0 + \int_{t_0}^{t} X(t, y (t))\, dt \text{ for } t \in I \text{ and } t_0 \le t.$$

Thus, y(t) is the function we are looking for.

QED.

2.3.11 The equations of variation

Let

$$X: \mathbb{R}\times\mathbb{R}^n \longrightarrow T\mathbb{R}^n$$

be a time dependent C^1 vector field. Then to the differential equation

$$\frac{dx}{dt} = X(t, x) \tag{1}$$

one can attach the differential equation

$$\frac{dy}{dt} = \frac{\partial X(t,x)}{\partial x}\cdot y \tag{2}$$

where $\frac{\partial X(t,x)}{\partial x}$ denotes the Jacobian matrix of the map

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n, x \mapsto X(t,x).$$

and y is considered to be a function taking values in $\mathbb{R}^{n\times n}$. The equation (1) and (2) together are called <u>equations of variation</u> of (1). Note that these are linear with respect to $y \in \mathbb{R}^n$.

Remarks

(i) The right hand side of the second equation of variation

$$\frac{dy}{dt} = \frac{\partial X(t,x)}{\partial x} \bullet y$$

is in general not a C^1 function: the Jacobian consits of the first derivatives of the coordinates of X(t,x). So we only can say, that it is a continuous function of x.

However, as a function of y it is a C^1 function[52]. Therefore, given a continuous function[53] x(t) defined on an interval containing t_0, the initial value problem

$$\frac{dy}{dt} = \frac{\partial X(t,x(t))}{\partial x(t)} \bullet y\,, \qquad y(t_0) = y_0\,, \tag{3}$$

has a unique (local) solution depending continuously upon t and the initial value y_0 (and possible continuous parameters of the vector field)[54]

(ii) If the function x(t) of (ii) is a solution

$$x(t) = \varphi(t, x)$$

of the initial value problem

$$\frac{dx}{dt} = X(t, x),\ x(t_0) = x,$$

then one can show that the solution y(t) of (3) in case y_0 is the identity matrix, is nothing but the Jacobian matrix of φ(t, x),

$$y(t) = \frac{\partial\varphi(t,x)}{\partial x}$$

In particular, φ(t, x) is a C^1 function of t and the initial values x. For details, see the book of Arnol'd [1].

2.3.12 Differentiability of the local phase flow

Let

$$X\colon \mathbb{R}\times\mathbb{R}^n \longrightarrow T\mathbb{R}^n$$

be a time dependent C^r vector field with $r \geq 1$. Then the local phase flows

$$(t, x) \mapsto g^t x$$

associated with the differential equation

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

are C^r maps.

Proof. Step 1: r =1.
This is just the assertion of Remark 2.3.11 (ii). Since we did not prove this remark, let us illustrate the situation in proving the claim under the stronger assumption that the vector field X is of class C^2.

Consider the equations of variation

---

[52] It is linear as a function of y.

[53] for example, a local solution of (1).

[54] The right hand side of the differential equation satisfies on every compact set $K \subseteq \mathbb{R}\times\mathbb{R}^n$ a Lipschitz condition with respect to y, for, the derivatives of $\frac{\partial X(t,x(t))}{\partial x(t)} \bullet y$ with respect to the coordinates of y are the colums of the matrix $\frac{\partial X(t,x(t))}{\partial x(t)}$, which are continousus, hence boundet on the compact set K. One has even a common Lipschitz constant for the set of all continuous functions x(t) bounded by a common constant.

$$\frac{dx}{dt} = X(t, x)$$
$$\frac{dy}{dt} = \frac{\partial X(t,x)}{\partial x} \bullet y$$

Since X(t, x) is a C^1 map, it satisfies a Lipschitz condition on each compact subset of the extended phase space, hence the Picard-Lindelöff theorem 2.3.9 applies to this system, so that every initial value problem has a unique local solution.

Consider the Picard approximatioins $\varphi_n(t,x)$, $\psi_n(t,x)$ for the equations of variation starting with constant functions

$$\varphi_0(t,x) = x$$

$$\psi_0(t,x) = Id = \begin{pmatrix} 1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & .. & 1 \end{pmatrix}$$

i..e.

$$\varphi_{n+1}(t,x) = x + \int_{t_0}^{t} X(t, \varphi_n(\tau,x))\, d\tau$$

$$\psi_{n+1}(t,x) = Id + \int_{t_0}^{t} \frac{\partial X(t,\varphi_n(t, x))}{\partial x} \bullet \psi_n(\tau,x)\, d\tau$$

