Mathematics for Informatics

Recap
(lecture 12 of 12)

Francesco Dolce

francesco.dolce@fjfi.cvut.cz

Czech Technical University in Prague

Fall 2020/2021

created: December 14, 2020, 12:39

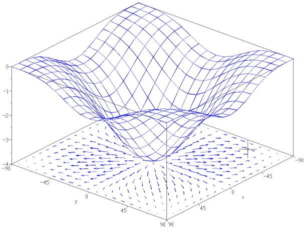
Recap

Gradient of a function

The **gradient** of a function $f(x_1, x_2, \ldots, x_n)$ at the (n-dimensional) point $b \in \mathbb{R}^n$ is the n-dimensional vector function $\nabla f(b)$ defined by

$$\nabla f(b) = \left(\frac{\partial f}{\partial x_1}(b), \frac{\partial f}{\partial x_2}(b), \ldots, \frac{\partial f}{\partial x_n}(b) \right).$$

Geometrical meaning: the gradient points is the direction of the greatest rate of increase of the function. Its magnitude equals the rate of increase.



Critical points – two variables

The **critical points** of a two variable function are those points where the **tangent plane** is parallel to the plane given by the x-axis and the y-axis or where the gradient does not exist.

- The first class of these points can be found as a solution of

$$\nabla f(x, y) = (0, 0)$$

which leads to the system of two equations for two variables

$$\begin{cases} \dfrac{\partial f}{\partial x}(x, y) &= 0 \\ \dfrac{\partial f}{\partial y}(x, y) &= 0 \end{cases} .$$

- In the second class there are the points where at at least one of the two partial derivates does not exist.

Hessian matrix

We can use the second derivative to decide the type of the critical point.

Definition

For a function $f(x_1, x_2, \ldots, x_n)$ *we define the* **Hessian matrix** *as*

$$\nabla^2 f(x_1, x_2, \ldots, x_n) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_1, \ldots, x_n) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_1, \ldots, x_n) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_1, \ldots, x_n) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x_1, \ldots, x_n) \end{pmatrix}$$

assuming that all the derivatives exist.

Positively and negatively definite

Definition

A matrix $A \in \mathbb{R}^{n,n}$ *is*

1. **positively definite** *if for all non-zero vectors* $a \in \mathbb{R}^n$ *it holds that* $aAa^T > 0$;
2. **negatively definite** *if for all non-zero vectors* $a \in \mathbb{R}^n$ *it holds that* $aAa^T < 0$;
3. **indefinite** *otherwise (not even positively/negatively semidefinite).*

Type of a critical point

Theorem

If $f : \mathbb{R}^n \to \mathbb{R}$ has all second partial derivative continuous at a critical point $b \in \mathbb{R}^n$, then

1. *if $\nabla^2 f(b)$ is* **positively definite***, then b is a point of* **strict local minimum***;*
2. *if $\nabla^2 f(b)$ is* **negatively definite***, then b is a point of* **strict local maximum***;*
3. *if $\nabla^2 f(b)$ is* **indefinite***, then b is a* **saddle point***.*

Sylvester's criterion on definiteness

For an $n \times n$ dimensional **symmetric** matrix A we define the **principal minors**:

- M_1 is the upper left 1-by-1 corner of A,
- M_2 is the upper left 2-by-2 corner of A,
- ...
- M_n is the upper left n-by-n corner of A.

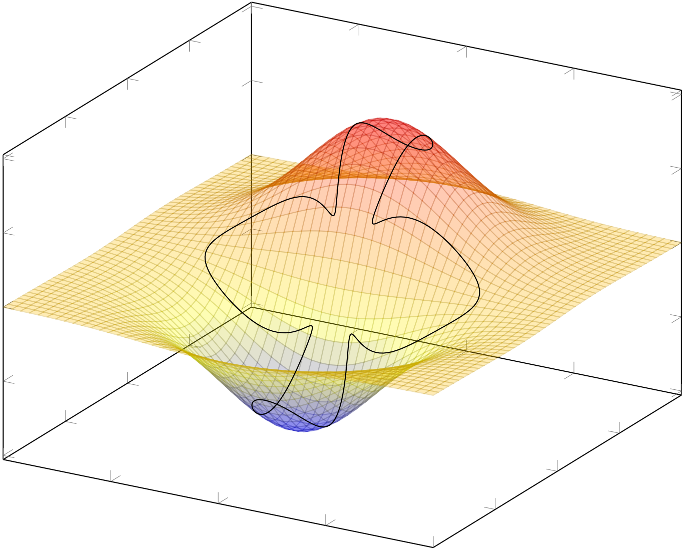
Theorem

Let $A \in \mathbb{R}^{n,n}$ be a symmetric matrix.

