UCT Mathematics

Václav Mácha

Once, these lecture notes will contain mathematical knowledge needed to pass through math exam at the University of Chemistry and Technology. They are released online and they are available for free. On the other hand, my work on this text is still not finished and thus it may contain some mistake. In case you find any, let me know.

Contents

1 Vectors and vector spaces

1.1 Vector spaces

Definition 1.1. *A set V endowed with operations $+$ (sum) and $.$ (multiplication by a real number) which satisfy $u + v \in V$ for all $u, v \in V$ and $\alpha.u \in V$ for all $u \in V$ and $\alpha \in \mathbb{R}$ is called* vector space *(or a linear space) if the following properties are true:*

i) $u + v = v + u$ *for all* $u, v \in V$,

ii) $u + (v + w) = (u + v) + w$ *for all* $u, w \in V$,

iii) $\exists 0 \in V$ *for which it holds that* $0 + v = v$ *for all* v,

iv) for all v *there is an element* $-v$ *such that* $v + (-v) = 0$,

v) $\alpha.(\beta.v) = (\alpha.\beta).v$ *for all* $\alpha, \beta \in \mathbb{R}$ *and for all* $v \in V$,

vi) $1.v = v$ *for all* $v \in V$,

vii) $(\alpha + \beta).v = \alpha.v + \beta.v$ *for all* $\alpha, \beta \in \mathbb{R}$ *and for all* $v \in V$,

viii) $\alpha.(v + w) = \alpha.v + \alpha.w$ *for all* $\alpha \in \mathbb{R}$ *and for all* $v, w \in V$.

An element of the vector space is called vector.

Examples:

- The space of ordered pairs of real numbers $(u, v) \in \mathbb{R}^2$ with summation and product defined as
$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2), \quad \alpha(u_1, v_1) = (\alpha u_1, \alpha v_1)$$
for all $(u_1, v_1), (u_2, v_2) \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$ is a vector space.

- In general, all ordered $n-$tuples of real numbers $(u_1, u_2, \ldots, u_n) \in \mathbb{R}^n$ for $n \in \mathbb{N}$ form a vector space.

- The set S of all $(x, y) \in \mathbb{R}^2$ satisfying
$$x + 2y = 0 \tag{1}$$
is a vector space. Since this is a subset of the vector space mentioned above, it is enough to verify that $((x_1, y_1), (x_2, y_2) \in S) \Rightarrow (x_1 + x_2, y_1 + y_2) \in S$ and $(\alpha \in \mathbb{R} \& (x, y) \in S) \Rightarrow (\alpha x, \alpha y) \in S$. So let (x_1, y_1) and (x_2, y_2) satisfy (1). Then $(x_1 + x_2, y_1 + y_2)$ also satisfies (1) since
$$x_1 + x_2 + 2(y_1 + y_2) = x_1 + 2y_1 + x_2 + 2y_2 = 0.$$
Next, let $\alpha \in \mathbb{R}$ be arbitrary number and let (x, y) satisfies (1). Then
$$\alpha x + 2\alpha y = \alpha(x + 2y) = 0$$
and $(\alpha x, \alpha y) \in S$.

- On the other hand, the set S of all pairs $(x, y) \in \mathbb{R}^2$ satisfying
$$x + 2y = 1$$
is not a vector space. For example, a zero vector $(0, 0)$ does not belong to S and the third property from the definition of vector space is not fulfilled.

- The set of polynomials is a vector space.
- The set of polynomials of degree 2 is not a vector space. In particular, a zero polynomial does not belong to this set as the zero polynomial has not degree 2.
- On the other hand, the set of polynomials of degree 0, 1 or 2 is a vector space.

Definition 1.2. *Let V be a vector space and let $S \subset V$ be such that*

i) $\forall s_1, s_2 \in S$, $s_1 + s_2 \in S$ *and*

ii) $\forall \alpha \in \mathbb{R}$ *and* $\forall s \in S$ *we have* $\alpha s \in S$.

Then S itself is a vector space and we say that S is a subspace *of V. If S is nonempty and $S \neq V$ then we will say that S is a proper subspace.*

Definition 1.3. *Let V be a vector space, $n \in \mathbb{N}$ and $\{u_i\}_{i=1}^n \subset V$. Their* linear combination *is any vector w of the form*

$$w = \sum_{i=1}^{n} \alpha_i u_i$$

where α_i are real numbers.

Examples:

- Consider a vector space $\mathbb{R}^3$. The vector $(2, 5, 3)$ is a linear combination of $(1, 1, 0)$ and $(0, 1, 1)$ because
$$(2, 5, 3) = 2(1, 1, 0) + 3(0, 1, 1).$$
- On the other hand, $(0, -2, 1)$ is not a linear combination of $(1, 1, 0)$ and $(0, 1, 1)$. Indeed, if it was, then there would be two numbers α and β such that
$$(0, -2, 1) = \alpha(1, 1, 0) + \beta(0, 1, 1).$$
This equation can be rewritten as a system
$$\begin{aligned} 0 &= \alpha \\ -2 &= \alpha + \beta \\ 1 &= \beta \end{aligned}$$
and we deduce that it is impossible to find α and β such that these equations are fulfilled.

Definition 1.4. *The set of all linear combinations of $v_1,\ v_2, \ldots,\ v_n$ is called a* linear span *of a set $\{v_1, v_2, \ldots, v_n\}$. Precisely,*

$$\mathrm{span}\{v_1, v_2, \ldots, v_n\} = \left\{ \sum_{i=1}^{n} \alpha_i v_i,\ \alpha_i \in \mathbb{R} \right\}.$$

Definition 1.5. *Vectors $v_1, v_2, \ldots, v_n \in V$ are said to be* linearly dependent *if the equation*

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

has a nontrivial solution (i.e. a solution $\alpha_1,\ \alpha_2, \ldots,\ \alpha_n$ where at least one coefficient is zero). Vectors $v_1, v_2, \ldots, v_n \in V$ are linearly independent *if the equation*

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

has only solution $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

Definition 1.6. *Let $V = \text{span}\{v_1, v_2, \dots, v_n\}$. Then we say that $\{v_1, v_2, \dots, v_n\}$ generates V.*

Observation 1.1. *Let $v_1, v_2, \dots, v_n$ be linearly dependent. Then one of the vectors is a linear combination of the remaining vectors. Precisely, there is $i \in \{1, \dots, n\}$ such that $v_i \in \text{span}\{\{v_1, v_2, \dots, v_n\} \setminus \{v_i\}\}$.*

Proof. According to assumptions, there is $i \in \{1, \dots, n\}$ such that

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$$

has a solution with $\alpha_i \neq 0$. Assume, without lost of generality, that $i = 1$. We may rearrange the equation as

$$v_1 = -\frac{\alpha_2}{\alpha_1} v_2 - \frac{\alpha_3}{\alpha_1} v_3 - \ldots - \frac{\alpha_n}{\alpha_1} v_n.$$

□

Corollary 1.1. *Let $v_1 \in \text{span}\{v_2, \dots, v_n\}$. Then*

$$\text{span}\{v_2, \dots, v_n\} = \text{span}\{v_1, v_2, \dots, v_n\}.$$

Proof. Clearly, $\text{span}\{v_2, \dots, v_n\} \subset \text{span}\{v_1, v_2, \dots, v_n\}$. Next, let

$$v = \sum_{i=1}^{n} \alpha_i v_i.$$

Since $v_1 = \sum_{i=2}^{n} \beta_i v_i$ for some $\beta_i \in \mathbb{R}$, we get

$$v = \sum_{i=2}^{n} (\alpha_i + \alpha_1 \beta_i) v_i$$

and $v \in \text{span}\{v_2, \dots, v_n\}$. □

Definition 1.7. *Let $\{v_1, \dots, v_n\}$ be a set of linearly independent vectors that generates V. Then $\{v_1, \dots, v_n\}$ is a* basis *of V.*

Theorem 1.1. *Every two basis of a vector space V has the same number of elements.*

Definition 1.8. *We say that V is of dimension $n \in \mathbb{N}$ iff every basis has n elements.*

Examples:

- The set $\{(1,0), (0,1)\} \subset \mathbb{R}^2$ is a basis. Indeed, every vector $(a, b) \in \mathbb{R}^2$ can be written as $a(1,0) + b(0,1)$. Moreover, the vectors are linearly independent since $\alpha_1(1,0) + \alpha_2(0,1) = 0$ has only the trivial solution. Thus, the dimension of $\mathbb{R}^2$ is 2.
- Vectors $\{1, x, x^2\}$ form a basis of a vector space containing polynomials of degree at most two. The dimension of this vector space is thus 3.

1.2 Some exercises

:

1. Let (a, b) be a multiple of (c, d) with $abcd \neq 0$. Show that (a, c) is a multiple of (b, d).
2. Consider $\mathbb{R}^2$ with sum defined as follows: $(a, b) + (c, d) = (a + d, b + c)$ and with a usual multiplication. Is such space a vector space?

2 Matrices

2.1 Notions

Definition 2.1. *A* matrix *is a table of numbers arranged in rows and columns. Namely, let m, n be natural numbers. Then*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{i=1,j=1}^{m,n}$$

The matrix A has $m-$rows and $n-$columns. The matrix A is said to be of type (m, n).

Example A matrix

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & 2 & -1 \end{pmatrix}$$

has two rows and three columns and it is of type $(2, 3)$ (or it is of type two by three).

Operations with matrices Let $A = (a_{ij})_{i=1,j=1}^{m,n}$ and $B = (b_{ij})_{i=1,j=1}^{m,n}$ be two matrices of the same type. Then we define

$$A + B = (a_{ij} + b_{ij})_{i=1,j=1}^{m,n}.$$

Let $\alpha \in \mathbb{R}$. Then $\alpha A = (\alpha a_{ij})_{i=1,j=1}^{m,n}$.
For a matrix $A = (a_{ij})_{i=1,j=1}^{m,n}$ we define a transpose matrix A^T as

$$A^T = (a_{ji})_{j=1,i=1}^{n,m}$$

Let A be of type (m, n) and B be of type (n, p). Then $C := AB$ of type (m, p) is defined as

$$C = (c_{ij})_{i=1,j=1}^{m,p}$$

where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Example

- $$\begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 0 & 1 & -2 \end{pmatrix} + \begin{pmatrix} 2 & 2 & 2 & -5 \\ 1 & 1 & -3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 4 & -5 \\ 1 & 1 & -2 & 2 \end{pmatrix}.$$

- $$3 \begin{pmatrix} 1 & \frac{1}{2} \\ 2 & 2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & \frac{3}{2} \\ 6 & 6 \\ -9 & 3 \end{pmatrix}$$

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

  or

  $$\begin{pmatrix} 1 & 1 & 3 \\ 2 & -1 & 1 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ 1 & -1 \\ 3 & 1 \end{pmatrix}$$

or

$$\begin{pmatrix}3 & -1 & -1 & 0\end{pmatrix}^T = \begin{pmatrix}3\\-1\\-1\\0\end{pmatrix}$$

- $$\begin{pmatrix}-1 & 1 & 0\\2 & 0 & 1\end{pmatrix}\begin{pmatrix}2 & -1\\-1 & -1\\0 & 1\end{pmatrix} = \begin{pmatrix}-3 & 0\\4 & -1\end{pmatrix}.$$

Remark 2.1. *Matrices of a given type (m, n) forms a vector space.*

Remark 2.2. *Warning:*

$$AB \neq BA.$$

Definition 2.2. *A* rank of matrix *A is a dimension of vector space generated by its columns. It is denoted by $rankA$.*

Observation 2.1. *It holds that $rankA = rankA^T$.*

Definition 2.3. *Elementary transformation of a matrix is*

- *scaling the entire row with a nonzero real number or*
- *interchanging the rows within a matrix or*
- *adding $\alpha-$multiple of one row to another for an arbitrary $\alpha \in \mathbb{R}$.*

Let A arise from B by an elementary transformation. Then we write $A \sim B$.

Example

$$\begin{pmatrix}2 & 1\\-1 & 2\end{pmatrix} \sim \begin{pmatrix}-1 & 2\\2 & 1\end{pmatrix} \sim \begin{pmatrix}-2 & 4\\2 & 1\end{pmatrix} \sim \begin{pmatrix}-2 & 4\\6 & -7\end{pmatrix}.$$

Definition 2.4. *A* leading coefficient *of a row is the first non-zero coefficient in that row. We say that matrix A is in* echelon form *if the leading coefficient (also called the pivot) of a nonzero row is always strictly to the right of the leading coefficient of the row above it.*

Example Consider following matrices:

$$A = \begin{pmatrix}-1 & -1 & 3 & 0\\0 & 0 & 2 & 1\\0 & 0 & 0 & -1\\0 & 0 & 0 & 0\end{pmatrix} \qquad B = \begin{pmatrix}-1 & -1 & 3 & 0\\0 & 2 & 2 & 1\\1 & 0 & -1 & -1\\0 & 0 & 0 & 3\end{pmatrix}$$

The matrix A is in echelon form whereas the matrix B is not in echelon form.

Observation 2.2. *Let A be in echelon form. Then its rank is equal to the number of non-zero rows.*

The Gauss elimination method

The Gauss elimination method is a sequence of elementary transformations which transform a given matrix A into an echelon form. As an example, we take a matrix

$$A = \begin{pmatrix}2 & 2 & -2\\4 & 1 & 0\\5 & 2 & -1\end{pmatrix}.$$

In the first step, we use elementary transformation in order to get rid of 4 in the second row and 5 in the last row. So we add (-1) times the first row to the second and $-5/2$ times the first row to the last one. We get

$$A \sim \begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \\ 0 & -3 & 4 \end{pmatrix}.$$

Next, we want to eliminate the second element in the last row. In order to do so, we add (-1) times the secon row to the last one to get

$$\begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \\ 0 & -3 & 4 \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \\ 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 2 & 2 & -2 \\ 0 & -3 & 4 \end{pmatrix}.$$

Here we use the fact that the zero row can be ommitted without any serious consequence.

As a remark we want to point out that A has a rank two and that means that the vectors $(2, 2, -2)$, $(4, 1, 0)$ and $(5, 2, -1)$ are linearly dependent.

2.2 Systems of linear equations

Systems of equations

We are going to deal with system of m linear equations with n unknowns $x_1,\ x_2, \ldots,\ x_n$.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots &= \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

We use notation $x = (x_1, x_2, \ldots, x_n)$, $b = (b_1, b_2, \ldots, b_n)$ and $A = (a_{ij})_{i=1,j=1}^{mn}$. Then the above system may be rewritten as

$$Ax^T = b^T.$$

The system of equations will be represented by an augmented matrix – i.e. a matrix $(A|b^T)$ where $A = (a_{i,j})_{i=1,j=1}^{mn}$ and b^T is the column on the right hand side. For example, a system of equations

$$\begin{aligned} 2x + 5y &= 10 \\ 3x + 4y &= 24 \end{aligned}$$

is represented by an augmented matrix

$$\begin{pmatrix} 2 & 5 & | & 10 \\ 3 & 4 & | & 24 \end{pmatrix}.$$

Such matrix consists of two parts – matrix $A = \begin{pmatrix} 2 & 5 \\ 3 & 4 \end{pmatrix}$ and a vector of right hand side $b = (10, 24)$. Let solve the system by Gauss elimination:

$$\begin{pmatrix} 2 & 5 & | & 10 \\ 3 & 4 & | & 24 \end{pmatrix} \sim \begin{pmatrix} 6 & 15 & | & 30 \\ 3 & 4 & | & 24 \end{pmatrix} \sim \begin{pmatrix} 6 & 15 & | & 30 \\ 6 & 8 & | & 48 \end{pmatrix} \sim \begin{pmatrix} 6 & 15 & | & 30 \\ 0 & -7 & | & 18 \end{pmatrix}$$

The last row of the last matrix represent an equation

$$-7y = 18 \;\Rightarrow\; y = -\frac{18}{7}.$$

The first row of the last matrix represent

$$6x + 15y = 30$$

and once we plug there $y = -\frac{18}{7}$ we deduce

$$x = \frac{80}{7}.$$

Theorem 2.1 (Frobenius)**.** *A system of linear equations has solution if and only if* $rank A = rank(A|b^T)$.

Example: Solve

$$\begin{aligned} -x + y + z &= 0 \\ 2y + x + z &= 1 \\ 2z + 3y &= 2. \end{aligned}$$

We have

$$\begin{pmatrix} -1 & 1 & 1 & | & 0 \\ 1 & 2 & 1 & | & 1 \\ 0 & 3 & 2 & | & 2 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 1 & | & 0 \\ 0 & 3 & 2 & | & 1 \\ 0 & 3 & 2 & | & 2 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 1 & | & 0 \\ 0 & 3 & 2 & | & 1 \\ 0 & 0 & 0 & | & 1 \end{pmatrix}$$

and, according to the Frobenius theorem, there is no solution to the given system. Let us emphasize that the last row represents an equation

$$0x + 0y + 0z = 1.$$

Example Let find all solutions to the system

$$\begin{aligned} 2x + y - z &= 3 \\ x - 2y + 3z &= -1 \end{aligned}$$

We use the Gauss elimination in order to deduce

$$\begin{pmatrix} 2 & 1 & -1 & | & 3 \\ 1 & -2 & 3 & | & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 3 & | & -1 \\ 2 & 1 & -1 & | & 3 \end{pmatrix} \sim \begin{pmatrix} \color{red}{1} & -2 & 3 & | & -1 \\ 0 & \color{red}{5} & -7 & | & 5 \end{pmatrix}$$

The red terms are the leading terms. The corresponding unknowns should be expressed by others. The unknown which does not have a corresponding leading term should be chosen as a parameter. Here we take $z = t$ where $t \in \mathbb{R}$ is a parameter. The last row of the last matrix yields $5y - 7t = 5$ and thus $y = \frac{7}{5}t + 1$. We deduce from the first row that $x = 1 - \frac{1}{5}t$. All solutions are of the form

$$(x, y, z) = (1, 1, 0) + t\left(-\frac{1}{5}, \frac{7}{5}, 1\right).$$

2.3 Inverse matrices

Definition 2.5. *A matrix I of type (n,n) is called an* identity matrix *if $I = (a_{ij})_{i=1,j=1}^{nn}$, $a_{ii} = 1$ for all $i \in \{1,\ldots,n\}$ and $a_{ij} = 0$ whenever $i \neq j$. For example,*

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

for $n = 3$. It holds that $AI = IA = A$ for every matrix A of type (n,n).
Let A by a matrix of type (n,n). If there is a matrix B of type (n,n) such that

$$AB = BA = I$$

then B will be called an inverse matrix *to A and we use notation $B = A^{-1}$.*

The Gauss elimination may be used to find A^{-1}. In particular, one has to write down an augmented matrix $(A|I)$ and use elementary transformations to get (I, B). If this is possible, then $B = A^{-1}$.
Example Find A^{-1} to $A = \begin{pmatrix} 2 & -1 \\ 3 & -3 \end{pmatrix}$:

$$\begin{pmatrix} 2 & -1 & | & 1 & 0 \\ 3 & -3 & | & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & | & 1 & 0 \\ 1 & -2 & | & -1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & | & -1 & 1 \\ 2 & -1 & | & 1 & 0 \end{pmatrix}$$
$$\sim \begin{pmatrix} 1 & -2 & | & -1 & 1 \\ 0 & 3 & | & 3 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & | & -1 & 1 \\ 0 & 1 & | & 1 & -\frac{2}{3} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & | & 1 & -\frac{1}{3} \\ 0 & 1 & | & 1 & -\frac{2}{3} \end{pmatrix}$$

Consequently, $A^{-1} = \begin{pmatrix} 1 & -\frac{1}{3} \\ 1 & -\frac{2}{3} \end{pmatrix}$.

Definition 2.6. *A* square matrix *is a matrix of type (n,n) for some $n \in \mathbb{N}$.*
A square matrix A is called regular *if there is A^{-1}. Otherwise it is called* singular.

Observation 2.3. *Let A be a regular matrix. Then a system $Ax^T = b^T$ has a unique solution.*

Proof. Indeed, it suffices to apply A^{-1} from the left side on both sides of equation

$$Ax^T = b^T$$

to obtain

$$x^T = A^{-1}b^T.$$

□

Example The above proof describes another way how to solve a system of equations. Namely, we can first find A^{-1} and then $x^T = A^{-1}b^T$. Let solve the following two systems

$$\begin{aligned} 2x + y + z &= 3 \\ x + 3z &= -7 \\ 2x + y &= 1 \end{aligned}$$

and

$$\begin{aligned} 2x + y + z &= 0 \\ x + 3z &= 3 \\ 2x + y &= -1. \end{aligned}$$

Note that the matrix A of the systems (without the right hand side) is always the same. We compute A^{-1} as follows

$$\begin{pmatrix} 2 & 1 & 1 & | & 1 & 0 & 0 \\ 1 & 0 & 3 & | & 0 & 1 & 0 \\ 2 & 1 & 0 & | & 0 & 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 2 & 1 & 1 & | & 1 & 0 & 0 \\ 2 & 1 & 0 & | & 0 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & -5 & | & 1 & -2 & 0 \\ 0 & 1 & -6 & | & 0 & -2 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & -5 & | & 1 & -2 & 0 \\ 0 & 0 & -1 & | & -1 & 0 & 1 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 3 & | & 0 & 1 & 0 \\ 0 & 1 & -5 & | & 1 & -2 & 0 \\ 0 & 0 & 1 & | & 1 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & | & -3 & 1 & 3 \\ 0 & 1 & 0 & | & 6 & -2 & -5 \\ 0 & 0 & 1 & | & 1 & 0 & -1 \end{pmatrix}$$

Thus, the first system has a solution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3 & 1 & 3 \\ 6 & -2 & -5 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ -7 \\ 1 \end{pmatrix} = \begin{pmatrix} -13 \\ 27 \\ 2 \end{pmatrix}$$

and the second system has a solution

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3 & 1 & 3 \\ 6 & -2 & -5 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

2.4 Determinant

Definition 2.7. *Let A be a square matrix of type $(1,1)$ – i.e., $A = (a)$ for some $a \in \mathbb{R}$. The determinant of such matrix A is $\det A = a$.*
Let $A = (a_{i,j})$ be a square matrix of type (n,n). We denote by M_{ij} the determinant of a matrix $(n-1, n-1)$ which arises from A by leaving out the $i-$th row and $j-$th column. Choose $k \in \{1, \ldots, n\}$. Then

$$\det A = (-1)^{k+1} a_{k1} M_{k1} + (-1)^{k+2} a_{k2} M_{k2} + \ldots + (-1)^{k+n} a_{kn} M_{kn} = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} M_{kj}.$$

Examples:
Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Then $\det A = a_{11}a_{22} - a_{12}a_{21}$.
Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Then

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Observation 2.4.