From the definition of $\varphi_0(t)$ and $\psi_0(t)$ we see that

$$\psi_0(t,x) = \begin{pmatrix} 1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & .. & 1 \end{pmatrix} = \frac{\partial \varphi_0(t,x)}{\partial x}$$

is the Jacobian matrix of $\varphi_0(t)$.[55] But then the same holds for the other approximations,

$$\psi_n(t,x) = \frac{\partial \varphi_n(t,x)}{\partial x} \qquad (1)$$

As we know[56] these approximations converge for t in a sufficiently small interval to a local solution $\varphi(t, x)$, $\psi(t, x)$ of the initial value problem

$$x(t_0,) = x$$
$$y(t_0,) = Id$$

and these solutions are continuous functions of t and x. From (1) we see, passing to the limit, that

$$\psi(t,x) = \frac{\partial \varphi(t,x)}{\partial x} \qquad (1)$$

i.e. the second component of the solution is the Jacobian matrix of the first component. In particular, this Jacobian matrix exists and is continuous. We have proved that $\varphi(t, x)$ is a C^1 function of x (hence of x and t).

Step 2. $r > 1$.

We may assume that the theorem is true with r replaced by r-1. By assumption, the system

---

[55] The entry in position (i,j) is $\frac{\partial x_i}{\partial x_j} = \delta_{ij}$.

[56] see also the remarks of 2.3.11.

$$\frac{dx}{dt} = X(t, x)$$
$$\frac{dy}{dt} = \frac{\partial X(t,x)}{\partial x} \bullet y$$

satisfies the induction hypothesis. Hence its local phase flows are C^{r-1} functions, i.e. the local solutions

$\varphi(t, x)$
$\psi(t, x)$

of the initial value problem $x(t_0) = x$, $y(t_0) = Id$ are C^{r-1} functions of x. Since $\psi(t, x)$ is the matrix of derivatives of $\varphi(t, x)$, this means that $\varphi(t, x)$ is a C^r function with respect to x (hence of x and t)[57].

QED.

Appendix

The directional derivative of a function

Let

$$f: U \longrightarrow \mathbb{R}$$

be a real valued function defined on an open set $U \subseteq \mathbb{R}^n$ of real n-space,

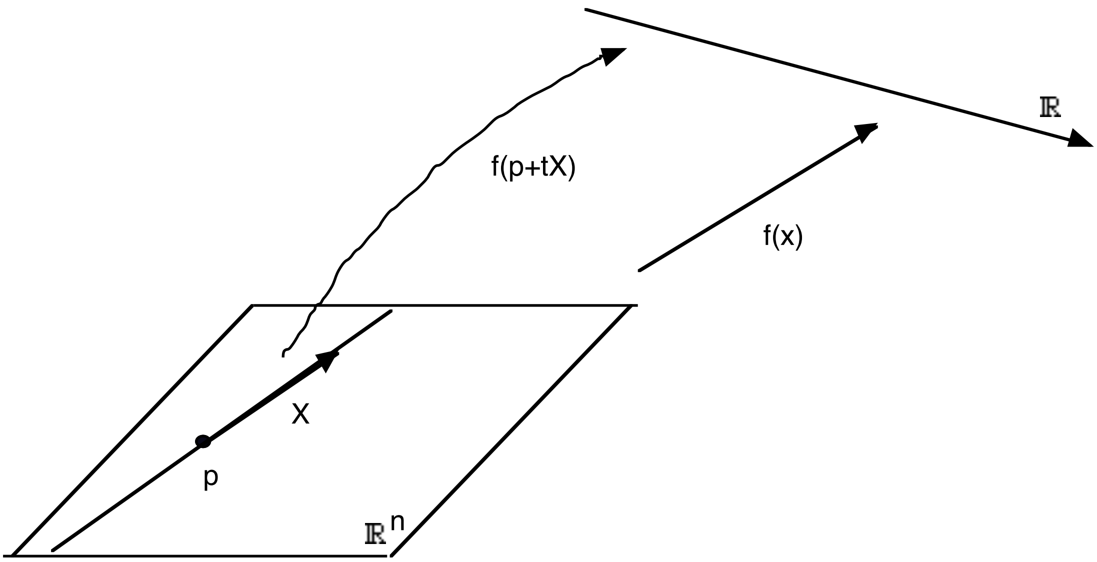
$$p \in U$$

be a point and

$$X \in \mathbb{R}^n$$

be a vektor at p. Then

$$X(f) = X_p(f) := \frac{df(p+t\bullet X)}{dt}\Big|_{t=0}$$



is called the derivative of f at p in the direction of X. If it exists, f is called differentiable in the the direction of X at p. The function f is called differentiable at p if it is differentiabe in every direction (at p) . The function is called differentiable if it is differentiable at ervery point p of its domain U of definition.

---

[57] Note that $\frac{\partial}{\partial t}\varphi(t, x) = X(t, \varphi(t, x))$ is a C^{s-2} function (by induction hypothesis).