- *A is* **positively definite** *if and only if the determinants of all principal minors are positive.*
- *A is* **negatively definite** *if and only if the determinant of M_i is negative for odd i and positive for even i.*

Motivation

Find the maximum and minimum points **when walking along the black line**:



The problem

Let $f : \mathbb{R}^n \to \mathbb{R}$. Find (local) maxima and minima of f **subject to**

$$\begin{cases} g_1(x_1, x_2, \ldots, x_n) & = & 0 \\ g_2(x_1, x_2, \ldots, x_n) & = & 0 \\ & \vdots & \\ g_p(x_1, x_2, \ldots, x_n) & = & 0. \end{cases}$$

Set $\mathcal{G} = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid g_i(x_1, x_2, \ldots, x_n) = 0, i = 1, 2, \ldots, p\}$.

1. The functions f and g_i, with $i = 1, 2, \ldots, p$, have continuous second partial derivatives.
2. The gradients $\nabla g_1(x), \nabla g_2(x), \ldots, \nabla g_p(x)$ form a linearly independent set for all $x \in \mathcal{G}$.

Necessary condition

Theorem

Assume f has a local extremum in $x^* = (x_1^*, \ldots, x_n^*) \in \mathcal{G}$ subject to $\mathcal{G}$.
Then there exist $\mu_1^*, \ldots, \mu_p^*$ such that the **Lagrangian function** *L given by*

$$L(x_1, \ldots, x_n, \mu_1, \ldots, \mu_p) = f(x_1, \ldots, x_n) + \sum_{i=1}^{p} \mu_i g_i(x_1, \ldots, x_n)$$

has zero partial derivatives with respect to $x_1, \ldots, x_n$ at the point x^*.
In other words, the following system of equations is true:

$$\begin{cases} \dfrac{\partial f}{\partial x_1}(x^*) + \mu_1^* \dfrac{\partial g_1}{\partial x_1}(x^*) + \cdots + \mu_p^* \dfrac{\partial g_p}{\partial x_1}(x^*) = 0 \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(x^*) + \mu_1^* \dfrac{\partial g_1}{\partial x_n}(x^*) + \cdots + \mu_p^* \dfrac{\partial g_p}{\partial x_n}(x^*) = 0 \end{cases}$$

Sufficient condition

Theorem

Let $x^* = (x_1^*, \ldots, x_n^*) \in \mathbb{R}^n$ *and* $\mu^* = (\mu_1^*, \ldots, \mu_p^*) \in \mathbb{R}^p$ *such that*

1. *the Lagrangian function* $L(x_1, \ldots, x_n, \mu_1, \ldots, \mu_p)$ *has zero partial derivatives with respect to* $x_1, \ldots, x_n$ *at the point* $(x^*, \mu^*) \in \mathbb{R}^{n+p}$*;*
2. *the Lagrangian function* $L(x_1, \ldots, x_n, \mu_1, \ldots, \mu_p)$ *has zero partial derivatives with respect to* $\mu_1, \ldots, \mu_p$ *at the point* $(x^*, \mu^*) \in \mathbb{R}^{n+p}$*;*
3. *for all non-zero* $y \in \mathbb{R}^n$ *satisfying* $y \cdot \nabla g_i(x^*) = 0$ *for* $i = 1, 2, \ldots, p$ *we have*

$$y \left(\nabla^2 f(x^*) + \sum_{i=1}^{p} \mu_i^* \nabla^2 g_i(x^*) \right) y^T > 0.$$

Thus, the function f *has a strict local minimum at* x^* *(subject to* $\mathcal{G}$*).*

If we replace in (iii) the condition "> 0" by "< 0", we obtain a sufficient condition of a strict local maximum.

How to calculate a double integral?

The following statement can be derived from the definition.