- *Let B arise from A by multiplying one row by a real number α. Then $\det B = \alpha \det A$.*
- *Let B arise from A by switching two rows. Then $\det B = -\det A$.*
- *Let B arise from A by adding $\alpha-$multiple of one row to another one. Then $\det B = \det A$.*

Observation 2.5. *Let A be a square matrix having zeros under the main diagonal (i.e., $a_{ij} = 0$ for $i > j$). Then $\det A = a_{11}a_{22}a_{33}\dots a_{nn}$.*

Example Compute $\det A$ for

$$A = \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & 3 & -3 & 1 \\ 2 & -3 & 0 & 2 \\ 0 & 0 & 3 & -1 \end{pmatrix}.$$

According to the rules for transformations, we have

$$\det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & 3 & -3 & 1 \\ 2 & -3 & 0 & 2 \\ 0 & 0 & 3 & -1 \end{pmatrix} = \det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & 3 & -3 & 1 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & 3 & -1 \end{pmatrix}$$
$$= -\det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & -1 & 0 & 6 \\ 0 & 3 & -3 & 1 \\ 0 & 0 & 3 & -1 \end{pmatrix} = -\det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & -3 & 19 \\ 0 & 0 & 3 & -1 \end{pmatrix}$$
$$= -\det \begin{pmatrix} -1 & 1 & 0 & 2 \\ 0 & -1 & 0 & 6 \\ 0 & 0 & -3 & 19 \\ 0 & 0 & 0 & 18 \end{pmatrix} = 54.$$

Theorem 2.2. *Let A be $n \times n$ matrix. Statements following are equivalent:*

- $\det A = 0$.
- *$Ax^T = 0$ has a nontrivial solution.*
- *A is a singular matrix matrix.*
- *$rank A = n$.*
- *Rows of A are linearly dependent vectors.*
- *Columns of A are linearly dependent vectors.*

Theorem 2.3 (the Cramer rule)**.** *Consider a system $Ax^T = b^T$. Assume that A is a regular n by n matrix. Let $j \in \{1, \dots, n\}$ and denote by A_j a matrix arising from A by replacing $j-$th column by a vector b^T. Then*

$$x_j = \frac{\det A_j}{\det A}.$$

Example We use the Cramer rule to solve

$$\begin{aligned} 3x - 2y + 4z &= 3 \\ -2x + 5y + z &= 5 \\ x + y - 5z &= 0 \end{aligned}$$

We have $A = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 5 & 1 \\ 1 & 1 & -5 \end{pmatrix}$ and $\det A = -88$.

Further, $A_x = \begin{pmatrix} 3 & -2 & 4 \\ 5 & 5 & 1 \\ 0 & 1 & -5 \end{pmatrix}$ and $\det A_x = -108$. Consequently, $x = \frac{-108}{-88} = \frac{27}{22}$.

Next, $A_y = \begin{pmatrix} 3 & 3 & 4 \\ -2 & 5 & 1 \\ 1 & 0 & -5 \end{pmatrix}$ and $\det A_y = -122$. Consequently $y = \frac{-122}{-88} = \frac{61}{44}$.

Finally, $A_z = \begin{pmatrix} 3 & -2 & 3 \\ -2 & 5 & 5 \\ 1 & 1 & 0 \end{pmatrix}$ and $\det A_z = -46$. Consequently $z = \frac{-46}{-88} = \frac{23}{44}$.

2.5 Eigenvalues and eigenvectors

Definition 2.8. *Let A be a square matrix. We are looking for λ for which there is a nontrivial solution to*

$$Ax^T = \lambda x^T.$$

Such number λ is called eigenvalue.

This means that

$$(A - \lambda I)x^T = 0.$$

This equation has a nontrivial solution only if $A - \lambda I$ is a singular matrix. Consequently, λ is an eigenvalue if and only if

$$\det(A - \lambda I) = 0.$$

Definition 2.9. *Let λ be an eigenvalue of A. A vector v solving*

$$(A - \lambda I)v = 0$$

is called an eigenvector *corresponding to λ.*

Remark 2.3. *If v is an eigenvector then tv is also an eigenvector for all $t \in \mathbb{R}$.*
Let v and w be eigenvectors corresponding to the same eigenvalue. Then $tv + sw$ is also an eigenvector for all $t, s \in \mathbb{R}$.
Generally, let u_i, $i = \{1, \dots, k\}$ be eigenvectors corresponding to λ. Then all their linear combinations are also eigenvectors corresponding to λ.
In what follows, if we say that there is only one eigenvector v, we mean that there is just one-dimensional space of eigenvectors spanned by v. If we say that there are two eigenvectors v, w, we mean that there is two-dimensional space of eigenvectors spanned by v, w. And so on.

Example Find all eigenvalues and eigenvectors to $A = \begin{pmatrix} 5 & 1 \\ 4 & 5 \end{pmatrix}$.
First, we find eigenvalues by solving

$$0 = \det\left(\begin{pmatrix} 5 & 1 \\ 4 & 5 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} 5-\lambda & 1 \\ 4 & 5-\lambda \end{pmatrix}$$
$$= 25 - 10\lambda + \lambda^2 - 4 = \lambda^2 - 10\lambda + 21.$$

We obtain

$$\lambda_1 = 3,\ \lambda_2 = 7.$$

Consider first $\lambda_1 = 3$. Then we have to solve $\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0$. We have

$$\begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 \end{pmatrix}$$

and we take $y = t$ and $x = -\frac{t}{2}$. Thus $(x, y) = t(-1/2, 1)$ and $v_1 = (-1/2, 1)$ is an eigenvector related to $\lambda = 3$.
Consider $\lambda_2 = 7$. Then

$$\begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} \sim \begin{pmatrix} -2 & 1 \end{pmatrix}.$$

and we take $y = t$ and $x = \frac{t}{2}$. Consequently, $v_2 = (1/2, 1)$ is an eigenvector related to the eigenvalue $\lambda = 7$.

Exercise: Find eigenvalues and eigenvectors to $A = \begin{pmatrix} 10 & -9 \\ 4 & -2 \end{pmatrix}$.
First, we have to solve

$$0 = \det\begin{pmatrix} 10-\lambda & -9 \\ 4 & -2-\lambda \end{pmatrix} = \lambda^2 - 8\lambda + 16.$$

This yields the only solution $\lambda_1 = 4$. To find an eigenvector we solve

$$\begin{pmatrix} 6 & -9 \\ 4 & -6 \end{pmatrix} \sim \begin{pmatrix} 2 & -3 \end{pmatrix}.$$

Thus, $(3/2, 1)$ is an eigenvector.

Exercise: Find eigenvalues and eigenvectors to $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.
Solve

$$0 = \det\begin{pmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

We get $\lambda = 1$. To find eigenvalues we have to solve

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

The solutions are of the form $s(1, 0) + t(0, 1)$ for all real numbers $s, t \in \mathbb{R}$.

Summary Let λ be a single root of $\det(A - \lambda I)$. Then there is just one corresponding eigenvector. Let λ be a double root of $\det(A - \lambda I)$. Then there might be one corresponding eigenvector or two corresponding eigenvectors.

2.6 Some exercises

1. Write $X = A^T A - 2A$ for $A = \begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix}$.

2. Find a matrix X such that $XA = B$ where $A = \begin{pmatrix} 2 & 1 \\ -4 & -3 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 2 \\ 6 & 4 \end{pmatrix}$.

3. Prove Observation 2.2 for a matrix 3×3.

4. Prove Observation 2.5 for a matrix 3×3 and 4×4.

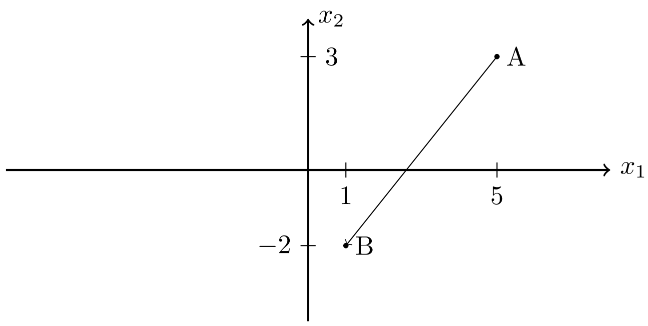
5. Find all eigenvectors and eigenvalues for $A = \begin{pmatrix} -2 & -8 \\ 1 & 2 \end{pmatrix}$. This time you should consider complex numbers.

3 Euclidean space

3.1 Euclidean space, Distance

Euclidean space

We consider a space $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n), x_1, \dots, x_n \in \mathbb{R}\}$ – $n-$tuples can be understood as points or as vectors. Points cannot be added together, however, $A + v$ is a point for arbitrary point A and arbitrary vector v. Simultaneously, $B - A$ is a vector for every two points A and B. We denote this vector by $\vec{AB}$. Here $A = (5, 3)$, $B = (1, -2)$ and $\vec{AB} = B - A = (-4, -5)$. It



holds that $B = A + \vec{AB}$.

Distance

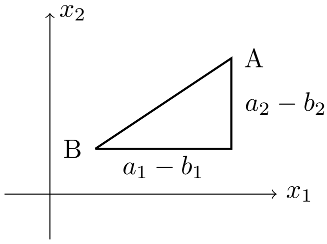
Let $A = (a_1, a_2)$ and $B = (b_1, b_2)$ be points in the two dimensional space. The Pythagorean theorem yields that

$$\varrho(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

This can be generalized to n dimensions as follows:

Definition 3.1. *Let A and B be points in $\mathbb{R}^n$. Then we define the* distance *between A and B as*

$$\varrho(A, B) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}.$$



Example: The distance between $A = (1, -3, 5)$ and $B = (2, 4, 0)$ is $\varrho(A, B) = \sqrt{(1-2)^2 + (-3-4)^2 + (5-0)^2} = \sqrt{75}$.

Observation 3.1. *Let A, B, C be points in $\mathbb{R}^n$. Then*

- $\varrho(A, B) \geq 0$,
- $\varrho(A, B) = 0 \;\Leftrightarrow\; A = B$,
- $\varrho(A, B) = \varrho(B, A)$,
- $\varrho(A, C) \leq \varrho(A, B) + \varrho(B, C)$.

3.2 Scalar product, norm and angles

Definition 3.2. *: Let $u = (u_1, u_2, \dots, u_n)$ and $v = (v_1, v_2, \dots, v_n)$. Then we define a* scalar product *of u and v as*

$$u \cdot v = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

Observation 3.2. *Let $u, v, w \in \mathbb{R}^n$ be arbitrary vectors and $\alpha \in \mathbb{R}$. Then:*

- $u \cdot v = v \cdot u$,
- $u \cdot (v + w) = u \cdot v + u \cdot w$,
- $(\alpha u) \cdot v = \alpha(u \cdot v)$,
- $u \cdot u \geq 0$ *and* $u \cdot u = 0 \;\Leftrightarrow\; u = 0$.

Definition 3.3. *A* norm *of a vector $v \in \mathbb{R}^n$ is a number*

$$\|v\| = \sqrt{v \cdot v}$$

Observation 3.3. *Let $u, v \in \mathbb{R}^n$ be arbitrary vectors and $\alpha \in \mathbb{R}$. Then:*

- $\|v\| \geq 0$,
- $\|v\| = 0 \;\Leftrightarrow\; v = 0$,
- $\|\alpha v\| = |\alpha| \|v\|$,
- $\|u + v\| \leq \|u\| + \|v\|$,
- $|u \cdot v| \leq \|u\| \|v\|$.

Proof. The first, second and third properties follows directly from the definition.
To prove the fifth property we consider a vector $u+tv$, $t\in\mathbb{R}$. It holds

$$0\le (u+tv)\cdot(u+tv)=\|u\|^2+2t(u\cdot v)+t^2\|v\|^2.$$

The right hand side is a second order polynomial of variable t. Since it is always positive, its discriminant should be negative. Thus

$$4(u\cdot v)^2-4\|u\|^2\|v\|^2<0.$$

This directly yields the claim.
To prove the fourth property it is enough to write

$$\|u+v\|^2=(u+v)\cdot(u+v)=\|u\|^2+2u\cdot v+\|v\|^2$$
$$\le \|u\|^2+2\|u\|\|v\|+\|v\|^2=(\|u\|+\|v\|)^2.$$

□

Recall also that

$$\varrho(A,B)=\|\vec{BA}\|.$$

This follows directly from the definitions.

Definition 3.4. *An* angle *between two nonzero vectors $u,v\in\mathbb{R}^n$ is a number $\varphi\in[0,\pi]$ fulfilling*

$$\cos\varphi=\frac{u\cdot v}{\|u\|\|v\|}.$$

Example Consider a triangle ABC where $A=(2,0,4)$, $B=(1,5,3)$ and $C=(-1,2,3)$. Lets determine the length of the sides of the triangle and lets determine all internal angles. We have

$$\varrho(A,B)=\sqrt{1^2+5^2+1^2}=\sqrt{27},\ \varrho(A,C)=\sqrt{14},\ \varrho(B,C)=\sqrt{13}.$$

Let $\alpha=\sphericalangle BAC$. We have

$$\cos\alpha=\frac{\vec{AB}\cdot\vec{AC}}{\|\vec{AB}\|\|\vec{AC}\|}=\frac{14}{\sqrt{27}\sqrt{14}}=\sqrt{\frac{14}{27}}$$

and we use a calculator to deduce $\alpha=0.7669$. Next, $\gamma=\sphericalangle ACB$ fulfills

$$\cos\gamma=\frac{\vec{AC}\cdot\vec{BC}}{\|\vec{AC}\|\|\vec{BC}\|}=\frac{0}{\|\vec{AC}\|\|\vec{BC}\|}=0$$

and $\gamma=\frac{\pi}{2}$.

Remark 3.1. *We would like to recall that vectors $u,v\in\mathbb{R}^n$ are perpendicular iff $u\cdot v=0$.*

In what follows we will use symbols $i=(1,0,0)$, $j=(0,1,0)$ and $k=(0,0,1)$.

Definition 3.5. *Let $u=(u_1,u_2,u_3)$ and $v=(v_1,v_2,v_3)$ be vectors in $\mathbb{R}^3$. We define their* cross product *(or* vector product*) as*

$$u\times v=\det\begin{pmatrix} i & j & k\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{pmatrix}.$$

Thus

$$u\times v=(u_2v_3-u_3v_2,u_3v_1-u_1v_3,u_1v_2-u_2v_1).$$

Example As an example, let compute

$$(-3,2,1)\times(1,0,5)=(10-0,1-(-15),0-2)=(10,16,-2).$$

Observation 3.4. *Let $u,v,w\in\mathbb{R}^3$ be arbitrary vectors and $\alpha\in\mathbb{R}$. Then*

- $u\times v=-v\times u$,
- $(\alpha u)\times v=\alpha(u\times v)$,
- $u\times(v+w)=u\times v+u\times w$,
- *let $z=u\times v$. Then $z\cdot u=0$ and $z\cdot v=0$,*
- *assume that u is a nonzero vector, then $u\times v=0$ if and only if $v=tu$ for some $t\in\mathbb{R}$.*

Warning It is not true that $u\times(v\times w)=(u\times v)\times w$ for all $u,v,w\in\mathbb{R}^3$.

Definition 3.6. *Let $u,v,w\in\mathbb{R}^3$ be vectors. Their* triple product *(also a* box product *or* mixed product*) is a number*

$$u\cdot(v\times w).$$

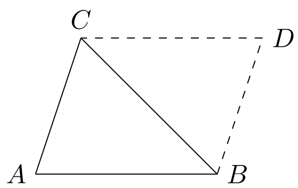
It holds that

$$u\cdot(v\times w)=v\cdot(w\times u)=w\cdot(u\times v)=\det\begin{pmatrix}u\\v\\w\end{pmatrix}.$$

Further properties of the cross product and triple product

Observation 3.5. *Let $u,v\in\mathbb{R}^3$. Then $\|u\times v\|=\sin\phi\|u\|\|v\|$.*

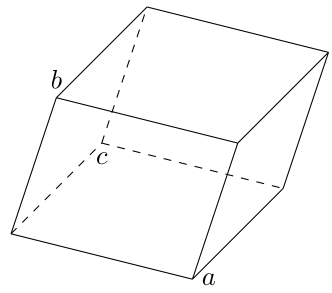
Example Compute the surface of a triangle ABC where $A=(2,0,-1)$, $B=(3,2,1)$ and $C=(-2,5,2)$. Denote $u=\vec{AB}=(1,2,2)$ and $v=\vec{AC}=(-4,5,3)$. The surface of the



parallelogram $ABCD$ is $\|u\times v\|$. Thus, the surface of the triangle is thus

$$\frac{1}{2}\|u\times v\|=\frac{1}{2}\|(1,2,2)\times(-4,5,3)\|=\frac{1}{2}\|(-4,-11,13)\|=\frac{1}{2}\sqrt{306}\approx 8.7.$$

Observation 3.6. *The volume of a parallelepiped generated by three vectors $a,b,c\in\mathbb{R}^3$ is equal to $|a\cdot(b\times c)|$.*



3.3 Lines

Let $A, B \in \mathbb{R}^n$. The line (determined uniquely) passing through A and B consists of points

$$X = A + t\vec{AB}, \ t \in \mathbb{R}.$$

Recall that $\vec{AB} = B - A$, thus for $t = 0$ we get $X = A$, for $t = 1$ we get $X = B$. Thus, $X = A + t\vec{AB}$, $t \in [0, 1]$ is an equation of the line segment AB. For $t = 1/2$ we get the midpoint S of the line segment AB and thus we can write

$$S = A + \frac{1}{2}(B - A) = A + \frac{1}{2}B - \frac{1}{2}A = \frac{1}{2}(A + B).$$

The vector $\vec{AB}$ is called a *direction vector*. The parametric equations are not uniquely given. A point C belongs to a line $X = A + tv, \ t \in \mathbb{R}$ if there is $t' \in \mathbb{R}$ such that $C = X + t'v$.

The mutual position of two lines

Two lines $p : X = A + tv$ and $q : X = B + tu$ are

- identical, if $u \| v$ and $A \in q$,
- parallel, if $u \| v$ and $A \notin q$,
- intersecting lines, if u is not parallel to v and there is a point $P \in p \cap q$. The point P is called an intersection,
- skew lines, if u is not parallel to v and there is no intersection.

Example Find a mutual position of AB and CD for $A = (2, -5, -2)$, $B = (0, -3, 0)$, $C = (4, 1, 2)$ and $D = (-1, -2, 1)$.
We have

$$p : X = (2, -5, -2) + t(-2, 2, 2), \qquad q : X = (4, 1, 2) + t(-5, -3, -1).$$

Clearly, $(-2, 2, 2)$ and $(-5, -3, -3)$ are not parallel to each other. Is there an intersection? If yes, then there are $s, t \in \mathbb{R}$ such that

$$P = (2, -5, -2) + t(-2, 2, 2) = (4, 1, 2) + s(-5, -3, -1)$$

and we deduce

$$\begin{aligned} 2 - 2t &= 4 - 5s \\ -5 + 2t &= 1 - 3s \\ -2 + 2t &= 2 - s. \end{aligned} \tag{2}$$

We reformulate this as

$$\begin{aligned} 5s - 2t &= 2 \\ 3s + 2t &= 6 \\ s + 2t &= 4 \end{aligned}$$

$$\begin{pmatrix} 5 & -2 & | & 2 \\ 3 & 2 & | & 6 \\ 1 & 2 & | & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & | & 4 \\ 5 & -2 & | & 2 \\ 3 & 2 & | & 6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & | & 4 \\ 0 & -12 & | & -18 \\ 0 & -4 & | & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & | & 4 \\ 0 & 2 & | & 3 \end{pmatrix}.$$

We get $t = 3/2$ and $s = 1$. These two lines are intersection lines and the point of intersection is $(-1, -2, 1)$.

Distance between a point and a line

The distance between A and a line $p : X = B + t\vec{BC}$ is

$$\inf_{X \in p} \varrho(A, X).$$

Problem Find a distance between $A = (2, 3)$ and $p : X = (1, 0) + t(-1, 1)$.
We have to find a minimum of $\varrho(A, X(t)) = \|A - X(t)\| = \|(1, 3) - t(-1, 1)\| = \sqrt{(1 + t, 3 - t)}$. That is $\sqrt{(1 + t)^2 + (3 - t)^2}$. It is enough to find a minimum of $f(t) = (1 + t)^2 + (3 - t)^2 = 2t^2 - 4t + 10$. We have $f'(t) = 4t - 4$ and the minimum is at the point $t = 1$. We have $f(1) = 8$ and the distance is $\sqrt{8}$.

Distance between two lines

The distance between two identical lines is zero. The distance between intersecting lines is also zero. The distance between two parallel lines $p : X = A + tv$ and $q : X = B + tu$ is the distance between A and q. The distance between two skew lines will be defined later.

Lines in two dimensions

Consider a line

$$p : \begin{aligned} x &= a + ut \\ y &= b + vt \end{aligned}$$

passing through $A = (a, b)$ and with a direction vector (u, v). Assume $u \neq 0$. Then we may deduce from the first equation that $t = \frac{1}{u}(x - a)$. We plug this into the second equation to get $y = b + \frac{v}{u}(x - a)$ which is equivalent to

$$-\frac{v}{u}x + y - b + \frac{v}{u}a = 0.$$

Note that $\left(-\frac{v}{u}, 1\right)$ is perpendicular to (u, v).
The latter equation is called a normal form of a line. .

Example Find a normal form of a line

$$p : \begin{aligned} x &= 1 + 2t \\ y &= 1 - t. \end{aligned}$$

The direction vector is $(2, -1)$ and the normal vector is $(1, 2)$ (one of many). Consequently, the normal form is

$$x + 2y + c = 0 \tag{3}$$

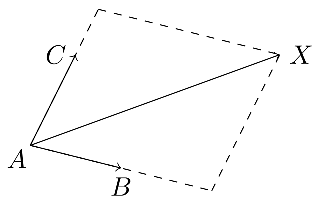
for certain value of $c \in \mathbb{R}$. We know that $(1, 1) \in p$ and thus (3) should be satisfied for $x = 1$, $y = 1$. We obtain $c = -3$ and the resulting normal form is

$$x + 2y - 3 = 0.$$

3.4 Planes in 3 dimensions

Let $A, B, C \in \mathbb{R}^3$ be points which are not collinear (there is no line passing through all of them). Then $u = \vec{AB}$ and $v = \vec{AC}$ are not parallel and the (uniquely determined) plane ABC consists of points in a form

$$X = A + su + tv, \;\; s, t \in \mathbb{R}.$$



Mutual position of a line and a plane

Let have a plane $X = A + tu + sv$ and a line $X = B + rw$. Their intersection consists of all points for which there are parameters $r, t, s \in \mathbb{R}$ such that

$$A + tu + sv = B + rw.$$

There might be infinitely many points in the intersection – in that case the line belongs to the plane – there might be just one point in the intersection – in that case the line intersects the plane – or there might be no intersection at all – in that case the line is parallel to the plane.