The function f is called continuously differentiable, if it is differentiable and the derivvatives $X_p(f)$ are continuous functions of the point $p \in U$ for every vector X.

The directional derivative of a vector valued function

A vector valued function

$$f: U \longrightarrow \mathbb{R}^m,\ x \mapsto f(x) = \begin{pmatrix} f_1(x) \\ \dots \\ f_n(x) \end{pmatrix},$$

is called (continuously) differentiable, if every coordinate function f_i is (continuously) differentiable. In this case,

$$X(f) := \begin{pmatrix} X(f_1) \\ \dots \\ X(f_n) \end{pmatrix}$$

is called the derivative of f at p in the direction of X.

Example

The derivative of f in the direction of the i-th standard unit vector

$$e_i = \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$$

whose i-th coordinate is 1 and whose other coordinates are zero, is denoted

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial x_i}\Big|_p$$

i.e., for every f defined at the point p,

$$e_i(f) = \frac{df(p+te_i)}{dt}\Big|_{t=0} = \frac{df(p_1,\dots,p_{i-1}, p_i+t, ,p_{i+1},\dots,p_n)}{dt}\Big|_{t=0} = \frac{\partial f}{\partial x_i}\Big|_p = \frac{\partial f}{\partial x_i}(p).$$

The Jacobian matrix of a vector valued function

The matrix

$$\frac{\partial f}{\partial x}(p) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \dots & \frac{\partial f_1}{\partial x_n}(p) \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(p) & \dots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix} = (\frac{\partial f_i}{\partial x_j}(p)) = (\frac{\partial f_i}{\partial x_j}(p))_{i=1,\dots,n,j=1,\dots m}$$

is called Jacobian matrix ore functional matrix of f at p. Its determinant is called Jacobian ore functional determinant. and is denoted

$$\left| \frac{\partial f}{\partial x} \right| (p) = \det \frac{\partial f}{\partial x}(p).$$

The differential of a vector valued function

For every differentiable vector valued function

$$f: U \longrightarrow \mathbb{R}^m,\ x \mapsto f(x) = \begin{pmatrix} f_1(x) \\ \dots \\ f_n(x) \end{pmatrix},$$

defined on an open set $U \subseteq \mathbb{R}^n$ and every point $p \in U$, matrix multiplication by the Jacobian matrix defined a linear map

$$d_p f: \mathbb{R}^n \longrightarrow \mathbb{R}^m,\ x \mapsto \frac{\partial f(p)}{\partial x} \bullet x \text{ , i.e., } \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \dots & \frac{\partial f_1}{\partial x_n}(p) \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(p) & \dots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix} \bullet \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix},$$

which is called the linearization or the differential of f at p. Soon we will see that this map also exists in the more general context of manifolds, where it looks like this:

Remarks

(i) Given a C^r map of C^r manifolds

$$f: M \longrightarrow M'$$

(with $r \geq 2$) and a point $p \in M$, than there is a linear map of real vector spaces

$$d_p f: T_p(M) \longrightarrow T_{p'}(M')$$

with $p' := f(p) \in M'$, also called differential of f at p.

Using this notation, two important theorems (which will soon be proved in the analyis course) can be translated as follows into the language of manifolds.

(ii) Inverse function theorem. Assume that $d_p f$ is an isomorphism, then f is locally at p a diffeomorphism, i.e. there are open sets

$$U \subseteq M \text{ and } U' \subseteq M'$$

such that

1. $p \in U$, $p' \in U'$, $f(U) = U'$.
2. $f|_U : U \longrightarrow U'$ is a diffeomorphism.

(iii) Implicite function theorem. Assume that $d_p f$ is surjective. Then

$$f^{-1}(p')$$

is locally at p a submanifold of M, i.e. there is a chart araound p,

$$x: U \longrightarrow V \subseteq \mathbb{R}^n,\ u \mapsto \begin{pmatrix} x_1(u) \\ \dots \\ x_n(u) \end{pmatrix},\ p \in U,$$

such that $U \cap f^{-1}(p')$ is defined by a system of linear (!) equation, say

$$U \cap f^{-1}(p') = \{u \in U \mid x_1(u) = \dots = x_m(u) = 0\}.$$

Chain rule

Let

$$f\colon U \longrightarrow \mathbb{R}^m,\ x = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} f_1(x) \\ \dots \\ f_m(x) \end{pmatrix},$$

be a continuously differentiable function defined in an open set $U \subseteq \mathbb{R}^n$ of n-space. Moreover, let

$$\varphi\colon I \longrightarrow U,\ t \mapsto \begin{pmatrix} \varphi_1(t) \\ \dots \\ \varphi_n(t) \end{pmatrix},$$

be a continuously differentiable function defined on an open set $I \subseteq \mathbb{R}$. Then, vor every $t_0 \in I$,

$$\frac{df(\varphi(t))}{dt}\Big|_{t=t_0} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\varphi(t_0)) \cdot \frac{d\varphi_i(t)}{dt}(t_0)$$