Example: Find the mutual position of a line $p : DE$ and a plane $\sigma : ABC$ where

$$A = (1, 2, -3), \; B = (3, 0, 2), \; C = (0, 2, 3), \; D = (1, 1, 2), \; E = (3, 2, 0).$$

We deduce

$$\begin{aligned} &p : X = (1, 1, 2) + r(2, 1, -2), \\ &\sigma : X = (1, 2, -3) + t(2, -2, 5) + s(-1, 0, 6). \end{aligned}$$

Thus

$$\begin{aligned} 1 + 2r &= 1 + 2t - s \\ 1 + r &= 2 - 2t \\ 2 - 2r &= -3 + 5t + 6s. \end{aligned}$$

and this might be reformulated as

$$\begin{aligned} 2r + s - 2t &= 0 \\ r + 2t &= 1 \\ -2r - 6s - 5t &= -5 \end{aligned}$$

We solve this by the Gauss elimination

$$\begin{pmatrix} 2 & 1 & -2 & | & 0 \\ 1 & 0 & 2 & | & 1 \\ -2 & -6 & -5 & | & -5 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 & | & 1 \\ 2 & 1 & -2 & | & 0 \\ -2 & -6 & -5 & | & -5 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 2 & | & 1 \\ 0 & 1 & -6 & | & -2 \\ 0 & -6 & -1 & | & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 2 & | & 1 \\ 0 & 1 & -6 & | & -2 \\ 0 & 0 & -37 & | & -15 \end{pmatrix}$$

and we get $t = 15/37$, $s = 16/37$, $r = 7/37$. Therefore, there is an intersection consisting of a point

$$P = \left(\frac{51}{37}, \frac{44}{37}, \frac{60}{37}\right).$$

Normal form of a plane

Normal form of a plane is an equation

$$ax + by + cz + d = 0.$$

Here (a, b, c) is a vector perpendicular to the direction vectors u and v. Let σ be given by

$$\sigma : X = (-7, 0, 0) + t(-2, -3, 0) + s(7, 12, 3).$$

A vector perpendicular to $u = (-2, -3, 0)$, $v = (7, 12, 3)$ can be obtained (for example) as $n := u \times v$. We get

$$n = (-9, 6, -3)$$

and the equation is of the form

$$-9x + 6y - 3z + d = 0$$

and since $(-7, 0, 0)$ is in the plane, we obtain $d = -63$. We can further divide the whole equation by 3 to get

$$-3x + 2y - z - 21 = 0.$$

Example: Determine the distance between the plane $\sigma : 3x - 2y + z + 21 = 0$ and $A = (3, 0, -2)$.

Clearly, the point P in the plane which is nearest to A should satisfy that $\vec{PA}$ is perpendicular to both direction vectors (and thus it is parallel to the normal vector). In particular, the direction vector of the line PA is $(3, -2, 1)$ and thus it can be written as

$$p : X = (3, 0, -2) + t(3, -2, 1).$$

Let find an intersection of p and σ. Since such point should be in p, we get $x = 3 + 3t$, $y = -2t$, $z = -2 + t$ and we plug these relations into the normal form. We obtain

$$9 + 9t + 4t - 2 + t + 21 = 0$$

which yields $t = -2$ and the intersection is $P = (-3, 4, -4)$. Thus, the distance between A and σ is

$$\varrho(A, P) = \sqrt{6^2 + 4^2 + 2^2} = \sqrt{56}.$$

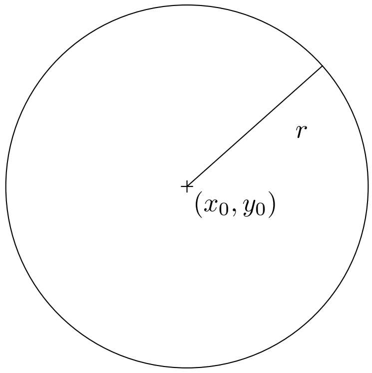
Based on the same considerations as above we may deduce the following.

Observation 3.7. *The distance between $X = (x_0, y_0, z_0)$ and $\sigma : ax + by + cz + d = 0$ is equal to*

$$\frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

The distance between two skew lines

Let $p : X = A + tu$ and $q : X = B + tv$ be two skew lines. Then the distance between them is equal to the distance between line p and a plane $\sigma : X = B + rv + su$. (Note that p and σ has no intersection and thus the distance between p and σ is equal to the distance between A and σ.)



3.5 Few words about topology

Definition 3.7. *An open ball centered at* $(x_0, y_0) \in \mathbb{R}^2$ *with radius* $r \in (0, \infty)$ *is a set*

$$B_r(x_0, y_0) = \{(x, y) \in \mathbb{R}^2, \|(x, y) - (x_0, y_0)\| < r\}.$$

Definition 3.8. *A set* $M \subset \mathbb{R}^2$ *is* open *if for every* $(x_0, y_0) \in \mathbb{M}$ *there is* $r > 0$ *such that* $B_r(x_0, y_0) \subset M$.
A set M *is called* closed *if* $\mathbb{R}^2 \setminus M$ *is open.*

Example A set $M := (0, 1) \times (0, 1)$ is open. Indeed, let $(a, b) \in M$. Define $\delta = \min\{a, b, 1 - a, 1 - b\}$. Since $a \in (0, 1)$ and $b \in (0, 1)$ we have $\delta > 0$. Necessarily, $B_{\delta/2}(a, b) \subset M$. On the other hand, a set $M := [0, 1] \times (0, 1)$ is not open. Consider for example a point $(1, 1/2) \in M$. Then every ball $B_r(1, 1/2)$ contains a point $(1 + r/2, 1/2)$ which is outside of M. Note that M is not closed. Why?

Remark 3.2.
- *$\emptyset$ and $\mathbb{R}^2$ are open sets (and closed sets as well),*
- *a union of open sets is an open set,*
- *an intersection of two open sets is an open set,*
- *a union of two closed sets is a closed set,*
- *an intersection of closed sets is a closed set.*

Observation 3.8. *Let* $f : \mathbb{R}^2 \mapsto \mathbb{R}$ *be a continuous function. Then* $f^{-1}(A)$ *is an open set for every* $A \subset \mathbb{R}$ *open. Similarly,* $f^{-1}(B)$ *is a closed set for every* $B \subset \mathbb{R}$ *closed.*

Question What is a *continuous function*? We will see later.
For now: A projection $p : \mathbb{R}^2 \mapsto \mathbb{R}$, $p(x, y) = x$ is a continuous function (as well as projection $q(x, y) = y$). A sum, difference and product of two continuous functions are continuous functions. A quotient of two continuous function is again a continuous function. A composition of two continuous function is a continuous function.

Example Let consider a set

$$M := \{(x, y), x \in (-1, 1),\ y < x^2\}.$$

Is this set open? First, $f(x,y) = |x|$ is a continuous function. Indeed, $f(x,y) = |p(x,y)|$ is a composition of p and $|\cdot|$. Thus, $f^{-1}((-\infty,1)) = \{(x,y) \in \mathbb{R}^2, x \in (-1,1)\}$ is an open set.
Next, $g(x,y) = y - x^2$ is a continuous function. Indeed, $g(x,y) = q(x,y) - p(x,y)^2$. Consequently, $f^{-1}((-\infty,0)) = \{(x,y) \in \mathbb{R}^2, y - x^2 < 0\} = \{(x,y) \in \mathbb{R}^2, y < x^2\}$.
Since $M = f^{-1}((-\infty,1)) \cap g^{-1}((-\infty,0))$, we deduce that M is open.

Definition 3.9. *An* interior *of set $M \subset \mathbb{R}^2$ is a set M^0 of all points (x_0,y_0) for which there is $r > 0$ such that $B_r(x_0,y_0) \subset M$. Equivalently, it is the biggest open set contained in M.*
A closure *of a set $M \subset \mathbb{R}^2$ is a set $\overline{M}$ defined as $\overline{M} := \mathbb{R}^2 \setminus (\mathbb{R}^2 \setminus M)^0$. Equivalently, it is the smallest closed set containing M.*
A boundary *of a set M is denoted by ∂M and it is defined as $\overline{M} \setminus M^0$.*

Example Consider $M = [0,1] \times (0,1)$. Then $M^0 = (0,1) \times (0,1)$ and $\overline{M} = [0,1] \times [0,1]$. We deduce that

$$\partial M = \overline{M} \setminus M^0 = ([0,1] \times \{0,1\},\ \{0,1\} \times [0,1]) .$$

Definition 3.10. *Let $M \subset \mathbb{R}^2$. A point $(x_0,y_0) \in \mathbb{R}^2$ is a* limit point *of M if $B_r(x_0,y_0) \cap M \neq \emptyset$ for every $r > 0$.*
A point $(x_0,y_0) \in M$ is an isolated point *of M if there is $r > 0$ such that $B_r(x_0,y_0) \cap M = \{(x_0,y_0)\}$.*

Example Consider a set $M := \{(x,y) \in \mathbb{R}, y = 0, x = 1/n,\ n \in \mathbb{N}\}$. We claim, that $(0,0)$ is a limit point of M. Indeed, let $r > 0$. Then there is n_r such that $n_r > 1/r$ and, clearly, $(1/n_r, 0) \in M$ is such point that $\|(1/n_r,0) - (0,0)\| < r$ and thus $B_r(0,0) \cap M = (1/n_r, 0)$.

3.6 Some exercises

1. Find a distance between lines AB and CD where $A = (1,0,1)$, $B = (2,2,1)$, $C = (2,1,4)$ and $D = (-1,0,1)$.

2. Prove, that all points of $M := \{(x,y) \in \mathbb{R}, y = 0, x = 1/n,\ n \in \mathbb{N}\}$ are isolated points.

4 Functions of two variables

4.1 Introduction

Definition 4.1. *Let $M \subset \mathbb{R}^2$ be a nonempty set. A* real function of two variables *defined on a set M is a formula f which assigns a (unique) real number y to every pair $(x_1,x_2) \in M$. We use the notation*

$$y = f(x_1,x_2).$$

To denote the function itself we use a notation $f : M \mapsto \mathbb{R}$. The set M is called a domain *of f and we write $M = \text{Dom } f$.*

Usually, the function will be given only by its formula without any specific domain. In that case, we assume that the domain is a maximal set for which has the formula sense. For example, a function

$$f(x_1,x_2) = \log(x_1 + x_2)$$

is defined on a set

$$\text{Dom } f = \{(x_1,x_2) \in \mathbb{R}^2,\ x_1 + x_2 > 0\}.$$

Problem Find (and sketch) a maximal set $M \subset \mathbb{R}^2$ of such pairs (x_1, x_2) for which the function

$$f(x_1, x_2) = \frac{1}{\sqrt{x_1^2 + x_2 - 1}}.$$

Necessarily, $\sqrt{x_1^2 + x_2 - 1} > 0$ and we deduce that the function has sense for all pairs satisfying

$$x_1^2 + x_2 - 1 > 0.$$

Definition 4.2. *Let $z = f(x, y)$ be a function of two variables. The* graph *of f is a set*

$$\mathit{graph} f = \{(x, y, f(x, y) \in \mathbb{R}^3,\ (x, y) \in \text{Dom } f\}.$$

Definition 4.3. *A* contour line *C at height $z_0 \in \mathbb{R}$ is a set*

$$\{(x, y) \in \mathbb{R}^2,\ f(x, y) = z_0\}.$$

Example Find contour lines at heights $z_0 = -2, -1, 0, 1, 2$ for a function

$$f(x, y) = \frac{x^2 + y^2}{2x}.$$

Algebra of functions of two variables:
Sum, product and division is defined 'pointwisely'. Consider, for example, functions $f(x, y) = e^{xy}$ and $g(x, y) = \sqrt{1 - x^2 - y^2}$. Then

- $(f + g)(x, y) = e^{xy} + \sqrt{1 - x^2 - y^2}$,
- $(fg)(x, y) = e^{xy}\sqrt{1 - x^2 - y^2}$,
- $\frac{f}{g}(x, y) = \frac{e^{xy}}{\sqrt{1-x^2-y^2}}$. Beware, here we have to exclude from the domain all points where g equals zero.

Composition of functions: Let $f : M \mapsto \mathbb{R}^2$ (this means that there are two functions $f_1 : M \mapsto \mathbb{R}$ and $f_2 : M \mapsto \mathbb{R}$ and $f = (f_1, f_2)$) and $g : \mathbb{R}^2 \mapsto \mathbb{R}$. Then a *composition* is a function $h = g \circ f$ defined as

$$h(x, y) = g(f_1(x, y), f_2(x, y)).$$

Similarly, if $f : M \mapsto \mathbb{R}$ and $g : \mathbb{R} \mapsto \mathbb{R}$ then $h = g \circ f$ is defined as $h(x, y) = g(f(x, y))$

We can also introduce the boundedness of a function $f : M \subset \mathbb{R}^2 \mapsto \mathbb{R}$. This can be done similarly to the one dimensional case. The precise definition of a bounded function is left as an exercise.

4.2 Continuity

Definition We say that $f : M \mapsto \mathbb{R}$ is continuous at a point $(x_0, y_0) \in M$ if

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall (x, y) \in (M \cap B_\delta(x_0, y_0)),\ |f(x, y) - f(x_0, y_0)| < \varepsilon.$$

Let $N \subset M$ and let $f : M \mapsto \mathbb{R}$ be continuous at all points $(x_0, y_0) \in N$. Then we say that f is continuous on N. If f is continuous on Dom f then we simply say that f is continuous.

Observation 4.1. *Let f_1 and f_2 be continuous functions. Then*

$$f_1 + f_2,\ f_1 - f_2 \text{ and } f_1 f_2$$

are continuous function. Moreover, $\frac{f_1}{f_2}$ is a continuous function on a set $\{(x,y) \in \mathbb{R}^2, f_2(x,y) \neq 0\}$. Further, $f_1 \circ f_2$ is also a continuous function. We remind that $f(x,y) = x$ and $f(x,y) = y$ are continuous function.

Example A function

$$f(x,y) = \frac{x + \sqrt{x+y}}{1 + \cos^2 x}$$

is continuous for all $(x,y) \in \mathbb{R}^2$.

4.3 Limits

Definition 4.4. *Let (x_0, y_0) be a limit point of $M \subset \mathbb{R}^2$ and let $f : M \mapsto \mathbb{R}$. We say that a limit of f at the point (x_0, y_0) is $A \in \mathbb{R}$ if*

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall (x,y) \in (M \cap B_\delta(x_0, y_0)),\ |f(x,y) - A| < \varepsilon.$$

We write $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = A$.
We say that a limit of f at the point (x_0, y_0) is ∞ if

$$\forall M > 0,\ \exists \delta > 0,\ \forall (x,y) \in (M \cap B_\delta(x_0, y_0)),\ f(x,y) > M.$$

We write $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = \infty$.
We say that a limit of f at the point (x_0, y_0) is $-\infty$ if $\lim_{(x,y)\to(x_0,y_0)} -f(x,y) = -\infty$.

Observation 4.2 (Arithmetic of limits)**.** *Let f and g be two functions and let (x_0, y_0) be a limit point of* Dom f *and of* Dom g. *Then*

$$\begin{aligned}
\lim_{(x,y)\to(x_0,y_0)} (f+g)(x,y) &= \lim_{(x,y)\to(x_0,y_0)} f(x,y) + \lim_{(x,y)\to(x_0,y_0)} g(x,y) \\
\lim_{(x,y)\to(x_0,y_0)} fg(x,y) &= \lim_{(x,y)\to(x_0,y_0)} f(x,y) \lim_{(x,y)\to(x_0,y_0)} g(x,y) \\
\lim_{(x,y)\to(x_0,y_0)} \frac{f}{g}(x,y) &= \frac{\lim_{(x,y)\to(x_0,y_0)} f(x,y)}{\lim_{(x,y)\to(x_0,y_0)} g(x,y)}.
\end{aligned}$$

assuming the right hand side is well defined.

The numbers $\infty - \infty$, $0 \cdot \infty$, $\frac{0}{0}$, $\frac{\infty}{\infty}$ are not well defined (similarly to the one dimensional case).

Observation 4.3. *A function f is continuous at point $(x_0, y_0) \in$* Dom f *if and only if* $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0, y_0)$.

Example Consider a function

$$f(x,y) = \frac{x^2 y^2}{x^2 y^2 + (x-y)^2}.$$

This function is not defined at $(0,0)$. It is possible to define the value $f(0,0)$ in such a way that f is continuous? In particular, does there exists a finite limit

$$\lim_{(x,y)\to(0,0)} f(x,y)?$$

First, we approach $(0,0)$ along the line $y=0$. We have

$$\lim_{(x,0)\to(0,0)} f(x,0) = \lim_{x\to 0} \frac{0}{x^2} = 0.$$

Next, we approach $(0,0)$ along the line $x=y$. We have

$$\lim_{(x,x)\to(0,0)} f(x,x) = \lim_{x\to 0} \frac{x^4}{x^4} = 1.$$

As a result, $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Lemma 4.1 (Sandwich lemma)**.** *Let f,g,h be three functions defined on $B_\delta(x_0,y_0)\setminus\{(x_0,y_0)\}$ for some $\delta>0$. Assume*

$$\forall (x,y)\in B_\delta(x_0,y_0)\setminus\{(x_0,y_0)\},\ \ g(x,y)\le f(x,y)\le h(x,y).$$

If $\lim_{(x,y)\to(x_0,y_0)} g(x,y) = \lim_{(x,y)\to(x_0,y_0)} h(x,y) = A\in\mathbb{R}$ then also

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = A.$$

Corollary 4.1. $\lim_{(x,y)\to(x_0,y_0)} |f(x,y)| = 0 \Rightarrow \lim_{(x,y)\to(x_0,y_0)} f(x,y) = 0.$

Example Compute

$$\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{x^2+y^2}}.$$

We use notation $f(x,y)=\frac{xy}{\sqrt{x^2+y^2}}$. First of all, we have $\lim_{x\to0} f(x,0)=0$ and $\lim_{y\to0} f(0,y)=0$. Thus, if there is a limit, it is equal to 0. We use the well known AM-GM inequality $(2|xy|\le(x^2+y^2))$ to deduce

$$0\le \frac{|xy|}{\sqrt{x^2+y^2}} \le \frac{x^2+y^2}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2}\to 0$$

as $(x,y)\to 0$. The sandwich lemma yields $\lim_{(x,y)\to(0,0)} |f(x,y)|=0$ and we have just proven that the given limit is equal to 0.

4.4 Derivatives

Definition 4.5. *We define* partial derivatives *with respect to x and with respect to y as*

$$\frac{\partial f}{\partial x}(x_0,y_0) = \lim_{h\to0}\frac{f(x_0+h,y_0)-f(x_0,y_0)}{h}$$
$$\frac{\partial f}{\partial y}(x_0,y_0) = \lim_{h\to0}\frac{f(x_0,y_0+h)-f(x_0,y_0)}{h}.$$

Example Let compute $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for a function

$$f(x,y)=3x^2y+x^2+\log(x^2+y^2).$$

Let first compute $\frac{\partial f}{\partial x}$. In that case we treat y as a constant and we deduce that

$$\frac{\partial f}{\partial x} = 6xy+2x+\frac{2x}{x^2+y^2}.$$

In order to compute $\frac{\partial f}{\partial y}$ we treat x as a constant and we get

$$\frac{\partial f}{\partial y} = 3x^2+\frac{2y}{x^2+y^2}.$$

Definition 4.6. *We define second order partial derivatives as follows*

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

Analogously we define the third and higher order partial derivatives.

Example Let compute the first and second order derivatives for $f(x,y) = \frac{x}{y} - e^{xy}$. We have

$$\frac{\partial f}{\partial x} = \frac{1}{y} - ye^{xy}, \quad \frac{\partial f}{\partial y} = -\frac{x}{y^2} - xe^{xy}$$
$$\frac{\partial^2 f}{\partial x^2} = -y^2 e^{xy}, \quad \frac{\partial^2 f}{\partial y \partial x} = -\frac{1}{y^2} - e^{xy} - xye^{xy}$$
$$\frac{\partial^2 f}{\partial y^2} = 2\frac{x}{y^3} - x^2 e^{xy}, \quad \frac{\partial^2 f}{\partial x \partial y} = -\frac{1}{y^2} - e^{xy} - xye^{xy}.$$

Observation 4.4. *Let the second order derivative of a function f be continuous. Then*

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Definition 4.7. *The* derivative *of a function f at point (x_0, y_0) in direction the of a vector $v \in \mathbb{R}^2$ is a number*

$$Df((x_0,y_0),v) = \lim_{h \to 0} \frac{f((x_0,y_0)+hv) - f(x_0,y_0)}{h}.$$

Example The derivative $Df((1,2),(3,-4))$ of a function $f(x,y) = \arctan(xy)$ is

$$Df((1,2),(3,-4)) = \lim_{h \to 0} \frac{\arctan((1+3h)(2-4h)) - \arctan 2}{h}$$
$$= \lim_{h \to 0} \frac{\frac{1}{1+(1+3h)^2(2-4h)^2}(2-24h)}{1} = \frac{2}{5}$$

Remark 4.1. *It holds that*

$$\frac{\partial f}{\partial x}(x,y) = Df((x,y),(1,0)), \quad \frac{\partial f}{\partial y}(x,y) = Df((x,y),(0,1)).$$

Definition 4.8. *Let $(x_0,y_0) \in \mathrm{Dom}\, f$. A vector*

$$\nabla f(x_0,y_0) = \left(\frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\right)$$

is called the gradient *of f at point (x_0,y_0).*

Observation 4.5. *It holds that*

$$Df((x_0,y_0),v) = \nabla f(x_0,y_0) \cdot v = \|\nabla f(x_0,y_0)\| \|v\| \cos\alpha,$$

where α is an angle between ∇f and v.

Corollary 4.2. *As a consequence, the function increases at most in the direction of the gradient.*

Theorem 4.1 (Chain rule – derivative of a composed function)**.** *Let $n = 1$ or 2 and let $f : \mathbb{R}^n \mapsto \mathbb{R}^2$ and $g : \mathbb{R}^2 \mapsto \mathbb{R}$. Then*

$$\frac{\partial (g \circ f)}{\partial x_i} = \frac{\partial g}{\partial y_1}\frac{\partial f_1}{\partial x_i} + \frac{\partial g}{\partial y_2}\frac{\partial f_2}{\partial x_i}, \; i = \{1, n\}.$$

Example Let $f(x) = g(\sin x, \cos x)$. Then

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial a}\cos x - \frac{\partial g}{\partial b}\sin x$$

where we use a notation $g = g(a, b)$.

4.5 Differential

Consider a function $f : \mathbb{R}^2 \mapsto \mathbb{R}$. We try to compute an increment of a function if we move from the point (x_0, y_0) to the point $(x_0 + h, y_0 + k9)$, i.e., $\Delta f(x_0, y_0) = f(x_0 + h, y_0 + k) - f(x_0, y_0)$. It can be written as

$$\Delta f(x_0, y_0) = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) + f(x_0 + h, y_0) - f(x_0, y_0).$$

Assuming $|h|$ and $|k|$ are sufficiently small we can us an approximation

$$f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) \sim \frac{\partial f}{\partial x}(x_0 + h, y_0)k$$
$$f(x_0 + h, y_0) - f(x_0, y_0) \sim \frac{\partial f}{\partial y}(x_0, y_0)h$$

Moreover, $\frac{\partial f}{\partial x}(x_0 + h, y_0) \sim \frac{\partial f}{\partial x}(x_0, y_0)$ if $f \in C^1$. This yields

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) \sim \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k.$$

We denote by $\mathrm{d}x$ the change in the x coordinate and $\mathrm{d}y$ the change in the y coordinate.