For short,

$$\frac{d(fg)}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \cdot \frac{d\varphi_i}{dt}$$

A trivial example

$f(x_1,x_2) = x_1^2 + x_2^2$

$\varphi_1(t) = \sin t$

$\varphi_2(t) = \cos t$

Then

$$\frac{\partial f}{\partial x_1} = 2x_1$$

$$\frac{\partial f}{\partial x_2} = 2x_2$$

Using the chain rule we get

$$\begin{aligned} \frac{df(\varphi_1(t),\varphi_2(t))}{dt} &= \frac{\partial f}{\partial x_1}(\text{sint } t, \text{cost } t)\frac{d\sin t}{dt} + \frac{\partial f}{\partial x_2}(\text{sint } t, \text{cost } t)\frac{d\cos t}{dt} \\ &= 2\sin t \cdot \cos t + 2\cos t \cdot (-\sin t) \\ &= 0 \end{aligned}$$

Of course we know that $f(\varphi_1(t), \varphi_2(t))$ ist constant. Hence its derivative should be zero.

A complex example (related to the first problem of series 2)

Let $U \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^n$ be open sets and

$$f: U \longrightarrow V, x \mapsto f(x) = \begin{pmatrix} f_1(x) \\ \dots \\ f_n(x) \end{pmatrix}$$

$$g: V \longrightarrow U, y \mapsto g(y) = \begin{pmatrix} g_1(y) \\ \dots \\ g_m(y) \end{pmatrix},$$

be continuously differentiable functions which are inverse to each other,

$$f\circ g = id: V \longrightarrow V \text{ and } g\circ f = id: U \longrightarrow U.$$

In particular,

$$g_i(f(x)) = x_i \text{ for every i.}$$

Therefore

$$\delta_{ij} = \frac{\partial x_i}{\partial x_j}(p) = \frac{\partial g_i(f(x_1;...;x_m))}{\partial x_j}(p) = \sum_{\ell=1}^{m} \frac{\partial g_i(y)}{\partial y_\ell}(f(p)) \cdot \frac{\partial f_\ell(x)}{\partial x_j}(p)$$

hence

$$Id = (\delta_{ij}) = \frac{\partial g}{\partial y}(f(p)) \cdot \frac{\partial f}{\partial x}(p)$$

For short

$$Id = \frac{\partial g}{\partial y} \cdot \frac{\partial f}{\partial x}$$

and similarly

$$Id = \frac{\partial f}{\partial x} \cdot \frac{\partial g}{\partial y}$$

In case m = n one obtains for the determinants

$$1 = |\frac{\partial f}{\partial x}| \cdot |\frac{\partial g}{\partial y}| = |\frac{\partial g}{\partial y}| \cdot |\frac{\partial f}{\partial x}|$$

Problem
Is it possible that m and n are different ?

Solution
To fix notation, we may assume that

$$m < n.$$

Consider the $n\times n$-matrix

$$Id = \frac{\partial f}{\partial x} \cdot \frac{\partial g}{\partial y} \tag{1}$$

We cannot pass to the product of the determinants, since the number m of colums of the first matrix $\frac{\partial f}{\partial x}$ and the number m of rows of the second matrix are too small. We can work around this problem adding colums to the first and rows to the second matrix which consits entirely of zeroes. Denote the resulting $n\times n$-matrix again by $\frac{\partial f}{\partial x}$ and $\frac{\partial g}{\partial y}$. This operation of adding zero entries does not change the product on the right of (1). But now we can pass to the determinants:

$$1 = |\frac{\partial f}{\partial x}| \cdot |\frac{\partial g}{\partial y}|$$

In particular

$$|\frac{\partial f}{\partial x}| \neq 0 \text{ and } |\frac{\partial g}{\partial y}| \neq 0.$$

In particular, these n×n matrices cannot contain columns or rows consisting entirely of zeros. There the number of zero columns or rows added to these matrices must be zero. We not add anything. But then the original matrices are already n×n matrices, which means

$$m = n.$$

Bibliography

Arrowsmith, D.K., Place, C.M.

[1] Ordinary differential equations with applications, Chapman and Hall, London 19982

Arnol'd, Vladimir

[1] Ordinary differential equations, Springer, Berlin 1992

Walter, Wolfgang

[1] Gewöhnliche Differentialgleichungen, Eine Einführung, Springer, Berlin 200

Index

Contents