Definition 4.9. *Let $f \in C^1$. Then*

$$\mathrm{d}f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\mathrm{d}x + \frac{\partial f}{\partial y}(x_0, y_0)\mathrm{d}y$$

is called the differential *of f at the point (x_0, y_0).*

The differential of a function can be used to determine approximate values. Let for example determine $\sqrt{(0.03)^2 + (2.89)^2}$. Consider a function $f(x, y) = \sqrt{x^2 + y^2}$. We have $\nabla f = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right)$. We choose $x_0 = 0$ and $y_0 = 3$. We have $\mathrm{d}x = 0.03$ and $\mathrm{d}y = -0.11$. It holds that

$$\sqrt{(0.03)^2 + (2.89)^2} \sim \sqrt{0^2 + 3^2} + 0 \cdot 0.03 + 1 \cdot (-0.11) = 2.89$$

It is worth to mention that $\mathrm{d}f(x_0, y_0) = \nabla f(x_0, y_0) \cdot (\mathrm{d}x, \mathrm{d}y)$.

Definition 4.10. *Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ have continuous partial derivatives at point (x_0, y_0). Then a* tangent plane *of the graph of f at point (x_0, y_0) is a plane with equation*

$$z = f(x_0, y_0) + \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0).$$

Example Let compute a tangent plane of the graph of f at point $(1, 2)$ for $f(x, y) = \sqrt{9 - x^2 - y^2}$. We have

$$\nabla f(x, y) = \left(-\frac{x}{\sqrt{9 - x^2 - y^2}}, -\frac{y}{\sqrt{9 - x^2 - y^2}} \right)$$

and $\nabla f(1, 2) = (-1/2, -1)$. Thus, the tangent plane is

$$z = 2 - 1/2(x - 1) - 1(y - 2) = 9/2 - x/2 - y.$$

4.6 The Taylor polynomial

An approximation by a differential is deduced above. In particular

$$f(x, y) \sim f(x_0, y_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0). \tag{4}$$

Recall that we use it to compute $\sqrt{(0.03)^2 + (2.89)^2}$.
The above considerations leads to the definition of the first-order Taylor polynomial at a point (x_0, y_0) as

$$T_1(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0).$$

Note that the graph of T_1 is also a tangent plane to the graph of the function f at the point (x_0, y_0) and it is the only plane which is the best approximation of the function near the point (x_0, y_0).

Definition 4.11. *We define the second order Taylor polynomial at a point (x_0, y_0) as*

$$\begin{aligned} T_2(x, y) = f(x_0, y_0) &+ \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + \\ &+ \frac{1}{2!}\left(\frac{\partial^2 f}{\partial x^2}(x_0, y_0)(x - x_0)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x - x_0)(y - y_0) \right. \\ &\left. \qquad + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)(y - y_0)^2 \right) \end{aligned}$$

Definition 4.12. *We define the third order Taylor polynomial at a point (x_0, y_0) as*

$$\begin{aligned} T_3(x, y) = f(x_0, y_0) &+ \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + \\ &+ \frac{1}{2!}\left(\frac{\partial^2 f}{\partial x^2}(x_0, y_0)(x - x_0)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x - x_0)(y - y_0) \right. \\ &\left. + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)(y - y_0)^2 \right) + \frac{1}{3!}\left(\frac{\partial^3 f}{\partial x^3}(x_0, y_0)(x - x_0)^3 + \right. \\ &\qquad 3\frac{\partial^3 f}{\partial x^2 \partial y}(x_0, y_0)(x - x_0)^2(y - y_0) + \\ &\qquad \left. + 3\frac{\partial^3 f}{\partial x \partial y^2}(x_0, y_0)(x - x_0)(y - y_0)^2 + \frac{\partial^3 f}{\partial y^3}(x_0, y_0)(y - y_0)^3 \right) \end{aligned}$$

Example We compute an approximate value $\sqrt{(0.03)^2+(2.89)^2}$ with the help of the second order Taylor polynomial. We choose $(x_0, y_0) = (0, 3)$ and we use notation $f(x, y) = \sqrt{x^2+y^2}$. We have $\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}$, $\frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}$, $\frac{\partial^2 f}{\partial x^2} = \frac{y^2}{(x^2+y^2)^{3/2}}$, $\frac{\partial^2 f}{\partial y^2} = \frac{x^2}{(x^2+y^2)^{3/2}}$, $\frac{\partial^2 f}{\partial x \partial y} = \frac{-xy}{(x^2+y^2)^{3/2}}$. We deduce that T_2 at $(0, 3)$ is

$$T_2(x, y) = 3 + (y-3) + \frac{1}{6}x^2$$

We get $T_2(0.03, 2.89) = 3 + (-0.11) + \frac{1}{6}0.0009 = 2.89015$.

Example It holds that

$$e^x \sin y \sim y + xy + \frac{1}{2}x^2 y - \frac{1}{6}y^3$$

assuming $|x|$ and $|y|$ are sufficiently small.
We compute the third order Taylor polynomial centered at $(x_0, y_0) = (0, 0)$. Denote $f(x, y) = e^x \sin y$. We have

$$\begin{aligned}
\frac{\partial f}{\partial x} &= e^x \sin y \quad \frac{\partial f}{\partial y} = e^x \cos y \\
\frac{\partial^2 f}{\partial x^2} &= e^x \sin y \quad \frac{\partial^2 f}{\partial y^2} = -e^x \sin y \quad \frac{\partial^2 f}{\partial x \partial y} = e^x \cos y \\
\frac{\partial^3 f}{\partial x^3} &= e^x \sin y \quad \frac{\partial^3 f}{\partial y \partial x^2} = e^x \cos y \\
\frac{\partial^3 f}{\partial y^3} &= -e^x \cos y \quad \frac{\partial^3 f}{\partial x \partial y^2} = -e^x \sin y.
\end{aligned}$$

Thus

$$\begin{aligned}
\frac{\partial f}{\partial x}(0,0) &= 0 \quad \frac{\partial f}{\partial y}(0,0) = 1 \\
\frac{\partial^2 f}{\partial x^2}(0,0) &= 0 \quad \frac{\partial^2 f}{\partial y^2}(0,0) = 0 \quad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 1 \\
\frac{\partial^3 f}{\partial x^3}(0,0) &= 0 \quad \frac{\partial^3 f}{\partial y \partial x^2}(0,0) = 1 \\
\frac{\partial^3 f}{\partial y^3}(0,0) &= -1 \quad \frac{\partial^3 f}{\partial x \partial y^2}(0,0) = 0.
\end{aligned}$$

and the corresponding third-order Taylor polynomial is indeed

$$y + xy + \frac{1}{2}x^2 y - \frac{1}{6}y^3.$$

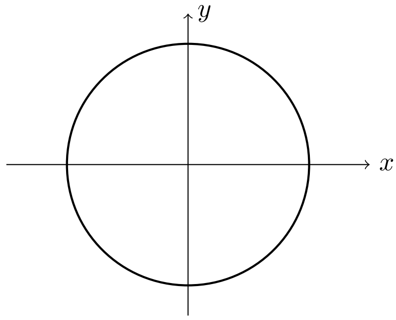
4.7 Implicit functions

Consider a set

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

The equation $x^2 + y^2 = 1$ defines two function $y_1(x)$ and $y_2(x)$ where

$$\begin{aligned}
y_1(x) &= \sqrt{1-x^2},\ \text{Dom}\ y_1(x) = [-1, 1], \\
y_2(x) &= -\sqrt{1-x^2},\ \text{Dom}\ y_2(x).
\end{aligned}$$



What if it is impossible to express y? Consider an equation

$$f(x, y) = 0.$$

What assumptions should be imposed in order to get uniquely defined function $y(x)$?

Theorem 4.2. *Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ and $(x_0, y_0) \in \mathbb{R}^2$ be given. If*

i) $f \in C^k$ *for some* $k \in \mathbb{N}$*,*

ii) $f(x_0, y_0) = 0$*,*

iii) $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$*,*

Then there is a uniquely determined function $y(x)$ of class C^k on a neighborhood of point x_0 such that $f(x, y(x)) = 0$ (precisely, there is $\epsilon > 0$ and a function $y(x)$ defined on $(x_0 - \varepsilon, x_0 + \varepsilon)$ such that $f(x, y(x)) = 0$.

Example Consider an equation

$$x^3 + y^3 - 3xy - 3 = 0.$$

Is there a function $y(x)$ determined by the given equation on the neighborhood of a point $(1, 2)$? According to the previous theorem, we have to verify three assumptions:
1, the function $f(x, y) = x^3 + y^3 - 3xy - 3$ should belong (at least) to C^1. That is true since $f(x, y)$ is a polynomial.
2, $f(1, 2)$ should be equal to zero (or, equivalently, the given equation should be satisfied at the given point). This is also true.
3, $\frac{\partial f}{\partial y} = 3y^2 - 3x$ and therefore $\frac{\partial f}{\partial y}(1, 2) = -3 \neq 0$ and the last assumption is also true.
As a result, there is a function $y(x)$ uniquely determined by the given equation in some neighborhood of point $x = 1$, $y = 2$.

Note that the last assumption in the implicit function theorem cannot be omited. Consider the first equation

$$x^2 + y^2 = 1$$

and let decide whether there is a function $y(x)$ given by that equation at the point $(1, 0)$. According to the picture, it is impossible (recall the vertical line test). The theorem may not be applied. Take $f(x, y) = x^2 + y^2 - 1$. We have

$$\frac{\partial f}{\partial y} = 2y, \ \frac{\partial f}{\partial y}(1, 0) = 0$$

and the third assumption is not fulfilled.
Or another example, consider a set

$$\{(x,y) \in \mathbb{R}^2,\ x^2 - y^2 = 0\}.$$

Is this set a graph of some function around a point $(0,0)$? Once again, we have $f(x,y) = x^2 - y^2$, $\frac{\partial f}{\partial y} = -2y$ and the last assumption of the implicit function theorem is not fulfilled.

Further analysis of the implicitly given function

In order to examine further qualitative properties of the given function we have to compute derivatives at the given points. The easiest method is to differentiate the given equation with respect to x (and to assume that y is in fact a function of x).
Example: Consider an equation

$$e^{2x} + e^y + x + 2y - 2 = 0.$$

This defines on a neighborhood of $(0,0)$ a function $y(x)$. Indeed, let $f(x,y) = e^{2x} + e^y + x + 2y - 2$. Then f is of class C^k for every $k \in \mathbb{N}$, $f(0,0) = 0$ and $\frac{\partial f}{\partial y} = e^y + 2$ which yields $\frac{\partial f}{\partial y}(0,0) = 3 \neq 0$. Let compute $y'''(0)$ (note that the third derivative exists as $f \in C^3$).

Let differentiate the equation with respect to x. We have

$$2e^{2x} + e^y y' + 1 + 2y' = 0$$

and we plug here $x = 0$ and $y = 0$ in order to get

$$2 + y'(0) + 1 + 2y'(0) = 0$$

which yields $y'(0) = -1$.
We differentiate once again with respect to x to get

$$4e^{2x} + e^y y'^2 + e^y y'' + 2y'' = 0$$

and we plug here $x = 0$, $y = 0$ and $y' = -1$. We get

$$4 + 1 + 3y''(0) = 0$$

yielding $y''(0) = -\frac{5}{3}$. We differentiate the equation for the third time in order to get

$$8e^{2x} + e^y y'^3 + e^y 2y'y'' + e^y y'y'' + e^y y''' + 2y''' = 0$$

and once again we plug there $x = 0$, $y = 0$, $y' = -1$ and $y'' = -\frac{5}{3}$. We get

$$8 - 1 + \frac{10}{3} + \frac{5}{3} + 3y''' = 0$$

which gives

$$y'''(0) = -4.$$

In particular, we may write

$$0 = \frac{\partial f(x, y(x))}{\partial x} = \frac{\partial f(x,y)}{\partial x} + \frac{\partial f(x,y)}{\partial y}\frac{\partial y}{\partial x}$$

which gives

$$y'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}.$$

4.8 Extremes

Similarly to the one-dimensional case, we talk about local and global extremes.

Definition 4.13. *Let $f : M \subset \mathbb{R}^2 \mapsto \mathbb{R}$. We say that f attains a* local maximum *at a point $(x_0, y_0) \in M^0$ if there is $r > 0$ such that $f(x_0, y_0) \geq f(x, y)$ for all $(x, y) \in B_r(x_0, y_0)$.*
We say that f attains a local minimum *at a point $(x_0, y_0) \in M^0$ if there is $r > 0$ such that $f(x_0, y_0) \leq f(x, y)$ for all $(x, y) \in B_r(x_0, y_0)$.*

Definition 4.14. *Let $f : M \subset \mathbb{R}^2 \mapsto \mathbb{R}$. We say that f attains its maximum on M at a point (x_0, y_0) if $f(x_0, y_0) \geq f(x, y)$ for all $(x, y) \in M$. Similarly, f attains its minimum on M at a point (x_0, y_0) if $f(x_0, y_0) \leq f(x, y)$ for all $(x, y) \in M$.*

Local extremes

Assume $f \in C^1$ Let f has a local extrem at (x_0, y_0). Then $g(x) = f(x, y_0)$ has also a local extreme at x_0 and, therefore, $g'(x_0) = 0$. Similarly, $h(y) = f(x_0, y)$ has a local extreme at y_0 and thus $h'(y_0) = 0$. This leads to the following observation.

Observation 4.6. *Let $f \in C^1$ have a local extreme at (x_0, y_0). Then $\nabla f(x_0, y_0) = 0$.*

Definition 4.15. *A point $(x_0, y_0) \in M$ such that $\nabla f(x_0, y_0) = 0$ is called a* stationary point.

How to find all local extremes of given function?
Step 1: determine the stationary point.
Step 2: examine the possible extremes in the stationary point.
Reminder: in the one-dimensional case one has to treat the sign of the second derivative in order to decide if there is an extreme in a stationary point.

Example Let find all stationary points of $f(x, y) = x^2 - y^2$. We have $\nabla f(x, y) = (2x, -2y)$ and therefore the only stationary point is $(x_0, y_0) = (0, 0)$. Is there a maximum or minimum?

Definition 4.16. *Let $f \in C^2$. Then the Hess matrix of f is*

$$Hf = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}.$$

The determinant of Hf is called Hessian.

Observation 4.7. *Let $f \in C^2$ and let (x_0, y_0) be its stationary point. Then:*

1. *Let $\det Hf(x_0, y_0) > 0$, then*

   i if $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ then f attains a local minimum at (x_0, y_0),

   ii if $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ then f attains a local maximum at (x_0, y_0).

2. *Let $\det Hf(x_0, y_0) < 0$, then f does not have an extreme at (x_0, y_0) (saddle point).*

3. *Otherwise, we do not know anything.*

Example: Let go back to $f(x, y) = x^2 - y^2$. We already know that $(x_0, y_0) = (0, 0)$ is a stationary point. We have

$$Hf = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Thus $\det Hf(0, 0) = -4$ and there is no extreme at $(0, 0)$.

Another example Determine all local extremes of

$$f(x,y) = x^3 + 3xy^2 - 15x - 12y.$$

Step 1, stationary points:

$$\nabla f(x,y) = (3x^2 + 3y^2 - 15, 6xy - 12)$$

and we stationary points are solutions to

$$\begin{aligned} 3x^2 + 3y^2 - 15 &= 0 \\ 6xy - 12 &= 0 \end{aligned}$$

which is equivalent to

$$\begin{aligned} x^2 + y^2 - 5 &= 0 \\ xy &= 2. \end{aligned}$$

We deduce from the second equation that x and y are different from zero. The second equation yields $x = \frac{2}{y}$. We plug this into the first equation to deduce

$$\frac{4}{y^2} + y^2 - 5 = 0$$

which is equivalent to

$$y^4 - 5y^2 + 4 = 0.$$

We have $y^2 = 4$, $y^2 = 1$ and therefore there are four stationary points

$$A = (-1,-2),\ B = (1,2),\ C = (2,1),\ D = (-2,-1).$$

Step 2: We have

$$Hf = \begin{pmatrix} 6x & 6y \\ 6y & 6x \end{pmatrix}$$

Further,

$$Hf(A) = \begin{pmatrix} -6 & -12 \\ -12 & -6 \end{pmatrix},\ \det Hf(A) = -108$$

and A is a saddle point.

$$Hf(B) = \begin{pmatrix} 6 & 12 \\ 12 & 6 \end{pmatrix},\ \det Hf(B) = -108$$

and B is a saddle point.

$$Hf(C) = \begin{pmatrix} 12 & 6 \\ 6 & 12 \end{pmatrix},\ \det Hf(C) = 108$$

and C is a point of a local minimum. The value of the local minimum is $f(C) = -28$.

$$Hf(D) = \begin{pmatrix} -12 & -6 \\ -6 & -12 \end{pmatrix},\ \det Hf(D) = 108$$

and D is a point of a local maximum. The value of the local maximum is $f(D) = 28$.

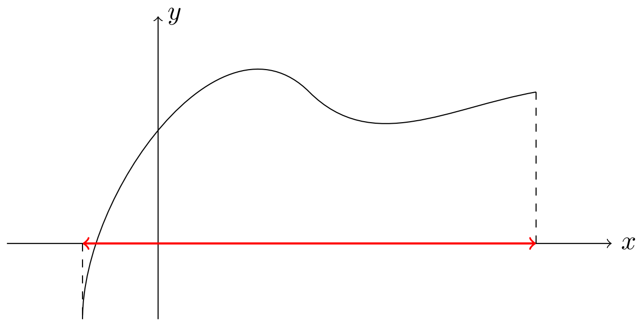
Global extremes

Definition 4.17. *A set $M \subset \mathbb{R}^2$ is* bounded *if there is $r > 0$ such that $M \subset B_r(0,0)$.*

Observation 4.8. *Let $f : M \subset \mathbb{R}^2 \mapsto \mathbb{R}$ be continuous. Let M be a bounded and closed set. Then there is (x_0, y_0) where f attains its minimum on M and there is (x_1, y_1) where M attains its maximum.*

One-dimensional case, reminder

Consider this function: One has to consider separately the interior of M and the 'boundary' of



M. Although the function whose graph is in the picture has two stationary points, just one of them is a point of a global extreme. The point of the global minimum is on the edge of M.

Boundary in the two dimensional case might be a bit complicated. In order to find extremes here, we use the Lagrange multipliers method.

Theorem 4.3. *Let $f : \mathrm{Dom}\, f \subset \mathbb{R}^2 \mapsto \mathbb{R}$ be of class C^1 and let it be defined on the neighborhood of a set M which is given as*

$$M = \{(x,y) \subset \mathbb{R}^2,\ g(x,y) = 0\}$$

for some function $g \in C^1$. Let $\nabla g \neq 0$. If there is an extreme of f with respect to the set M then there exists $\lambda \in \mathbb{R}$ such that

$$\nabla f + \lambda \nabla g = 0.$$

Example We show how to determine a maximum and minimum of $f(x,y) = -y^2 + x^2 + \frac{4}{3}x^3$ on a set $M = \{(x,y) \in \mathbb{R}^2,\ x^2 + y^2 = 4\}$.

The first question is whether there is a maximum and minimum. Our first claim is that the set M is closed. Why? Recall Observation 3.8. We define $g(x,y) = x^2 + y^2 - 4$ and then the set M is $g^{-1}(\{0\})$ and since $\{0\} \subset \mathbb{R}^2$ is a closed set, we deduce that M is also closed. Further, M is bounded since $M \subset B_3(0,0)$. Therefore, according to the very first observation of this talk there has to be a maximum and minimum of f on M.

Further, it holds that $\nabla g \neq 0$ for every $(x,y) \neq (0,0)$. Note that $(0,0) \notin M$ and thus we may use the Lagrange multipliers. We have

$$\nabla f(x,y) = (2x + 4x^2, -2y), \quad \nabla g(x,y) = (2x, 2y).$$

We end up with a system

$$\begin{aligned} 2x + 4x^2 + 2\lambda x &= 0 \\ -2y + 2\lambda y &= 0 \\ x^2 + y^2 &= 4. \end{aligned}$$

We deduce from the second equation that $y(2\lambda - 2) = 0$ and we get that either $y = 0$ or $\lambda = 1$. Consider first the case $y = 0$. Then the last equation yields $x^2 = 4$ and therefore $x = \pm 2$. We get two 'stationary' points

$$A = (2, 0), \ B = (-2, 0).$$

Next, assume $\lambda = 1$. The first equation then yields

$$4x + 4x^2 = 0$$

which gives $x = 0$ or $x = -1$.
Let $x = 0$. The last equation is then $y^2 = 4$ and we get $y = \pm 2$ and another two stationary points

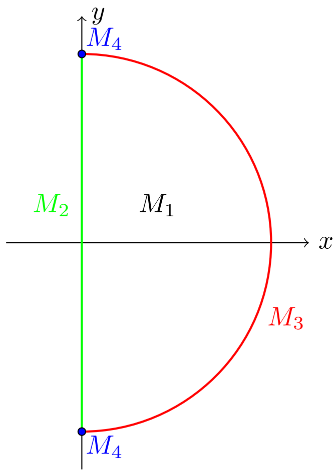
$$C = (0, 2), \ D = (0, -2).$$

Finally, let $x = -1$. Then we get $y^2 = 3$ and $y = \pm 3$ and we deduce another two stationary points

$$E = (-1, \sqrt{3}), \ D = (-1, -\sqrt{3}).$$

We have $f(A) = \frac{44}{3}$, $f(B) = -\frac{20}{3}$, $f(C) = -4$, $f(D) = -4$, $f(E) = -\frac{10}{3}$ and $f(F) = -\frac{10}{3}$. We deduce that the maximum is attained at the point $(2, 0)$ and its value is $\frac{44}{3}$, the minimum is attained at the point $(-2, 0)$ and its value is $-\frac{20}{3}$.

Example We find extremes of $f(x, y) = x^2 + y^2 - 12x + 16y$ on a set $M = \{(x, y) \subset \mathbb{R}^2, \ x^2 + y^2 \leq 25, x \geq 0\}$.



We dismantle the set into four pieces

$$\begin{aligned}
M_1 &= \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 25, x > 0\}, \\
M_2 &= \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 25, x = 0\}, \\
M_3 &= \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 25, x > 0\}, \\
M_4 &= \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 25, x = 0\}.
\end{aligned}$$

and we takcle each subset separately:
- Stationary points in $M_1 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 25, x > 0\}$:

We solve $\nabla f = 0$ which is

$$2x - 12 = 0$$
$$2y + 16 = 0$$

Therefore the stationary point is $(6, -8)$. However, $6^2 + (-8)^2 = 100 > 25$ and this point does not belong to M_1.

- Stationary points in $M_2 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 25, x = 0\}$:

We are going to consider a function $f(x, y)$ on line $x = 0$. Therefore it is enough to examine function $f(0, y) =: h(y)$. We have

$$h(y) = y^2 + 16y$$

and therefore $h'(y) = 2y + 16$. The resulting stationary point is $x = 0$, $y = -8$. However, $(-8)^2 > 25$ and the point $(0, -8)$ does not belong to M_2.

- Stationary points in $M_3 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 25, x = 0\}$:

There is a constraint $g(x, y) = x^2 + y^2 - 25$. We have $\nabla g = (2x, 2y)$ and we have $\nabla g \neq 0$ for every $(x, y) \in M_3$. The system $\nabla f + \lambda \nabla g = 0$ complemented with $g = 0$ has form

$$2x - 12 + 2x\lambda = 0$$
$$2y + 16 + 2y\lambda = 0$$
$$x^2 + y^2 = 25.$$

We may deduce that $y \neq 0$ (otherwise the second equation cannot be true) and $\lambda \neq 0$ (otherwise $x = 6$, $y = -8$ and the last equation is not fulfilled. The first and second equation might be rewritten as

$$x\lambda = 6 - x$$
$$y\lambda = -8 - y$$

and we divide the first equation by the second to get

$$\frac{\lambda x}{\lambda y} = \frac{6 - x}{-8 - y}.$$

This yields

$$\frac{x}{y} = \frac{x - 6}{8 + y}$$

and

$$8x + xy = xy - 6y$$

and therefore

$$y = -\frac{4}{3}x.$$

We plug this into the last equation ($x^2 + y^2 = 25$) to get

$$x^2 + \frac{16}{9}x^2 = 25$$

which yields $x = \pm 3$. Therefore we have two stationary points $(-3, 4)$ and $(3, -4)$, however, the first one does not belong to M_3. So we take into consideration only $A = (3, -4)$.

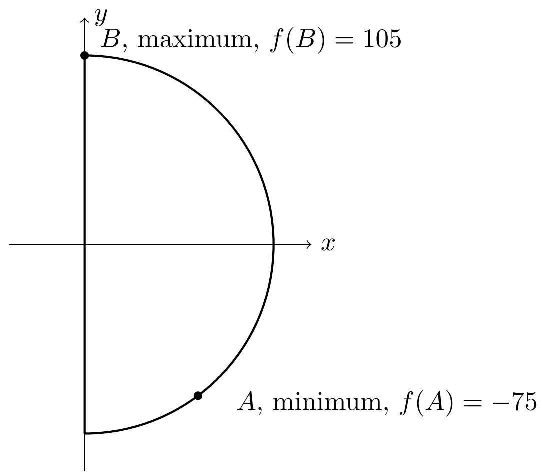
- Stationary points in $M_4 = \{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 25,\ x = 0\}$: This set consists only of two points. Indeed, let both equations holds at once. Then necessarily

$$y^2 = 25$$

and we have two points $B = (0, 5)$ and $C = (0, -5)$. These two points have to be considered as there might appear global extremes (although these points are not stationary).
- Final evaluation: We have just three points where the extremes might be attained:
$A = (3, -4)$, $B = (0, 5)$ and $C = (0, -5)$. We have

$$\begin{aligned} f(A) &= -75 \\ f(B) &= 105 \\ f(C) &= -55. \end{aligned}$$

We deduce that the minimum of f on set M is attained at the point $(3, -4)$ and its value is -75, the maximum of f on set M is attained at point $(0, 5)$ and its value is 105.



The least square method

We will solve the following exercise: Assume that the cost of a car (of one given type) depends linearly on its age, i.e.,

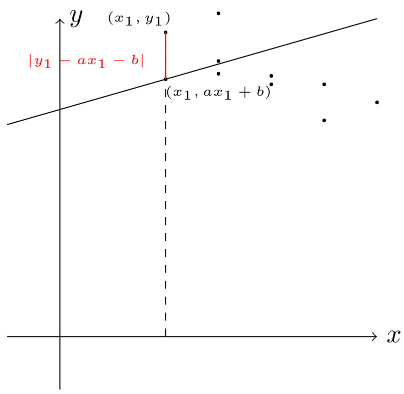
$$y = ax + b, \ a, b \in \mathbb{R},$$

where y is the price of a car and x is its age.
Our aim now is to determine this function (constants a and b) from the given sets of data. Below we have a table of particular cars (their price does not follow strictly the above rule since the price come from the free market)

x	2	3	3	3	4	4	5	5	6
y	28.7	24.8	26.0	30.5	23.8	24.6	23.8	20.4	22.1

To find the line which fits best to the given data, we use the least squares method. This means that we are going to minimize the 'distance' between the line $ax + b$ and the given data. We define such distance as sum of squares:

$$|y_1 - ax_1 - b|^2 + |y_2 - ax_2 - b|^2 + \ldots + |y_n - ax_n - b|^2 = \sum_{i=1}^{n} |y_i - ax_i - b|^2.$$



This sum of squares in infact a function f of variables a and b of the form

$$f(a,b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

and we are going to minimize this sum of squares. We compute the partial derivative

$$\frac{\partial f}{\partial a} = -2\sum_{i=1}^{n}(y_i - ax_i - b)x_i, \quad \frac{\partial f}{\partial b} = -2\sum_{i=1}^{n}(y_i - ax_i - b).$$

and we deduce that the stationary point of this function has to fulfill

$$\sum_{i=1}^{n}(y_i - ax_i - b)x_i = 0$$
$$\sum_{i=1}^{n}(y_i - ax_i - b) = 0.$$

Recall that unknowns are a and b. We reformulate this into

$$\left(\sum_{i=1}^{n} x_i^2\right) a + \left(\sum_{i=1}^{n} x_i\right) b = \sum_{i=1}^{n} x_i y_i$$
$$\left(\sum_{i=1}^{n} x_i\right) a + nb = \sum_{i=1}^{n} y_i.$$

Recall our example

x	2	3	3	3	4	4	5	5	6
y	28.7	24.8	26.0	30.5	23.8	24.6	23.8	20.4	22.1

where we have

$$n = 9,\ \sum_{i=1}^{9} x_i = 35,\ \sum_{i=1}^{9} x_i^2 = 149,\ \sum_{i=1}^{9} y_i = 224.7,\ \sum_{i=1}^{9} x_i y_i = 848.5.$$

We and up with equation

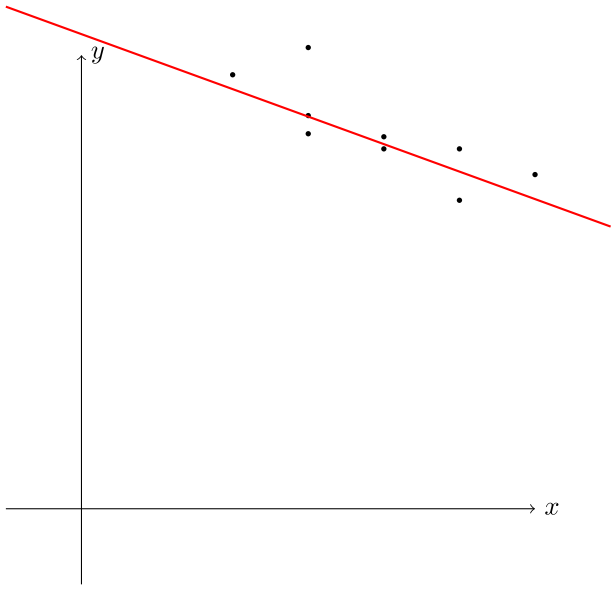
$$\begin{aligned} 149a + 35b &= 848.5 \\ 35a + 9b &= 224.7 \end{aligned}$$

which has (approximate) solution

$$a = -2.02, \quad b = 32.8.$$

Thus, the desired line has equation

$$y = -2.02x + 32.8$$



4.9 Double Integrals

Let start with a motivation - double integral over a rectangle: Assume we have a constant function $f(x,y) \equiv k > 0$ on a set $M = [a,b] \times [c,d]$. What is the volume of a prism $[a,b] \times [c,d] \times [0,k]$? Simple answer is $(b-a)(c-d)k$. In this particular case we write $\int_{[a,b]\times[c,d]} f(x,y)\ \mathrm{d}x\mathrm{d}y = (b-a)(c-d)k$.

Let M be a rectangle $[a,b] \times [c,d]$ and let $f(x,y)$ be a positive function defined on M. The value of the integral

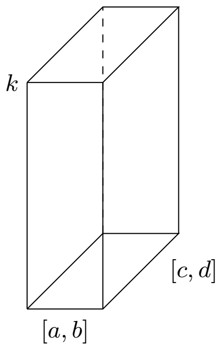
$$\int_M f(x,y)\ \mathrm{d}x\mathrm{d}y$$

is a volume of a set

$$S = \{(x,y,z) \in \mathbb{R}^3,\ (x,y) \in M,\ 0 \le z \le f(x,y)\}.$$

Observation Let f be continuous function on a rectangle M. Then there is an integral

$$\int_M f(x,y)\ \mathrm{d}x\mathrm{d}y.$$



Remark 4.2 (measurable sets)**.** *It is not necessary to define integrals only over rectangles. In particular, the set M can be 'almost arbitrary' and then the meaning of integral is the same as in the previous slide. The only condition is that the integral*

$$\int_M 1 \, \mathrm{d}x\mathrm{d}y$$

has value (and it might be even infinity). Such sets are called measurable sets *and we will not define them in the scope of this class. Let me just mention that not every set is measurable. On the other hand, it is very difficult to construct a non-measurable set. All sets appearing in this class are measurable. Interested students might look for the Banach-Tarski theorem.*

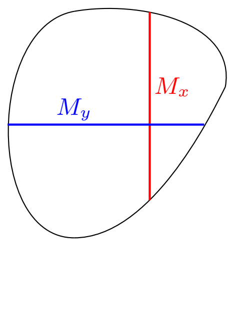
Remark 4.3 (measurable functions)**.** *Similarly, it is not necessary to define integrals only for continuous functions. Once again, there are functions called 'measurable functions' (and all continuous functions are measurable as well). And, similarly as before, it is very difficult to construct a non-measurable functions. In particular, every 'well-behaved' function is a measurable function and all functions appearing in this class are measurable.*

Definition 4.18. *Let $M \subset \mathbb{R}^2$. We define a* vertical cross-section *as*

$$M_x = \{y \in \mathbb{R}, \ (x, y) \in M\}.$$

Similarly, we define a horizontal cross-section *as*

$$M_y = \{x \in \mathbb{R}, \ (x, y) \in M\}.$$



Theorem 4.4 (Fubini)**.** *Let $M \subset \mathbb{R}^2$ is a measurable set and $f : M \to \mathbb{R}$ be a measurable function. Then*

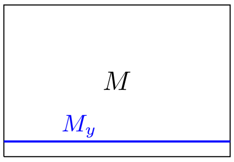
$$\int_M f(x,y)\ \mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}} \left(\int_{M_x} f(x,y)\ \mathrm{d}y \right)\ \mathrm{d}x = \int_{\mathbb{R}} \left(\int_{M_y} f(x,y)\ \mathrm{d}x \right)\ \mathrm{d}y.$$

assuming that the integral on the left hand side is well defined.

Example Compute

$$\int_M 5x^2y - 2y^3\ \mathrm{d}x\mathrm{d}y, \quad M = [2,5] \times [1,3].$$

We use notation $f(x,y) = 5x^2y - 2y^3$ and we have $M_y = [2,5]$ for $y \in [1,3]$ and $M_y = \emptyset$



otherwise. Thus

$$\int_M f(x,y)\ \mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}} \int_{M_y} f(x,y)\ \mathrm{d}x\mathrm{d}y = \int_1^3 \int_2^5 f(x,y)\ \mathrm{d}x\mathrm{d}y.$$

We have

$$\begin{aligned}\int_M 5x^2y - 2y^3\ \mathrm{d}x\mathrm{d}y &= \int_1^3 \left(\int_2^5 5x^2y - 2y^3\ \mathrm{d}x \right)\ \mathrm{d}y \\ &= \int_1^3 \left[\frac{5x^3y}{3} - 2xy^3 \right]_{x=2}^{x=5} \mathrm{d}y = \int_1^3 \frac{625}{3}y - 10y^3 - \frac{40}{3}y + 4y^3\ \mathrm{d}y \\ &= \int_1^3 195y - 6y^3\ \mathrm{d}y = \left[\frac{195}{2}y^2 - \frac{3}{2}y^4 \right]_1^3 = 660.\end{aligned}$$

Example Let compute integral

$$\int_M 2xe^y\ \mathrm{d}x\mathrm{d}y, \quad M = [0,2] \times [0,1].$$

We use the Fubini theorem to deduce

$$\int_M 2xe^y\ \mathrm{d}x\mathrm{d}y = \int_0^2 \left(\int_0^1 2xe^y\ \mathrm{d}y \right)\ \mathrm{d}x =$$

and since $2x$ is not a function of y, it can be put in front of the inner integral to obtain

$$= \int_0^2 2x \left(\int_0^1 e^y\ \mathrm{d}y \right)\ \mathrm{d}x = \int_0^2 2x\ \mathrm{d}x \int_0^1 e^y\ \mathrm{d}y.$$

Observation 4.9. *Let $f(x,y) = g(x)h(y)$ and let $M = [a,b] \times [c,d]$. Then*

$$\int_M f(x,y)\ \mathrm{d}x\mathrm{d}y = \int_a^b g(x)\ \mathrm{d}x \int_c^d h(y)\ \mathrm{d}y.$$

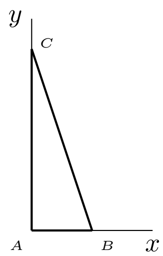
Back to the given integral. We have

$$\int_M 2xe^y \, \mathrm{d}x\mathrm{d}y = \int_0^2 2x \, \mathrm{d}x \int_0^1 e^y \, \mathrm{d}y = \left[x^2\right]_0^2 \left[e^y\right]_0^1 = 4(e-1)$$

Example Compute

$$\int_M x^2 + xy - 1 \, \mathrm{d}x\mathrm{d}y$$

where M is a triangle with vertices $A = \langle 0,0\rangle$, $B = \langle 2,0\rangle$ and $C = \langle 0,6\rangle$. Recall that the line



BC has an equation $y = 6 - 3x$. Therefore, the vertical cross-section has form $M_x = (0, 6-3x)$. and we deduce

$$\begin{aligned}
\int_M x^2 + xy - 1 \, \mathrm{d}x\mathrm{d}y &= \int_0^2 \left(\int_0^{6-3x} x^2 + xy - 1 \, \mathrm{d}y \right) \, \mathrm{d}x \\
&= \int_0^2 \left[x^2 y + \frac{xy^2}{2} - y \right]_{y=0}^{y=6-3x} \, \mathrm{d}x \\
&= \int_0^2 6x^2 - 3x^3 + 18x - 18x^2 + \frac{9}{2}x^3 - 6 + 3x \, \mathrm{d}x \\
&= \left[\frac{3}{8}x^4 - 4x^3 + \frac{21}{2}x^2 - 6x \right]_0^2 = 6 - 32 + 42 - 12 = 4.
\end{aligned}$$

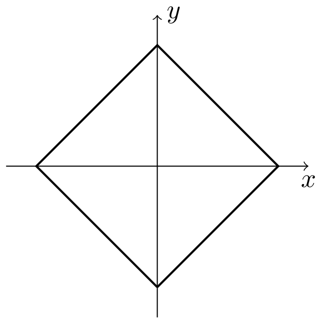
Observation 4.10 (Properties of integral)**.** *The following holds:*

- *Let f and g be (measurable) functions of two variables, $M \subset \mathbb{R}^2$ measurable set and $\alpha \in \mathbb{R}$. Then*
$$\int_M \alpha f(x,y) + g(x,y) \, \mathrm{d}x\mathrm{d}y = \alpha \int_M f(x,y) \, \mathrm{d}x\mathrm{d}y + \int_M g(x,y) \, \mathrm{d}x\mathrm{d}y.$$
- *Let $M = \bigcup_{i=1}^n M_i$. Then*
$$\int_M f(x,y) \, \mathrm{d}x\mathrm{d}y = \sum_{i=1}^n \int_{M_i} f(x,y) \, \mathrm{d}x\mathrm{d}y.$$
- *Let f be a measurable function, $M \subset \mathbb{R}^2$ be a measurable set and let f be non-negative on M (i.e. $f(x,y) \geq 0$ for all $(x,y) \in M$). Then*
$$\int_M f(x,y) \, \mathrm{d}x\mathrm{d}y \geq 0.$$

Example Compute

$$\int_M (x+y)^2 \; \mathrm{d}x\mathrm{d}y$$

where M is a square with vertices $A = \langle 0,-1\rangle$, $B = \langle 1,0\rangle$, $C = \langle 0,1\rangle$, $D = \langle -1,0\rangle$. Here we



divide M into two subsets, $M_1 = M \cap \{x < 0\}$ and $M_2 = M \cap \{x \geq 0\}$. We have

$$\begin{aligned}(M_1)_x &= (-x-1, x+1) \text{ for } x \in (-1,0),\\ (M_2)_x &= (x-1, -x+1) \text{ for } x \in [0,1).\end{aligned}$$

Therefore,

$$\begin{aligned}\int_M (x+y)^2 \; \mathrm{d}x\mathrm{d}y = \int_{-1}^{0} \left(\int_{-x-1}^{x+1} x^2 + 2xy + y^2 \; \mathrm{d}y \right) \; \mathrm{d}x \\ + \int_0^1 \left(\int_{x-1}^{-x+1} x^2 + 2xy + y^2 \; \mathrm{d}y \right) \; \mathrm{d}x.\end{aligned}$$

We compute

$$\begin{aligned}\int_{-1}^{0} \left(\int_{-x-1}^{x+1} x^2 + 2xy + y^2 \; \mathrm{d}y \right) \; \mathrm{d}x &= \int_{-1}^{0} \left[x^2 y + xy^2 + \frac{y^3}{3} \right]_{y=-x-1}^{y=x+1} \; \mathrm{d}x \\ &= \int_{-1}^{0} \frac{8}{3}x^3 + 4x^2 + 2x + \frac{2}{3} \; \mathrm{d}x = \left[\frac{2}{3}x^4 + \frac{4}{3}x^3 + x^2 + \frac{2}{3}x \right]_{-1}^{0} = \frac{1}{3}.\end{aligned}$$

Similarly

$$\begin{aligned}\int_{0}^{1} \left(\int_{x-1}^{1-x} x^2 + 2xy + y^2 \; \mathrm{d}y \right) \; \mathrm{d}x &= \int_{0}^{1} \left[x^2 y + xy^2 + \frac{y^3}{3} \right]_{y=x-1}^{y=1-x} \; \mathrm{d}x \\ &= \int_{0}^{1} -\frac{8}{3}x^3 + 4x^2 - 2x + \frac{2}{3} \; \mathrm{d}x = \left[-\frac{2}{3}x^4 + \frac{4}{3}x^3 - x^2 + \frac{2}{3}x \right]_{0}^{1} = \frac{1}{3}.\end{aligned}$$

Eventually, we obtain

$$\int_M (x+y)^2 \; \mathrm{d}x\mathrm{d}y = \frac{2}{3}.$$

4.10 Change of variables

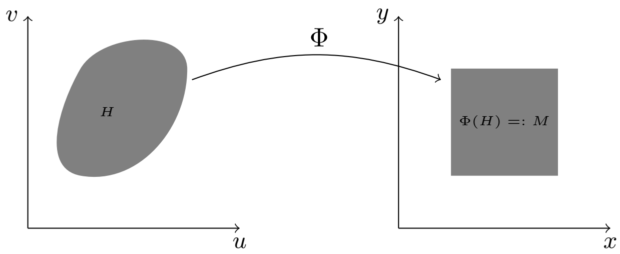
Recall one of the previous exercises:

$$\int_M (x+y)^2 \, \mathrm{d}x\mathrm{d}y, M \text{ is a square with vertices } A = (0,-1),$$
$$B = (1,0),\ C = (0,1),\ D = (-1,0).$$

The value of the integral is $\frac{2}{3}$ – that was deduced in the previous section, however, the computation was quite cumbersome. This time we present one better method how to compute the integral.
Recall, that the one-dimensional substitution method works in the following way

$$\int_a^b f(t)\ \mathrm{d}t = \int_\alpha^\beta f(\varphi(x))\varphi'(x)\ \mathrm{d}x.$$

This time, we consider a mapping $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$, $\Phi(u,v) = (\varphi(u,v), \psi(u,v))$ and we assume that



$x = \varphi(u,v)$, $y = \psi(u,v)$.

Definition 4.19. *A mapping $\Phi : H \subset \mathbb{R}^2 \mapsto \mathbb{R}^2$ satisfying*

- $\Phi \in C^1$,
- Φ *is an injection,*
- *The Jacobian matrix of Φ is regular,*

is called a regular mapping.

Definition 4.20. *Let $\Phi : H \subset \mathbb{R}^2 \mapsto \mathbb{R}^2$ have components $\varphi(u,v)$ and $\psi(u,v)$. Then the* Jacobian matrix *of Φ is a matrix*

$$J\Phi(u,v) = \begin{pmatrix} \frac{\partial\varphi}{\partial u}(u,v) & \frac{\partial\varphi}{\partial v}(u,v) \\ \frac{\partial\psi}{\partial u}(u,v) & \frac{\partial\psi}{\partial v}(u,v) \end{pmatrix}.$$

Its determinant is then called the Jacobian determinant.

Theorem 4.5. *Let $f(x,y)$ is a measurable function on $D \subset \mathbb{R}^2$ and let $\Phi = (\varphi,\psi) : H \subset \mathbb{R}^2 \mapsto M$ is a regular mapping. Then*

$$\int_M f(x,y)\ \mathrm{d}x\mathrm{d}y = \int_H f(\varphi(u,v), \psi(u,v))|\det J\Phi(u,v)|\ \mathrm{d}u\mathrm{d}v.$$

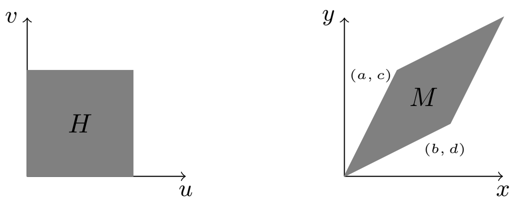
To show the role of the Jacobian determinant we consider a mapping

$$x = au + bv =: \varphi(u, v)$$
$$y = cu + dv =: \psi(u, v).$$

Here we have that

$$J\Phi(u, v) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Let $H = (0, 1) \times (0, 1)$. Then M is a parallelogram with sides represented by vectors (a, c) and (b, d).



The area of H is $\int_H 1 \; \mathrm{d}u\mathrm{d}v = 1$ and the area of M is $\int_M 1 \; \mathrm{d}x\mathrm{d}y = \left| \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right|$. Indeed, the area of the parallelogram is equal to $S = \sin\alpha \|(a, c)\| \|(b, d)\|$ and we may compute

$$\begin{aligned} S^2 = \sin^2\alpha \|(a, c)\|^2 \|(b, d)\|^2 &= \left(1 - \cos^2\alpha\right) \|(a, c)\|^2 \|(b, d)\|^2 \\ &= \|(a, c)\|^2 \|(b, d)\|^2 - ((a, c) \cdot (b, d))^2 \\ &= (a^2 + c^2)(b^2 + d^2) - (ab + cd)^2 = a^2 d^2 + c^2 b^2 - 2abcd \\ &= (ad - bc)^2. \end{aligned}$$

Therefore, there has to be a factor $|\det J\Phi|$ in order to get

$$\int_M 1 \; \mathrm{d}x\mathrm{d}y = \int_H 1 |\det J\Phi| \; \mathrm{d}u\mathrm{d}v.$$

Example: Let compute (once again) the integral

$$\int_M (x + y)^2 \; \mathrm{d}x\mathrm{d}y$$

where $M = \{(x, y) \in \mathbb{R}^2, \; -1 \le x + y \le 1, \; -1 \le x - y \le 1\}$. We establish new variables

$$u = x + y$$
$$v = x - y$$

and we deduce that

$$x = \frac{1}{2}(u + v)$$
$$y = \frac{1}{2}(u - v).$$

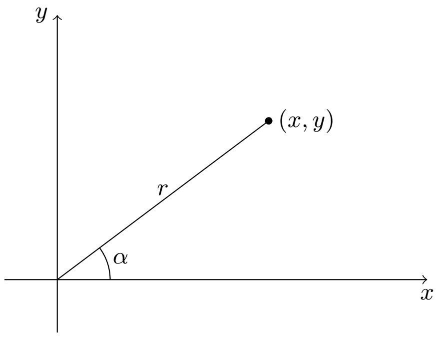
Thus, in this case, $\Phi(u,v) = \left(\frac{1}{2}(u+v), \frac{1}{2}(u-v)\right)$ and it holds that

$$J\Phi(u,v) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}, \ |\det J\Phi(u,v)| = \frac{1}{2}.$$

Let also mention that $\Phi(M) = \{(u,v) \in \mathbb{R}^2,\ u \in [-1,1],\ v \in [-1,1]\}$. Therefore we deduce that

$$\int_M (x+y)^2 \ \mathrm{d}x\mathrm{d}y = \int_{[-1,1]^2} u^2 \frac{1}{2} \ \mathrm{d}u\mathrm{d}v = \frac{1}{2}\int_{-1}^{1} u^2 \ \mathrm{d}u \int_{-1}^{1} 1 \ \mathrm{d}v = \int_{-1}^{1} u^2 \ \mathrm{d}u = \left[\frac{u^3}{3}\right]_{-1}^{1} = \frac{2}{3}.$$

4.11 Polar coordinates



We have

$$x = r\cos\alpha$$
$$y = r\sin\alpha.$$

(and also $r = \sqrt{x^2+y^2}$). Therefore we establish $\Phi(r,\alpha) = (r\cos\alpha, r\sin\alpha)$ and we infer

$$J\Phi(r,\alpha) = \begin{bmatrix} \cos\alpha & -r\sin\alpha \\ \sin\alpha & r\cos\alpha \end{bmatrix}$$

and

$$\det J\Phi(r,\alpha) = r\cos^2\alpha + r\sin^2\alpha = r.$$

Thus,

$$\int_M f(x,y) \ \mathrm{d}x\mathrm{d}y = \int_{\Phi^{-1}(M)} f(r\cos\alpha, r\sin\alpha) r \ \mathrm{d}r\mathrm{d}\alpha.$$

Example: Volume of a ball with radius R can be computed as twice the integral

$$\int_M \sqrt{R^2 - (x^2+y^2)} \ \mathrm{d}x\mathrm{d}y$$

where $M = \{(x,y) \in \mathbb{R}^2, x^2+y^2 \le R^2\}$. We have

$$\int_M \sqrt{R^2-(x^2+y^2)} \ \mathrm{d}x\mathrm{d}y = \int_0^{2\pi}\int_0^R \sqrt{R^2-r^2}r \ \mathrm{d}r\mathrm{d}\alpha$$

$$= 2\pi \int_0^R \sqrt{R^2-r^2}r \ \mathrm{d}r$$

and we use a (one-dimensional) substitution $t = R^2 - r^2$. In that case $\mathrm{d}t = -2r\ \mathrm{d}r$ and we have

$$2\pi \int_0^R \sqrt{R^2 - r^2} r\ \mathrm{d}r = -\pi \int_{R^2}^0 \sqrt{t}\ \mathrm{d}t = \pi \int_0^{R^2} \sqrt{t}\ \mathrm{d}t = \pi \left[\frac{t^{3/2}}{\frac{3}{2}}\right]_0^{R^2} = \pi \frac{2}{3} R^3.$$

Note that this is just one half of the demanded volume. Therefore, we have just deduced the well known relation

$$V = \frac{4}{3}\pi R^3.$$

Example Let compute an area of the set M which is given by the following conditions:

$$(x^2 + y^2)^2 \leq 2xy,\ x \geq 0,\ y \geq 0.$$

We use the polar coordinates. The second and third condition yields $\alpha \in [0, \pi/2]$. Next, we plug the polar coordinates into the first condition to deduce

$$r^4 \leq 2r^2 \cos\alpha \sin\alpha$$

and since $r > 0$, we may divide the inequality by r to get

$$r^2 \leq 2\cos\alpha\sin\alpha$$

and thus (recall that $\sin(2\alpha) = 2\sin\alpha\cos\alpha$)

$$r \leq \sqrt{\sin(2\alpha)}.$$

We deduce that

$$\int_M 1\ \mathrm{d}x\mathrm{d}y = \int_0^{\pi/2} \int_0^{\sqrt{\sin(2\alpha)}} r\ \mathrm{d}r\mathrm{d}\alpha = \int_0^{\pi/2} \left[\frac{r^2}{2}\right]_{r=0}^{r=\sqrt{\sin(2\alpha)}} \mathrm{d}\alpha = \frac{1}{2}\int_0^{\pi/2} \sin(2\alpha)\ \mathrm{d}\alpha$$
$$= -\frac{1}{4}\left[\cos(2\alpha)\right]_0^{\pi/2} = \frac{1}{2}$$

Example:Adjusted polar coordinates
Let compute an area of an ellipse which is given as

$$M = \left\{(x, y) \in \mathbb{R}^2,\ \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\}$$

for some positive reals a and b.
We define

$$x = ar\cos\alpha =: \varphi(r, \alpha)$$
$$y = br\sin\alpha =: \psi(r, \alpha)$$

It holds that $\Phi^{-1}(M) = (0, 1) \times (0, 2\pi)$. Furthermore, we have

$$J\Phi(r, \alpha) = \begin{pmatrix} a\cos\alpha & -ar\sin\alpha \\ b\sin\alpha & br\cos\alpha \end{pmatrix}$$

and therefore $\det J\Phi(r, \alpha) = abr$. We get

$$\int_M 1\ \mathrm{d}x\mathrm{d}y = \int_{(0,1)\times(0,2\pi)} abr\ \mathrm{d}r\mathrm{d}\alpha = \int_0^1 abr\mathrm{d}r \int_0^{2\pi} 1\ \mathrm{d}\alpha = \pi ab.$$

4.12 The Laplace integral

We use polar coordinates in order to deduce the value of integral

$$\int_0^\infty e^{-ax^2} \, \mathrm{d}x, \; a > 0.$$

Note that this integral cannot be evaluated by 'standard' one-dimensional methods. We denote $I := \int_0^\infty e^{-ax^2} \, \mathrm{d}x$ and $M = (0,\infty) \times (0,\infty)$. We have

$$\int_M e^{-ax^2-ay^2} \, \mathrm{d}x\mathrm{d}y = \int_0^\infty e^{-ax^2} \, \mathrm{d}x \int_0^\infty e^{-ay^2} \, \mathrm{d}y = I^2.$$

We use polar coordinates, i.e.

$$\begin{aligned} x &= r\cos\alpha \\ y &= r\sin\alpha \end{aligned}$$

Note that $\Phi((0,\infty) \times (0,\pi/2)) = M$. Therefore

$$\int_M e^{-a(x^2+y^2)} \, \mathrm{d}x\mathrm{d}y = \int_0^{\pi/2} \int_0^\infty e^{-ar^2} r \, \mathrm{d}r\mathrm{d}\alpha = \frac{\pi}{2} \int_0^\infty e^{-ar^2} r \, \mathrm{d}r$$

and we use a (one-dimensional) substitution $r^2 = t$ to get

$$\frac{\pi}{2} \int_0^\infty e^{-ar^2} r \, \mathrm{d}r = \frac{\pi}{4} \int_0^\infty e^{-at} \, \mathrm{d}t = -\frac{\pi}{4} \frac{1}{a} \left[e^{-at}\right]_0^\infty = \frac{\pi}{4a}.$$

We have just deduced that

$$I = \frac{1}{2}\sqrt{\frac{\pi}{a}}.$$

4.13 Some exercises

- Write precisely the definition of a bounded function.
- Sketch contour lines for a function $f(x,y) = x^2 - y^2$ at heights $z_0 = -3, -2, -1, 0, 1, 2, 3$.
- Consider a function $f(x,y) = \begin{cases} 0, & \text{for } (x,y) = 0 \\ 1, & \text{for } y = x^2, x \neq 0, \\ 0 & \text{otherwise} \end{cases}$
- Let $f(x) = g(\sin x, \cos x)$. Express $f''(x)$ in terms of the first and second derivatives of g.
- Write the second order Taylor polynomial for a function $f(x,y) = \arctan(x + 2y)$ at the point $(1,0)$.
- Try to write a formula for $T_4(x,y)$.
- Show that the equation

  $$\log(x+y) = x + y - xy - x^2 - y^2$$

  determine on a neighborhood of $(0,1)$ a function $y(x)$. Write an equation of the tangent line to the graph of that function at the given point (recall that log stands for the natural logarithm).

- Show that the equation
$$e^x - e^y - x - y = 0$$
determine on a neigborhood of $(0,0)$ a function $y(x)$. Is this function increasing, decreasing, convex or concave at the given point? Try to approximate that function by the second order Taylor polynomial.

- Determine all local extremes of
$$f(x,y) = xy + \frac{50}{x} + \frac{20}{y}.$$

- Find maximum and minimum of a function
$$f(x,y) = x^2 + 12xy + 2y^2$$
on a set $M = \{(x,y) \in \mathbb{R}^2,\ 4x^2 + y^2 = 25\}$.

- The bottom of a rectangular box costs twice as much per unit area as the sides and top. Find the shape for a given volume that will minimize cost.

- Suppose you are running a factory, producing some sort of widget that requires steel as a raw material. Your costs are predominantly human labor, which is \$20 per hour for your worker, and the steel itself, which runs for \$170 per ton. Suppose your revenue R is loosely modeled by the following equation
$$R(h,s) = 200h^{2/3}s^{1/3}$$
where h represents hours of labor and s represents tons of steel.
If your budget is \$20 000, what is the maximum possible revenue?

- Write cross-sections M_x and M_y for a triangle whose vertices are $\langle -1,-1\rangle$, $\langle -1,3\rangle$ and $\langle 3,-1\rangle$.

- Try to compute the integral
$$\int_M 5x^2y - 2y^3 \ \mathrm{d}x\mathrm{d}y, \quad M = [2,5] \times [1,3].$$
by a different approach, in particular, try to write $\int_{\mathbb{R}} \left(\int_{M_x} f(x,y)\ \mathrm{d}y \right)\ \mathrm{d}x$ and then compute it.

- There is one easy way how to prove that the parallelogram with sides represented by vectors (a,c) and (b,d) has an area equal to $\left| \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right|$ – it can be proven by use of a triple product. Try to find it.

- Compute an area of the set M given as
$$M = \{(x,y) \in \mathbb{R}^2,\ a^2 \le x^2 + y^2 \le b^2\}$$
where a and b are given positive constants.

5 Systems of ODEs

5.1 Introduction

Problem: Two large tanks, each holding 24 liters of a brine solution, are interconnected by pipes. Fresh water flows into tank A at a rate of 6 L/min, and fluid is drained out of tank B at the same rate; also 8 L/min of fluid are pumped from tank A to tank B, and 2 L/min from tank B to tank A. The liquids inside each tank are kept well stirred so that each mixture is homogeneous. If, initially, the brine solution in tank A contains x_0 kg of salt and that in tank B initially contains y_0 kg of salt, determine the mass of salt in each tank at time $t > 0$.

Let denote:
amount of salt in the first tank: x
amount of salt in the second tank: y
salt flowing out of the first tank per one minute: $\frac{x}{24}8$
salt flowing out of the second tank per one minute: $\frac{y}{24}2 + \frac{y}{24}6$
salt flowing into the first tank per one minute: $\frac{y}{24}2$
salt flowing into the second tank per one minute: $\frac{x}{24}8$
We arrive at the system

$$\begin{aligned} x' &= -\frac{1}{3}x + \frac{1}{12}y \\ y' &= \frac{1}{3}x - \frac{1}{3}y \end{aligned}$$

which can be rewritten as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & \frac{1}{12} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

This is in particular a system of first-order linear equations.

In what follows, we will tackle a system of ODEs of the form

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t)$$

where $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^T$ and $\mathbf{b}(t) = (b_1(t), \ldots, b_n(t))^T$ are $n-$dimensional vectors and A is an n by n square matrix.
We emphasize that higher order linear differential equations with constant coefficients might be rewritten into a system of first order linear equations. Indeed, consider

$$y'' + ky' + my = 0.$$

We denote $x = y'$ and then it holds that $x' = -kx - my$ and the above system might be rewritten as

$$\begin{aligned} x' &= -kx - my \\ y' &= x \end{aligned}$$

Theorem 5.1. *Assume A is a constant n by n matrix and let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be n linearly independent solutions to the homogeneous system*

$$\mathbf{x}'(t) = A\mathbf{x}(t) \tag{5}$$

on the interval I. Then every solution to (5) *on I can be expressed in the form*

$$\mathbf{x}(t) = c_1\mathbf{x_1}(t) + \ldots + c_n\mathbf{x}_n(t),$$

where $c_1, \ldots, c_n$ are real constants.

Definition 5.1. *A set of solutions $\{\mathbf{x_1}, \ldots, \mathbf{x}_n\}$ that are linearly independent is called a* fundamental solution set *for* (5)*.*

Theorem 5.2. *If $\mathbf{x}_p$ is a particular solution to the nonhomogeneous system*

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{b}(t) \tag{6}$$

on the interval I and $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ is a fundamental solution set on I for the corresponding homogeneous system $\mathbf{x}'(t) = A\mathbf{x}(t)$, then every solution to (6) *on I can be expressed in the form*

$$\mathbf{x}(t) = \mathbf{x}_p(t) + c_1\mathbf{x}_1(t) + \ldots + c_n\mathbf{x}_n(t),$$

where $c_1, \ldots, c_n$ are real constants.

Proof is left as an exercise for interested readers.

The above theorem yields an approach to solving linear systems of the form $\mathbf{x}' = A\mathbf{x} + b$:

1. Find a fundamental solution set for the corresponding homogeneous system $\mathbf{x}' = A\mathbf{x}$.
2. Find one particular solution to the non-homogeneous system.

5.2 Homogeneous systems with constant coefficients

We are going to solve

$$\mathbf{x}' = A\mathbf{x} \tag{7}$$

Let assume (and that is something usual in the case of linear system with constant coefficients) that the solution is of the form

$$\mathbf{x}(t) = e^{\lambda t}\mathbf{v}$$

where $\lambda \in \mathbb{R}$ and $\mathbf{v}$ is an $n-$dimensional vector constant in t. We have

$$\mathbf{x}'(t) = \lambda e^{\lambda t}\mathbf{v}$$

and once we plug this into (7), we deduce

$$\lambda e^{\lambda t}\mathbf{v} = Ae^{\lambda t}\mathbf{v}.$$

We may divide by $e^{\lambda t}$ to deduce

$$\lambda\mathbf{v} - A\mathbf{v} = 0.$$

As a result, λ is an eigenvalue and $\mathbf{v}$ is a corresponding eigenvector.

Example
Let try to solve the initial value problem given at the beginning of this lesson, i.e.,

$$\mathbf{x}' = \begin{pmatrix} -\frac{1}{3} & \frac{1}{12} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}\mathbf{x},\ \mathbf{x}(0) = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

To compute the eigenvalues of $A = \begin{pmatrix} -\frac{1}{3} & \frac{1}{12} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$ we have to compute a determinant

$$\det(A - \lambda I) = \det \begin{pmatrix} -\frac{1}{3} - \lambda & \frac{1}{12} \\ \frac{1}{3} & -\frac{1}{3} - \lambda \end{pmatrix} = \lambda^2 + \frac{2}{3}\lambda + \frac{1}{12}.$$

Therefore, the eigenvalues are solutions to

$$\lambda^2 + \frac{2}{3}\lambda + \frac{1}{12} = 0.$$

We get $\lambda_1 = -\frac{1}{2}$ and $\lambda_2 = -\frac{1}{6}$.

Let take λ_1. Then the corresponding eigenvector whould satisfy $\begin{pmatrix} \frac{1}{6} & \frac{1}{12} \\ \frac{1}{3} & \frac{1}{6} \end{pmatrix} \mathbf{v_1} = 0$ and this can be solved by GEM as follows

$$\begin{pmatrix} \frac{1}{6} & \frac{1}{12} \\ \frac{1}{3} & \frac{1}{6} \end{pmatrix} \sim \begin{pmatrix} \frac{1}{6} & \frac{1}{12} \end{pmatrix}.$$

The solution (one of many) is $\mathbf{v_1} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

Similarly, for λ_2 we have

$$\begin{pmatrix} -\frac{1}{6} & \frac{1}{12} \\ \frac{1}{3} & -\frac{1}{6} \end{pmatrix} \sim \begin{pmatrix} -\frac{1}{6} & \frac{1}{12} \end{pmatrix}$$

and the second eigenvector is $\mathbf{v_2} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

The set of all solution (the general solution) is

$$\mathbf{x}(t) = c_1 e^{-1/2t} \begin{pmatrix} -1 \\ 2 \end{pmatrix} + c_2 e^{-1/6t} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

In order to reach the initial condition we deduce that

$$\mathbf{x}(0) = c_1 \begin{pmatrix} -1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

and the constants c_1 and c_2 has to be determined from the equation

$$\begin{aligned} -c_1 + c_2 &= x_0 \\ 2c_1 + 2c_2 &= y_0. \end{aligned}$$

What if the eigenvalues are not real? And what if the eigenvalues are not distinct? (the characteristic polynomial has a double (triple, etc.) root?

5.2.1 Complex eigenvalues

Example Find a general solution of

$$\mathbf{x}' = \begin{pmatrix} -1 & 2 \\ -1 & -3 \end{pmatrix} \mathbf{x}.$$

To find the eigenvalues we have to solve

$$0 = \det \begin{pmatrix} -1-\lambda & 2 \\ -1 & -3-\lambda \end{pmatrix} = (1+\lambda)(3+\lambda) + 2 = \lambda^2 + 4\lambda + 5.$$

Therefore,

$$\begin{aligned} \lambda^2 + 4\lambda + 4 &= -1 \\ (\lambda+2)^2 &= -1 \\ \lambda + 2 &= \pm i. \end{aligned}$$

We get $\lambda_1 = -2 + i$, $\lambda_2 = -2 - i$. (Similarly, we may deduce that $\lambda_{1,2} = \frac{-4\pm\sqrt{-4}}{2}$). Consider λ_1. We have

$$\begin{pmatrix} 1-i & 2 \\ -1 & -1-i \end{pmatrix} \sim \begin{pmatrix} 1-i & 2 \end{pmatrix}$$

and the corresponding eigenvector is $\mathbf{v}_1 = (2, i-1)$. Here we note that $\lambda_2 = \overline{\lambda}_1$ and $\mathbf{v}_2 = \overline{\mathbf{v_1}}$ where $\overline{(\alpha + \beta i)} = \alpha - \beta i$.
We obtain that one solution is of the form

$$\mathbf{x}(t) = e^{(-2+i)t}(2, i-1) = e^{(-2+i)t}\left((2,-1) + i(0,1)\right).$$

Recall that

$$e^{a+bi} = e^a(\cos b + i \sin b).$$

Therefore, we can write

$$\begin{aligned} \mathbf{x}(t) &= e^{-2t}\left(\cos t + i \sin t\right)\left((2,-1) + i(0,1)\right) \\ &= e^{-2t}\left(\cos t(2,-1) - \sin t(0,1)\right) + ie^{-2t}\left(\sin t(2,-1) + \cos t(0,1)\right). \end{aligned}$$

The real part represents one solution, the imaginary part the second one. Thus, the general solution has a form

$$\mathbf{x}(t) = c_1 e^{-2t}\left(\cos t(2,-1) - \sin t(0,1)\right) + c_2 e^{-2t}\left(\sin t(2,-1) + \cos t(0,1)\right)$$

where c_1 and c_2 are arbitrary real constants. The same considerations lead to the following theorem.

Theorem 5.3. *If the real matrix A has complex eigenvalues $\alpha \pm \beta i$ with corresponding eigenvectors $\mathbf{a} + i\mathbf{b}$, then the two linearly independent real vector solutions to $\mathbf{x}' = A\mathbf{x}$ are*

$$\begin{aligned} &e^{\alpha t}\cos\beta t\mathbf{a} - e^{\alpha t}\sin\beta t\mathbf{b} \\ &e^{\alpha t}\sin\beta t\mathbf{a} + e^{\alpha t}\cos\beta t\mathbf{b}. \end{aligned}$$

5.2.2 Double roots

Here we distinquish two cases: either there are two linearly independent eigenvectors corresponding to one eigenvalue, or there is just one.
Example Solve

$$\mathbf{x}' = A\mathbf{x}$$

where $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. The characteristic equation is

$$0 = \det(A - \lambda I) = \begin{pmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

There is just one eigenvalue $\lambda = 1$ and the corresponding eigenvectors solves

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0.$$

We deduce that there are two corresponding eigenvectors $\mathbf{v}_1 = (1, 0)$ and $\mathbf{v}_2 = (0, 1)$ (actually, all linear combinations of these two are eigenvectors as well). Thus the generalized solution is of the form

$$\mathbf{x}(t) = c_1 e^t (1, 0) + c_2 e^t (0, 1)$$

for some $c_1, c_2 \in \mathbb{R}$.

Example
Consider now

$$\mathbf{x}' = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix} \mathbf{x}.$$

The characteristic equation is

$$0 = \det \begin{pmatrix} 1-\lambda & 0 & 0 \\ 1 & 3-\lambda & 0 \\ 0 & 1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2(3-\lambda)$$

and the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 1$. Take $\lambda_1 = 3$. Then

$$\begin{pmatrix} -2 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \end{pmatrix}$$

and the corresponding eigenvector is $(0, 2, 1)$. Thus, the fundamental solution set contains a function

$$\mathbf{x}(t) = e^{3t}(0, 2, 1).$$

Take $\lambda_2 = 1$. Then

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and the corresponding eigenvector is $\mathbf{v}_1 = (0, 0, 1)$. Thus, the fundamental solution set contains a function

$$\mathbf{x}(t) = e^t(0, 0, 1).$$

But we need one additional function in the fundamental solution set. How to get it?

Definition 5.2. *A* generalized eigenvector $\mathbf{w}$ *corresponding to an eigenvalue* λ *is a vector satisfying*

$$(A - \lambda I)\mathbf{w} = \mathbf{v}$$

where $\mathbf{v}$ *is an eigenvector corresponding to* λ.

Matrix exponential
The exponential function $e^x : \mathbb{R} \mapsto \mathbb{R}$ can be defined as an infinite sum

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Similarly, let A be a square matrix. Then we write

$$e^{At} = I + At + \frac{A^2t^2}{2} + \frac{A^3t^3}{6} + \ldots$$

It holds that

$$\left(e^{At}\right)' = 0 + A + A^2t + \frac{A^3t^2}{2} + \ldots = A\left(I + At + \frac{A^2t^2}{2} + \ldots\right) = Ae^{At}$$

and therefore the columns of the matrix e^{At} form the fundamental solution set of

$$\mathbf{x}(t) = A\mathbf{x}.$$

This also means that every solution is of the form $\mathbf{x}(t) = e^{At}\mathbf{v}$ where $\mathbf{v}$ is an arbitrary $n-$dimensional vector. Let $\mathbf{v}$ be an eigenvector. Then

$$\begin{aligned} e^{At}\mathbf{v} &= e^{\lambda t}e^{(A-\lambda t)}\mathbf{v} \\ &= e^{\lambda t}\left(I\mathbf{v} + t(A-\lambda I)\mathbf{v} + \frac{t^2}{2}(A-\lambda I)\mathbf{v} + \ldots\right) = e^{\lambda t}\mathbf{v} \end{aligned}$$

Let $\mathbf{w}$ is a generalized eigenvector. Then

$$(A-\lambda I)^2\mathbf{w} = (A-\lambda I)(A-\lambda I)\mathbf{w} = (A-\lambda I)\mathbf{v} = 0$$

and we deduce that

$$\begin{aligned} e^{At}\mathbf{w} &= e^{\lambda t}e^{(A-\lambda t)}\mathbf{w} \\ &= e^{\lambda t}\left(I\mathbf{w} + t(A-\lambda I)\mathbf{w} + \frac{t^2}{2}(A-\lambda I)\mathbf{w} + \ldots\right) = e^{\lambda t}(\mathbf{w} + t\mathbf{v}). \end{aligned}$$

Back to our example: we have

$$\mathbf{x}' = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix}\mathbf{x}.$$

We have already deduced that $\lambda_1 = 3$ has corresponding eigenvector $(0, 2, 1)$ and the double root $\lambda_2 = 1$ has a corresponding eigenvector $\mathbf{v} = (0, 0, 1)$. Now, we have to find a corresponding generalized eigenvector $\mathbf{w}$ which satisfies

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{pmatrix}\mathbf{w} = \mathbf{v}$$

and we use the Gauss elimination method to deduce

$$\begin{pmatrix} 0 & 0 & 0 & | & 0 \\ 1 & 2 & 0 & | & 0 \\ 0 & 1 & 0 & | & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 & | & 0 \\ 0 & 1 & 0 & | & 1 \end{pmatrix}$$

Solutions are of the form $(-2,1,0)+r(0,0,1)$ for any $r \in \mathbb{R}$. It is enough to choose one solution, say $(-2,1,0)$. According to our considerations, we deduce that one solution is of the form

$$\mathbf{x}(t) = e^t((-2,1,0)+t(0,0,1)).$$

Thus, the generalized solution for the given problem is

$$\mathbf{x}(t) = c_1 e^{3t}(0,2,1) + c_2 e^t(0,0,1) + c_3 e^t((-2,1,0)+t(0,0,1))$$

for some $c_1, c_2, c_3 \in \mathbb{R}$.

To summarize:

Observation 5.1. *Let the real matrix A has an eigenvalue $\lambda \in \mathbb{R}$ which is a double root of the characteristic equation. Let there be just one corresponding eigenvector $\mathbf{v}$ and let $\mathbf{w}$ be a generalized eigenvector. Then the fundamental solution set contains the functions*

$$e^{\lambda t}\mathbf{v}, \quad e^{\lambda t}\left(\mathbf{w}+t\mathbf{v}\right).$$

5.3 Non-zero right hand side

This time we tackle the problem

$$\mathbf{x}'(t) = A\mathbf{x}(t) + \mathbf{f}(t)$$

where $\mathbf{f}(t)$ is a nonzero vector-valued function. We already know how to find all solution to the corresponding homogeneous system

$$\mathbf{x}'(t) = A\mathbf{x}(t).$$

5.3.1 Undetermined coefficients

Example Let solve

$$\mathbf{x}'(t) = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \mathbf{x}(t) + t \begin{pmatrix} -9 \\ 0 \\ -18 \end{pmatrix}.$$

First, let find all solutions to the corresponding homogeneous system. The characteristic equation is

$$0 = \det \begin{pmatrix} 1-\lambda & -2 & 2 \\ -2 & 1-\lambda & 2 \\ 2 & 2 & 1-\lambda \end{pmatrix} = (\lambda-3)^2(\lambda+3).$$

For $\lambda_1 = 3$ we have

$$\begin{pmatrix} -2 & -2 & 2 \\ -2 & -2 & 2 \\ 2 & 2 & -2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$$

and the corresponding eigenvectors are $\mathbf{v}_1 = (1,0,1)$ and $\mathbf{v}_2 = (0,1,1)$. For $\lambda_2 = -3$ we have

$$\begin{pmatrix} 4 & -2 & 2 \\ -2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & 1 \\ -2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} \sim \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & 3 \end{pmatrix}$$

and the corresponding eigenvector is $\mathbf{v}_3 = (1,1,-1)$. All solutions to the homogeneous problem have form

$$\mathbf{x}(t) = e^{3t}(c_1(1,0,1) + c_2(0,1,1)) + e^{-3t}c_3(1,1,-1)$$

where $c_1, c_2, c_3 \in \mathbb{R}$.
Let find one particular solution to

$$\mathbf{x}'(t) = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \mathbf{x}(t) + t \begin{pmatrix} -9 \\ 0 \\ -18 \end{pmatrix}$$

We can assume that the solution is of the form $\mathbf{x}(t) = \mathbf{a}t + \mathbf{b}$ where $\mathbf{a}$ and $\mathbf{b}$ are vectors constant in time. Thus we have $\mathbf{x}'(t) = \mathbf{a}$ and we plug this into the given equation in order to deduce

$$\mathbf{a} = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} (\mathbf{a}t + \mathbf{b}) + t \begin{pmatrix} -9 \\ 0 \\ -18 \end{pmatrix}$$

We compare coefficients in order to deduce

$$0 = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \mathbf{a} + \begin{pmatrix} -9 \\ 0 \\ -18 \end{pmatrix}$$

and

$$\mathbf{a} = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \mathbf{b}.$$

To find $\mathbf{a}$ we use the GEM as follows

$$\begin{pmatrix} 1 & -2 & 2 & | & 9 \\ -2 & 1 & 2 & | & 0 \\ 2 & 2 & 1 & | & 18 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 2 & | & 9 \\ 0 & -3 & 6 & | & 18 \\ 0 & 6 & -3 & | & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 2 & | & 9 \\ 0 & -3 & 6 & | & 18 \\ 0 & 0 & 9 & | & 36 \end{pmatrix}$$

and we deduce that $\mathbf{a} = (5, 2, 4)$.
Next, we have

$$\begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \mathbf{b}$$

and we once again use the GEM to get

$$\begin{pmatrix} 1 & -2 & 2 & | & 5 \\ -2 & 1 & 2 & | & 2 \\ 2 & 2 & 1 & | & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 2 & | & 5 \\ 0 & -3 & 6 & | & 12 \\ 0 & 6 & -3 & | & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 & 2 & | & 5 \\ 0 & -3 & 6 & | & 12 \\ 0 & 0 & 9 & | & 18 \end{pmatrix}$$

and we have $\mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$. Thus, all solutions to the given equation are of the form

$$\mathbf{x}(t) = \begin{pmatrix} 5 \\ 2 \\ 4 \end{pmatrix} t + \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + e^{3t} \left(c_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right) + c_3 e^{-3t} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Observation 5.2. *Let $\mathbf{f}(t) = e^{rt} t^m \mathbf{g}$ where g is a constant vector. Then one solution is of the form*

$$\mathbf{x}_p(t) = e^{rt} \left(t^{m+s} \mathbf{a}_{m+s} + t^{m+s-1} \mathbf{a}_{m+s-1} + \ldots + t\mathbf{a}_1 + \mathbf{a}_0 \right)$$

where $\mathbf{a}_i$ are constant vectors and s is an appropriately chosen integer.

5.3.2 Variation of parameters

Let solve

$$\mathbf{x}' = A\mathbf{x} + \mathbf{f}(t).$$

Let $X(t)$ be a matrix whose columns are the elements of the fundamental system set. Then we know that all solutions to the appropriate homogeneous system are of the form

$$\mathbf{x}(t) = X(t)\mathbf{w}$$

where $\mathbf{w}$ is a constant vector. We are looking for a particular solution $\mathbf{x}_p$ of the form

$$\mathbf{x}_p(t) = X(t)\mathbf{w}(t).$$

Then

$$\mathbf{x}_p'(t) = X'(t)\mathbf{w}(t) + X(t)\mathbf{w}'(t)$$

and since $X'(t) = AX(t)$ we get

$$X(t)\mathbf{w}'(t) = \mathbf{f}(t).$$

Thus we have just deduced that

$$\mathbf{w}(t) = \int X^{-1}(t)\mathbf{f}(t)\ \mathrm{d}t.$$

Example
Let solve

$$\mathbf{x}'(t) = \begin{pmatrix} 2 & -3 \\ 1 & -2 \end{pmatrix}\mathbf{x}(t) + \begin{pmatrix} e^{2t} \\ 1 \end{pmatrix},\ \mathbf{x}(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

To solve the appropriate homogeneous system we have to find solutions to

$$0 = \det\begin{pmatrix} 2-\lambda & -3 \\ 1 & -2-\lambda \end{pmatrix} = \lambda^2 - 1$$

thus $\lambda_1 = 1$ and $\lambda_2 = -1$.
Take $\lambda_1 = 1$. We have

$$\begin{pmatrix} 1 & -3 \\ 1 & -3 \end{pmatrix} \sim \begin{pmatrix} 1 & -3 \end{pmatrix}$$

and $\mathbf{v}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.
Take $\lambda_2 = -1$. We have

$$\begin{pmatrix} 3 & -3 \\ 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & -1 \end{pmatrix}$$

and $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
The matrix whose columns are solutions is

$$X(t) = \begin{pmatrix} 3e^t & e^{-t} \\ e^t & e^{-t} \end{pmatrix}.$$

We infer that

$$X^{-1}(t) = \begin{pmatrix} \frac{1}{2}e^{-t} & -\frac{1}{2}e^{-t} \\ -\frac{1}{2}e^{t} & \frac{3}{2}e^{t} \end{pmatrix}.$$

Thus the desired vector $\mathbf{w}$ should satisfy

$$\mathbf{w}(t) = \int \begin{pmatrix} \frac{1}{2}e^{-t} & -\frac{1}{2}e^{-t} \\ -\frac{1}{2}e^{t} & \frac{3}{2}e^{t} \end{pmatrix} \begin{pmatrix} e^{2t} \\ 1 \end{pmatrix} \, \mathrm{d}t = \int \begin{pmatrix} \frac{1}{2}(e^{t} - e^{-t}) \\ -\frac{1}{2}e^{3t} + \frac{3}{2}e^{t} \end{pmatrix} \, \mathrm{d}t$$

and we deduce that

$$\mathbf{w}(t) = \begin{pmatrix} \frac{1}{2}(e^{t} + e^{-t}) \\ -\frac{1}{6}e^{3t} + \frac{3}{2}e^{t} \end{pmatrix}$$

Thus, all solutions are of the form

$$\mathbf{x}(t) = \begin{pmatrix} 3e^{t} & e^{-t} \\ e^{t} & e^{-t} \end{pmatrix} \begin{pmatrix} \frac{1}{2}(e^{t} + e^{-t}) \\ -\frac{1}{6}e^{3t} + \frac{3}{2}e^{t} \end{pmatrix} + \begin{pmatrix} 3e^{t} & e^{-t} \\ e^{t} & e^{-t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

where $c_1, c_2 \in \mathbb{R}$ are arbitrary.
Next, we have to employ the initial condition $\mathbf{x}(0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$. This yields

$$\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{7}{6} \end{pmatrix} + \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

We deduce that

$$\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} - \begin{pmatrix} \frac{25}{6} \\ \frac{13}{6} \end{pmatrix} = \begin{pmatrix} \frac{-31}{6} \\ -\frac{15}{6} \end{pmatrix}.$$

We deduce that $c_1 = -\frac{4}{3}$ and $c_2 = -\frac{7}{6}$. Finally,

$$\mathbf{x}(t) = \begin{pmatrix} 3e^{t} & e^{-t} \\ e^{t} & e^{-t} \end{pmatrix} \begin{pmatrix} \frac{1}{2}(e^{t} + e^{-t}) \\ -\frac{1}{6}e^{3t} + \frac{3}{2}e^{t} \end{pmatrix} + \begin{pmatrix} 3e^{t} & e^{-t} \\ e^{t} & e^{-t} \end{pmatrix} \begin{pmatrix} -\frac{4}{3} \\ -\frac{7}{6} \end{pmatrix}$$

is the demanded solution.

5.4 Systems in a plane

During this subsection we consider systems of the form

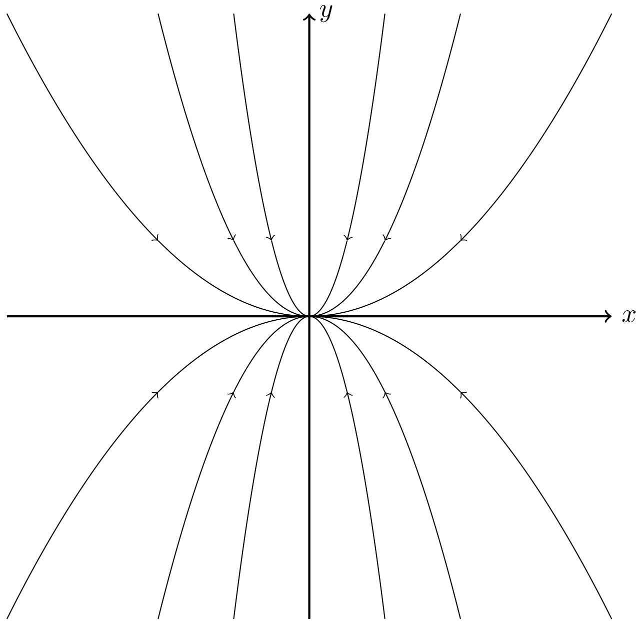
$$\begin{aligned} \frac{\partial x}{\partial t} &= f(x, y) \\ \frac{\partial y}{\partial t} &= g(x, y) \end{aligned}$$

Note that this time, the system is not necessarily linear, however, it is autonomous – the right hand side in $t-$independent.

Definition 5.3. *If $x(t)$ and $y(t)$ is a solution pair to the above mentioned system for t in the interval I, then a plot in the xy-plane of the parametrized curve $(x(t), y(t))$ for t in I, together with arrows indicating its direction with increasing t, is said to be a* trajectory *of the system. In such a context we call the xy-plane the* phase plane.

Example no. 1:

$$\begin{aligned} x' &= -x \\ y' &= -2y \end{aligned}$$



Recall that

$$\frac{\partial y}{\partial x} = \frac{\frac{\partial y}{\partial t}}{\frac{\partial x}{\partial t}} = \frac{-2y}{-x}.$$

This yields $y = cx^2$ and thus we get the picture as above. Note that, since $y'(t)$ and $x'(t)$ are negative for $x, y > 0$, we get trajectories aiming to the origin. See picture above.

Example no. 2:

$$x' = x$$
$$y' = 2y$$

This time, the picture is the same as above with only one exception – the arrows aim away of origin (see the picture below), try to justify why.

Definition 5.4. *: A point $(x_0, y_0) \in \mathbb{R}^2$ where $f(x_0, y_0) = g(x_0, y_0) = 0$ is called a* critical point *(or* equilibrium point*) of the given system. The corresponding solution $x \equiv x_0$, $y \equiv y_0$ is called an* equilibrium solution *(or* stationary solution*).*

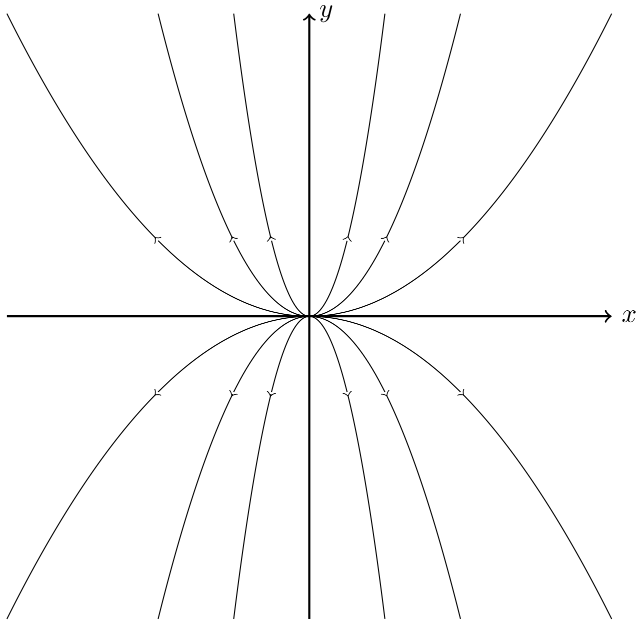
Observation 5.3. *Let $x(t)$ and $y(t)$ be a solution on $[0, \infty)$ to the given system (we assume f and g are continuous). If the limits*

$$\lim_{t\to\infty} x(t) = x_0, \quad \lim_{t\to\infty} y(t) = y_0$$

exist and are finite, then (x_0, y_0) is a critical point of the system.

Types of equilibrium points:

- Stable node (asymptotically stable)



- Unstable node
- Stable spiral (asymptotically stable)
- Unstable spiral
- Saddle (unstable)
- Center (stable, but not asymptiotically)

HERE SHOULD BE A PICTURE, TBD.
Example: Find the critical points and sketch trajectories in the phase plane for

$$\begin{aligned} x' &= -y(y-2) \\ y' &= (x-2)(y-2). \end{aligned}$$

What is the behavior of the solutions starting from $(3,0)$, $(5,0)$ and $(2,3)$?

Let consider a special case of a linear system in a plane, i.e.,

$$\begin{aligned} x' &= a_{11}x + a_{12}y + b_1 \\ y' &= a_{21}x + a_{22}y + b_2 \end{aligned}$$

which might be shortened to

$$\mathbf{x}' = A\mathbf{x} + \mathbf{b}$$

where $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ and

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

In what follows, we assume that $\mathbf{b} = 0$. This assumption will be commented later.

From what we know, we deduce that

- Let the eigenvalues λ_1 and λ_2 be real, distinct and both positive. Then $(0, 0)$ is an unstable node.
- Let the eigenvalues λ_1 and λ_2 be real, distinct and both negative. Then $(0, 0)$ is a stable node.
- Let the eigenvalues λ_1 and λ_2 be real and have oposite signs. Then $(0, 0)$ is a saddle point.
- Let the eigenvalues λ_1 and λ_2 be equal. Then $(0, 0)$ is either a proper node (stable or unstable) or an improper node (stable or unstable).
- Let the eigenvalues be comples, i.e. $\lambda_{12} = a \pm bi$ where $a, b \in \mathbb{R}$. Take the special case

$$\begin{aligned} x' &= ax - by \\ y' &= bx + ay. \end{aligned}$$

Take $z = x + iy$. Then we have

$$z' = x' + iy' = (ax - by) + i(bx + ay) = a(x + iy) + ib(x + iy) = (a + bi)z.$$

We use the polar coordinates $z(t) = r(t)e^{i\theta(t)}$ we arrive at

$$z'(t) = r'(t)e^{i\theta t} + ir(t)e^{i\theta(t)}\theta'(t) = (a + bi)r(t)e^{i\theta(t)}$$

and we infer

$$r'(t) + ir(t)\theta'(t) = ar(t) + bir(t).$$

Therefore (recall that r and θ are real functions),

$$r'(t) = ar(t), \qquad \theta'(t) = b.$$

What if $a = 0$? (pure imaginary roots). TBD.

Example Find and classify the critical point of the linear system

$$\begin{aligned} x' &= 2x + y - 3 \\ y' &= -3x - 2y - 4 \end{aligned}$$

SOLUTION SHOULD BE ADDED.

5.4.1 Almost linear systems

An *almost linear system* is a system of the form

$$\begin{aligned} x' &= a_{11}x + a_{12}y + f(x, y) \\ y' &= a_{21}x + a_{22}y + g(x, y) \end{aligned}$$

Here we assume that f and g are just small perturbations. In particular,

$$\lim_{\sqrt{x^2+y^2}\to 0} \frac{f(x, y)}{\sqrt{x^2 + y^2}} = 0, \qquad \lim_{\sqrt{x^2+y^2}\to 0} \frac{g(x, y)}{\sqrt{x^2 + y^2}} = 0.$$

The system

$$\mathbf{x}' = A\mathbf{x}, \qquad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is called a *corresponding linear system.*

Note, that the origin is an equilibrium point of the almost linear system.

Theorem 5.4. *The stability properties of the critical point at the origin for the almost linear system are the same as the stability properties of the origin for the corresponding linear system with one exception: When the eigenvalues are pure imaginary, the stability properties for the almost linear system cannot be deduced from the corresponding linear system.*

Competing species:
The population of two species x and y (independent on each other) might be governed by the logistic equations

$$\begin{aligned} x' &= k_1 x(C_1 - x) = k_1 C_1 x - k_1 x^2 = a_1 x - b_1 x^2 \\ y' &= k_2 x(C_2 - y) = a_2 y - b_2 y^2 \end{aligned}$$

Now assume that both species compete for the same food. In such a case, the capacity C might be exhausted by both x and y and therefore we assume

$$\begin{aligned} x' &= a_1 x - b_1 x^2 - c_1 xy \\ y' &= a_2 y - b_2 y^2 - c_2 xy \end{aligned}$$

Consider two competing species whose population is governed by

$$\begin{aligned} x' &= x(7 - x - 2y) = 7x - x^2 - 2xy \\ y' &= y(5 - y - x) = 5y - y^2 - xy. \end{aligned}$$

SOLUTION SHOULD BE ADDED

Now, consider two species x and y where x is a prey and y is a predator. Then, we get

$$\begin{aligned} x' &= ax - bxy \\ y' &= -cy + dxy. \end{aligned} \tag{8}$$

where a, b, c, d are positive constants. SOLUTION SHOULD BE ADDED

5.5 Some exercises

- Find a solution to

  $$\mathbf{x}'(t) = \begin{pmatrix} -3 & -1 \\ 2 & -1 \end{pmatrix} \mathbf{x}(t)$$

  that satisfies an initial condition $\mathbf{x}(0) = (-1, 0)$.

- Find a general solution to

  $$x'' + 4x' + 4 = 0$$

  using the theory of systems of linear differential equations of the first order.

- Find a solution to

  $$\mathbf{x}' = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \mathbf{x}$$

  fulfilling $\mathbf{x}(0) = (1, 1)$.

- Find a general solution to

$$\mathbf{x}'(t) = \begin{pmatrix} 6 & 1 \\ 4 & 3 \end{pmatrix} \mathbf{x}(t) + \begin{pmatrix} -11 \\ -5 \end{pmatrix}.$$

- Find a general solution to

$$\mathbf{x}'(t) = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{x}(t) + \begin{pmatrix} e^{2t} \\ \sin t \\ t \end{pmatrix}.$$

- Solve the related phase plane differential equation for

$$\begin{aligned} x' &= \frac{3}{y} \\ y' &= \frac{2}{x}. \end{aligned}$$

  Then sketch several representative trajectories.

- Find and classify the critical point of

$$\begin{aligned} x' &= -x + y + 4 \\ y' &= -y + 1. \end{aligned}$$

- Find all the critical points of

$$\begin{aligned} x' &= x + y \\ y' &= 5y - xy + 6 \end{aligned}$$

  and discuss their type and stability properties.

6 Contour integrals

6.1 Contour lines

Definition 6.1. *A* smooth curve *(or a contour line) in $\mathbb{R}^2$ is a set of points $\mathcal{K} \subset \mathbb{R}^2$ whose coordinates are given as*

$$x = x(t), \ y = y(t)$$

where

1. *$x(t)$ and $y(t)$ are continuous on some interval $I \subset \mathbb{R}$ (usual choice is $I = [a, b]$),*
2. *$x(t)$ and $y(t)$ are of class C^1,*
3. *$(x'(t), y'(t)) \neq 0$ for every $t \in [a, b]$.*

Remark 6.1. *- A smooth curve in $\mathbb{R}^3$ is defined analogously.*
- If the second and third conditions are fulfilled with exception of finitely many points, the curve is called piecewisely smooth.
- A function $r : I \mapsto \mathbb{R}^2$, $r(t) = (x(t), y(t))$ is called a parametrization of the curve.
- A curve is called simple if it does not cross itself. I.e., $r(t)$ is one-to-one.
- A curve is called closed if $r : [a, b] \to \mathbb{R}^2$ fulfills $r(a) = r(b)$.

Several examples:
- The curve $r : [0, 2\pi] \to \mathbb{R}^2$, $r(t) = (R\cos t, R\sin t)$ is a simple closed curve. (a circle with radius R)
- The curve $r : \mathbb{R} \to \mathbb{R}^3$, $r(t) = (R\cos t, R\sin t, bt)$ is a simple curve but it is not closed. (a helix)

Definition 6.2. *Let $r(t) : [a, b] \to \mathbb{R}^2$ (resp. $r(t) : [a, b] \to \mathbb{R}^3$) be a parametrization of a smooth curve $\mathcal{K}$ and let $P_0 \in \mathcal{K}$ and t_0 be such that $P_0 = r(t_0)$. Then the vector $r'(t)$ is a* tangent vector *of the curve $\mathcal{K}$ at the point P_0.*

Example Consider a helix $r : \mathbb{R} \mapsto \mathbb{R}^3$, $r(t) = (2\cos t, 2\sin t, 3t)$. Determine a tangent vector at $P_0 = (-2, 0, 3\pi)$.
First, we have to determine $t_0 \in \mathbb{R}$ such that $r(t_0) = P_0$. We infer $t_0 = \pi$. Further, we have

$$r'(t) = (-2\sin t, 2\cos t, 3)$$

and therefore

$$r'(\pi) = (0, -2, 3)$$

which is the desired tangent vector in $(-2, 0, 3\pi)$.

Definition 6.3. *Let $\mathcal{K}$ be a curve given by a parametrization $r : I_1 \to \mathcal{K}$. Next, let $\varphi : I_2 \to I_1$ be an arbitrary C^1 function. Then $r \circ \varphi : I_2 \to \mathcal{K}$ is also a parametrization of K assuming $\varphi' \neq 0$ everywhere on I_2. The function φ is called an* admissible change of parametrization.

Definition 6.4. *Let $r : [a, b] \to \mathbb{R}^d$, $d = 2, 3$ be a smooth parametrization of some curve. Then we say that $r(a)$ is the* initial point *and $r(b)$ is the* terminal point *of the curve.*

Observation 6.1. *Let $\varphi \in C^1$ and let $\varphi' \neq 0$ everywhere on an interval $I \subset \mathbb{R}$. Then either $\varphi'(t) > 0$ for every $t \in I$ or $\varphi'(t) < 0$ for every t in I.*
Let φ be a change of parametrization with $\varphi' > 0$. Then the initial point and the terminal point of the curve do not change.
Let φ be a change of parametrization with $\varphi' < 0$. Then the initial point and the terminal point of the curve interchange.

Example
Consider a curve given by parametrization $r : [0, 1] \to \mathbb{R}^3$

$$r(t) = (1 - 2t, 2 + 3t, 3 - 2t).$$

Consider a change $t = \varphi(s) = s - 2$. Then $\varphi : [2, 3] \to [0, 1]$ is one-to-one, $\varphi' = 1 \neq 0$ and we have a new parametrization

$$r_1(t) = (1 - 2(s - 2), 2 + 3(s - 2), 3 - 2(s - 2)) = (5 - 2s, -4 + 3s, 7 - 2s),\ s \in [2, 3].$$

Note that the initial point is $(1, 2, 3)$ and the terminal point is $(-1, 5, 1)$ for both parametrizations.

Observation 6.2. *Let $\mathcal{K}$ be a curve given by a parametrization $r_1(t)$. And let $r_2(s)$ be given as $r_1(\varphi(s))$ where φ is an admissible change of parametrization. Then we have*

$$r_1'(t) = r_2'(\varphi(s))\varphi'(s).$$

Corollary 6.1. *The change of parametrization does not change the direction of the tangent vector. It might change only its size or orientation.*

Definition 6.5. *Let $\mathcal{K}$ is given by a parametrization $r : I \to \mathbb{R}^d$ and let f be a real function defined on $\mathbb{R}^d$. Then the* restriction *of f to $\mathcal{K}$ $f\restriction_{\mathcal{K}} : I \to \mathbb{R}$ is a function $f \circ r$.*

Example Let $f(x, y, z) = x^2 + y^2 + z^2$ and let $\mathcal{K}$ be given by a parametrization $r(t) = (t, t^2, t^3),\ t \in [0, 1]$. Then

$$f\restriction_{\mathcal{K}} = f(t, t^2, t^3) = t^2 + t^4 + t^6,\ \mathrm{Dom}\ f\restriction_{\mathcal{K}} = [0, 1].$$

Recall the Riemann integral: for a function $f : [a, b] \to \mathbb{R}$ we define

$$(R) \int_a^b f(x)\ \mathrm{d}x = \lim_{n\to\infty} S_n(f)$$

where

$$S_n(f) = \frac{b-a}{n} \sum_{i=1}^{n} f(c_i)$$

where $c_i \in [x_{i-1}, x_i]$ and $x_i = a + \frac{i}{n}(b - a)$.

The length of a curve:

Let the curve $\mathcal{K}$ has a parametrization $r : [a, b] \mapsto \mathbb{R}^d$, $d = 2, 3$. In order to compute the length of a curve, we approximate the length by

$$\sum_{i=1}^{n} \|r(x_i) - r(x_{i-1})\|$$

where x_i are defined as above. By a mean value theorem there exists $\xi_i \in (x_{i-1}, x_i)$ such that

$$r(x_i) - r(x_{i-1}) = r'(\xi_i)(x_i - x_{i-1}) = r'(\xi_i)\frac{b-a}{n}.$$

(This is not completely true) We plug this into the sum in order to get the approximate length

$$\frac{b-a}{n} \sum_{i=1}^{n} \|r'(\xi_i)\|$$

and, due to the definition of the Riemann integral, we obtain that the length of the smooth curve l is given as

$$l = \int_a^b \|r'(t)\|\ \mathrm{d}t.$$

Definition 6.6. *A curve whose length is finite is called a* finite curve.

Definition 6.7. *For a finite curve $\mathcal{K}$ with parametrization $r : [a, b] \to \mathbb{R}^d$ we define an* integral of f on $\mathcal{K}$ *as follows*

$$\int_{\mathcal{K}} f(x, y, z)\ \mathrm{d}s = \int_a^b f(r(t))\|r'(t)\|\ \mathrm{d}t$$

Example Let compute an integral

$$\int_{\mathcal{K}} xyz\ \mathrm{d}s$$

where $\mathcal{K}$ is given as

$$x = t,\ y = \frac{1}{3}\sqrt{8t^3},\ z = \frac{1}{2}t^2,\ t \in [0,1].$$

Here we have $r(t) = (t, \frac{1}{3}\sqrt{8}t^{3/2}, \frac{1}{2}t^2)$. Therefore, $r'(t) = (1, \frac{1}{2}\sqrt{8t}, t)$ and $\|r'(t)\| = \sqrt{1+2t+t^2} = (1+t)$. We infer

$$\int_{\mathcal{K}} xyz\ \mathrm{d}s = \int_0^1 \frac{\sqrt{2}}{3}t^{9/2}(1+t)\ \mathrm{d}t = \frac{\sqrt{2}}{3}\left[\frac{2}{11}t^{11/2} + \frac{2}{13}t^{13/2}\right]_0^1 = \frac{\sqrt{2}}{3}\frac{26+22}{143} = \frac{16\sqrt{2}}{143}.$$

Example Determine the mass of a part of a helix $\mathcal{K}$ which is parametrized by $r(t) = (3\cos t, 3\sin t, 4t),\ t \in [0, 2\pi]$ and whose (linear) density is equal to $f(x,y,z) = z^2$.
Here we have $f(r(t)) = (4t)^2 = 16t^2$ and $r'(t) = (-3\sin t, 3\cos t, 4)$. Therefore $\|r'(t)\| = 5$ and we compute

$$\int_{\mathcal{K}} f(x,y,z)\ \mathrm{d}s = \int_0^{2\pi} 16t^2 5\ \mathrm{d}t = 80\left[\frac{t^3}{3}\right]_0^{2\pi} = \frac{640}{3}\pi^3.$$

6.2 Vector fields

Definition 6.8. *Let $d = 2, 3$. A* vector field *on a set $G \subset \mathbb{R}^d$ is a function*

$$F : G \to \mathbb{R}^d$$

(this means a function which assign a vector $F(X)$ to every $X \in G$.)

Example Consider a vector field

$$F(x,y) = \left(-\frac{y}{\sqrt{x^2+y^2}}, \frac{x}{\sqrt{x^2+y^2}}\right).$$

This field is defined on a set $G = \mathbb{R}^2 \setminus \{0\}$.
Example Consider a vector field

$$F(x,y,z) = \left(\frac{x}{\sqrt{x^2+y^2+z^2}}, \frac{y}{\sqrt{x^2+y^2+z^2}}, \frac{z}{\sqrt{x^2+y^2+z^2}}\right)$$

This field is defined on a set $G = \mathbb{R}^3 \setminus \{0\}$.

Similarly to the scalar case (i.e. scalar function defined on $\mathbb{R}^d$, $d = 2, 3$) we define restriction of a vector field F to a curve $\mathcal{K}$ given by a parametrization $r : [a,b] \to \mathbb{R}^d$ as a composition, i.e.

$$F{\restriction}_{\mathcal{K}} : [a,b] \to \mathbb{R}^d,\ F{\restriction}_{\mathcal{K}} = F \circ r.$$

Example Consider

$$F(x,y,z) = \left(\frac{x}{2+y}, \frac{y}{3+x}, \frac{xy}{1+z^2}\right)$$

and a curve $\mathcal{K}$ given by

$$r(t) = (\cos t, \sin t, t), t \in \mathbb{R}.$$

Then we have

$$F{\restriction}_{\mathcal{K}}(r(t)) = \left(\frac{\cos t}{2+\sin t}, \frac{\sin t}{3+\cos t}, \frac{\cos t \sin t}{1+t^2}\right).$$

6.3 Differential

Definition 6.9. *Let $r : [a, b] \to \mathbb{R}^3$ have coordinates $x(t), y(t), z(t)$. The* differential *of r at a point t_0 is a vector-valued function defined as*

$$dr = (x'(t_0)dt, y'(t_0)dt, z'(t_0)dt)$$

and we write

$$dr = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z).$$

Remark 6.2. *We may also write*

$$dr = r'(t)\mathrm{d}t.$$

Equivalently, we may define differentials in $\mathbb{R}^2$.

Definition 6.10. *Let $\mathcal{K}$ be a curve with parametrization $r : [a, b] \to \mathbb{R}^d$, $d = 2, 3$. We define an* integral of a vector field *F on $\mathcal{K}$ as*

$$\int_{\mathcal{K}} F\ dr = \int_a^b F(t) \cdot r'(t)\ \mathrm{d}t.$$

Example Compute $\int_{\mathcal{K}} F\ dr$ where $\mathcal{K}$ is given by a parametrization

$$r(t) = (2 - t^2, 1 + 2t, 2 + 3t^3),\ t \in [-1, 1]$$

and F is given as

$$F(x, y, z) = (xz, yz, xy).$$

We have $r'(t) = (-2t, 2, 9t^2)$ and therefore

$$\int_{\mathcal{K}} (xz, yz, xy)\ \mathrm{d}r = \int_{-1}^{1} ((2 - t^2)(2 + 3t^3), (1 + 2t)(2 + 3t^3), (2 - t^2)(1 + 2t)) \cdot (-2t, 2, 9t^2)\ \mathrm{d}t$$

$$= \int_{-1}^{1} 4 + 18t^2 + 46t^3 - 9t^4 - 18t^5 + 6t^6\ \mathrm{d}t = \frac{634}{35}.$$

Definition 6.11. *Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be two curves (whose corresponding parametrizations are $r_1 : [a, b] \to \mathbb{R}^d$ and $r_2 : [c, d] \to \mathbb{R}^d$) such that the terminal point of the first curve is the initial point of the second one. Then we define a (piecewisely smooth) curve $\mathcal{K}_1 \dot{+} \mathcal{K}_2$ as a set of points belonging to either $\mathcal{K}_1$ or $\mathcal{K}_2$ and whose parametrizaion (one of many) is a function $r : [a, d - c + b] \to \mathbb{R}^d$ defined as*

$$r(t) = \begin{cases} r_1(t) \text{ for } t \in [a, b] \\ r_2(t + c - b) \text{ for } t \in [b, d - c + b]. \end{cases}$$

Next, we define a curve $\dot{-}\mathcal{K}_1$ as a set of points of the curve $\mathcal{K}_1$ with parametrization $\dot{r}$ defined as

$$\dot{r}(t) = r(a + b - t),\ t \in [a, b].$$

Remark 6.3. *(Note that the terminal point of $\dot{-}\mathcal{K}_1$ is the initial point of $\mathcal{K}_1$ and vice versa.)*

Observation 6.3. *Let $\alpha \in \mathbb{R}$ and let F and G be vector fields. Then:*

1. $\int_{\mathcal{K}} (\alpha F + G)\ dr = \alpha \int_{\mathcal{K}} F\ dr + \int_{\mathcal{K}} G\ dr$,
2. $\int_{\mathcal{K}_1 \dot{+} \mathcal{K}_2} F\ dr = \int_{\mathcal{K}_1} F\ dr + \int_{\mathcal{K}_2} F\ dr$,
3. $\int_{\dot{-}\mathcal{K}} F\ dr = -\int_{\mathcal{K}} F\ dr$.

6.4 Differential forms

Definition 6.12. *Let* $F(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z))$ *be a vector field and let* $dr = (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$*. Then we may write that*

$$F \cdot dr = F_1(x, y, z)\mathrm{d}x + F_2(x, y, z)\mathrm{d}y + F_3(x, y, z)\mathrm{d}z.$$

The right hand side of this formula is called a differential form*. Therefore, there is a one-to-one relation between differential forms and vector fields.*

Remark 6.4. *We may also write that*

$$\int_{\mathcal{K}} F \cdot dr = \int_{\mathcal{K}} F_1(x, y, z)\mathrm{d}x + F_2(x, y, z)\mathrm{d}y + F_3(x, y, z)\mathrm{d}z.$$

Example Compute

$$\int_{\mathcal{K}} x \ \mathrm{d}x + xy \ \mathrm{d}y + (z - x) \ \mathrm{d}z$$

where $\mathcal{K}$ is given by

$$r(t) = (\cos t, \sin t, t), \ \ t \in [0, \pi].$$

We have $\mathrm{d}x = -\sin t \ \mathrm{d}t$, $\mathrm{d}y = \cos t \ \mathrm{d}t$ and $\mathrm{d}z = 1 \ \mathrm{d}t$ and therefore we may write

$$\begin{aligned} \int_{\mathcal{K}} x \ \mathrm{d}x + xy \ \mathrm{d}y + (z - x) \ \mathrm{d}z &= \int_0^{\pi} \cos t(-\sin t) \ \mathrm{d}t + \cos t \sin t \cos t \ \mathrm{d}t + (t - \cos t) \ \mathrm{d}t \\ &= \frac{2}{3} + \frac{\pi^2}{2}. \end{aligned}$$

6.5 Potential

Let

$$w = \log \sqrt{x^2 + y^2 + z^2}.$$

then

$$\mathrm{d}w = \frac{x}{x^2 + y^2 + z^2}\mathrm{d}x + \frac{y}{x^2 + y^2 + z^2}\mathrm{d}y + \frac{z}{x^2 + y^2 + z^2}\mathrm{d}z.$$

We say, that w is a potential of a differential form

$$F(x, y, z) = \frac{x}{x^2 + y^2 + z^2}\mathrm{d}x + \frac{y}{x^2 + y^2 + z^2}\mathrm{d}y + \frac{z}{x^2 + y^2 + z^2}\mathrm{d}z.$$

Definition 6.13. *Let* $U(x, y, z) : G \subset \mathbb{R}^3 \mapsto \mathbb{R}$ *and let* $F = \nabla U$*. Then we say that* F *has a* potential U*.*

Observation 6.4. *Let* F *has a potential* U *and let* $\mathcal{K}$ *be a curve parametrized by* $r : [a, b] \mapsto \mathbb{R}^3$*. Then*

$$\int_{\mathcal{K}} F \ dr = U(r(b)) - U(r(a)).$$

Or, equivalently, the integral of F *on* $\mathcal{K}$ *is equal to the difference of the value of the potential* U *in the terminal point of the curve and in the initial point of the curve.*

Proof. Assume that $r(t) = (x(t), y(t), z(t))$. It holds that

$$\int_{\mathcal{K}} F \, \mathrm{d}r = \int_a^b \frac{\partial}{\partial x} U(r(t)) x'(t) + \frac{\partial}{\partial y} U(r(t)) y'(t) + \frac{\partial}{\partial z} U(r(t)) z'(t) \, \mathrm{d}t$$
$$= \int_a^b \frac{\partial}{\partial t} U(r(t)) \, \mathrm{d}t = U(r(b)) - U(r(a)).$$

□

Example Verify, that $U(x, y) = -\arctan \frac{x}{y}$ is a potential of

$$F(x, y) = \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right).$$

on an upper half-plane $\{(x, y) \in \mathbb{R}^2, y > 0\}$. Then compute $\int_{\mathcal{K}} F \, dr$ where $\mathcal{K}$ is a curve with an initial point $A = (1, 1)$ and terminal point $B = (2\sqrt{3}, 2)$.
First, we have to differentiate U with respect to x and y. We have

$$\frac{\partial U}{\partial x} = -\frac{1}{1 + \frac{x^2}{y^2}} \frac{1}{y} = -\frac{y}{x^2 + y^2} = F_1(x, y)$$

and

$$\frac{\partial U}{\partial y} = -\frac{1}{1 + \frac{x^2}{y^2}} \left(-\frac{x}{y^2} \right) = \frac{x}{x^2 + y^2} = F_2(x, y).$$

We infer that F has potential U. Next,

$$\int_{\mathcal{K}} F \, dr = U(2\sqrt{3}, 2) - U(1, 1) = -\arctan \sqrt{3} + \arctan 1 = -\frac{\pi}{3} + \frac{\pi}{4} = -\frac{\pi}{12}.$$

Observation 6.5.
- *Let F have a potential. Then the value $\int_{\mathcal{K}} F \, dr$ is independent of the curve – only the initial and the terminal points matter.*
- *Let $\mathcal{K}$ be a closed curve and let F have a potential. Then $\int_{\mathcal{K}} F \, dr = 0$.*

Definition 6.14. *An open set $G \subset \mathbb{R}^2$ is called* simply connected *if for every closed curve $\mathcal{K} \subset G$ it holds that its interior belongs to G.*

Not every field F has a potential. The following hold:

Theorem 6.1. *Let $F(x, y) = (F_1(x, y), F_2(x, y))$ belong to C^1 on a simply connected domain G. Then F has potential if and only if*

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}.$$

$3d$ analogy: First, we gave a vague definition of a simply connected set. Roughly speaking, simply connected set is an open set which do not contain holes and which consists only of one connected component.

Theorem 6.2. *A vector field $F(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z)) \in C^1$ defined on a simply connected set G has potential if and only if*

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x}, \; \frac{\partial F_1}{\partial z} = \frac{\partial F_3}{\partial x}, \; \frac{\partial F_2}{\partial z} = \frac{\partial F_3}{\partial y}.$$

Example Prove that

$$F(x,y) = (x^2 - y^2, 5 - 2xy)$$

defined on the whole plane has a potential. Then find its potential.
First, we deduce that

$$\frac{\partial F_1}{\partial y} = -2y = \frac{\partial F_2}{\partial x}.$$

Next, since $F_1 = \frac{\partial U}{\partial x}$, we deduce that

$$U(x,y) = \int F_1(x,y)\ \mathrm{d}x = \int x^2 - y^2\ \mathrm{d}x = \frac{x^3}{3} - xy^2 + c(y).$$

Similarly, since $F_2 = \frac{\partial U}{\partial y}$, we deduce that

$$U(x,y) = \int F_2(x,y)\ \mathrm{d}y = \int 5 - 2xy\ \mathrm{d}y = 5y - xy^2 + c(x).$$

We deduce by comparing these two relations that

$$U(x,y) = \frac{x^3}{3} - xy^2 + 5y + c,\ \ c \in \mathbb{R}.$$

6.6 Some exercises

- Try to sketch a curve
  $$r(t) = (4 - 2t, 3 + 6t - 4t^2)$$
  and find a tangent vector at the point $P_0 = (2,5)$.
- Try to find a curve $\mathcal{K}$ with parametrization $r : [a,b] \to \mathbb{R}^3$ such that $r(b) - r(a) = (0,0,1)$ and, simultaneously, there is no point where $r'(t)$ is parallel to $(0,0,1)$. (Such curve contradicts the mean value property mentioned in the text.)
- Let $r_1 : [a,b] \to \mathbb{R}^d$ and $r_2 : [c,d] \to \mathbb{R}^d$ be two different parametrization of $\mathcal{K}$ with the same orientation (i.e., the initial and terminal points are similar). Prove that
  $$\int_a^b F(r_1(t)) \cdot r_1'(t)\ \mathrm{d}t = \int_c^d F(r_2(u)) \cdot r_2'(u)\ \mathrm{d}u.$$
- Consider a vector field
  $$F(x,y) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$$
  which is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$. Compute
  $$\int_{\mathcal{K}} F\ dr$$
  where $\mathcal{K}$ is a unit circle with center at the origin.
- Consider a vector field
  $$F(x,y) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right)$$
  which is continuous on $\mathbb{R}^2 \setminus \{(0,0)\}$. Compute
  $$\int_{\mathcal{K}} F\ dr$$
  where $\mathcal{K}$ is a unit circle with center at the origin.