Theorem

If f is integrable over $D = [a, b] \times [c, d]$ and one of the integrals

$$\int_a^b \left(\int_c^d f(x, y) dy \right) dx \quad \text{or} \quad \int_c^d \left(\int_a^b f(x, y) dx \right) dy$$

exists, then it is equal to

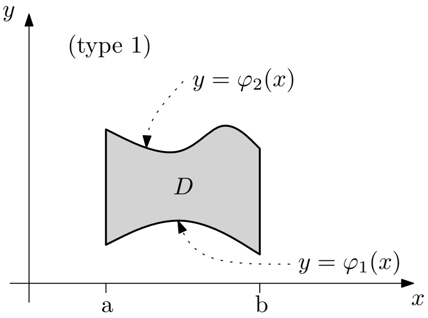
$$\iint_D f(x, y) dxdy.$$

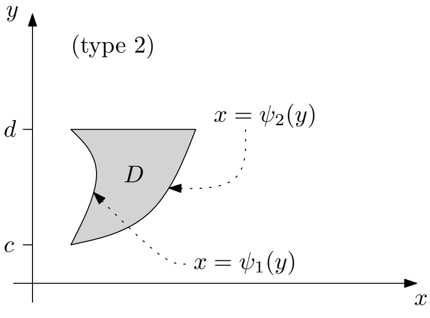
And if D is not a rectangle?
The definition is very similar: we approximate D using smaller and smaller rectangular areas...

Special types of domain D (1/2)

We will consider the following two types of the domain D.

- (type 1) $x \in [a, b]$ and y is bounded by $\varphi_1(x)$ and $\varphi_2(x)$;
- (type 2) $y \in [c, d]$ and x is bounded by $\psi_1(y)$ and $\psi_2(y)$.





Special types of domain D (2/2)

Double integrals over such D are calculated as follows.

Theorem

If the integral on the right side exists, then we have (for such a domain D):

- *if D is of type 1, then*

$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\mathrm{d}y \right) \mathrm{d}x;$$

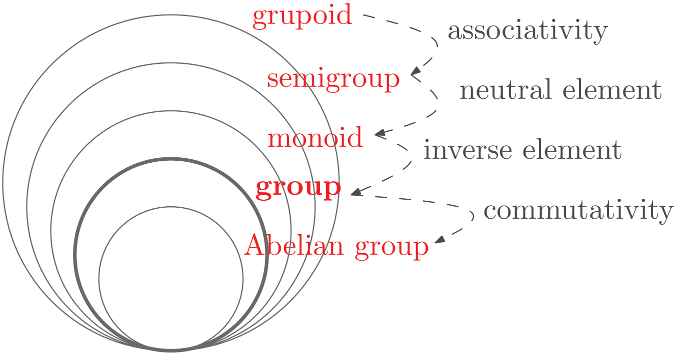
- *if D is of type 2, then*

$$\iint_D f(x,y)\mathrm{d}x\mathrm{d}y = \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\mathrm{d}x \right) \mathrm{d}y.$$

Recap

Sets with one binary operation

We call an arbitrary pair "a set and a binary operation" a groupoid. Adding another requirements we get further notions.



Definitions and elementary properties

Definition

- *An ordered pair* $(M, \circ)$, *where* M *is an arbitrary non-empty set and* $\circ$ *is a binary operation on* M, *is called a* **groupoid**.
- *A groupoid* $(M, \circ)$ *such that* $\circ$ *is associative is called a* **semigroup**.
- *A semigroup* $(M, \circ)$ *such that there exists a* **neutral element** e *satisfying*

$$\forall\, a \in M \quad \textit{holds} \quad e \circ a = a \circ e = a$$

  is called a **monoid**.
- *A monoid* $(M, \circ)$ *such that for each* $a \in M$ *there exists an* **inverse element** $a^{-1} \in M$ *satisfying*

$$a^{-1} \circ a = a \circ a^{-1} = e$$

  is called a **group**.
- *Moreover, if* $\circ$ *is commutative, we say that a group* $(M, \circ)$ *is a* **commutative (Abelian) group**.

Cayley tables for finite groups

If the set M from the pair $(M, \circ)$ has a finite number of elements, its structure (with the given operation $\circ$) could be completely represented by the **Cayley table**. Its construction is obvious from the following example.

Example

Let us consider $(\mathbb{Z}_4, +_4)$*, i.e., the set of numbers* $\{0, 1, 2, 3\}$ *with addition modulo* 4*.*
Since the set has 4 *elements, the Cayley table has* 4 *rows and* 4 *columns:*

$+_4$	0	1	2	3
0				
1				
2				1
3				

Cayley tables for finite groups

If the set M from the pair $(M, \circ)$ has a finite number of elements, its structure (with the given operation $\circ$) could be completely represented by the **Cayley table**. Its construction is obvious from the following example.

Example

Let us consider $(\mathbb{Z}_4, +_4)$*, i.e., the set of numbers* $\{0, 1, 2, 3\}$ *with addition modulo* 4*.*

Since the set has 4 *elements, the Cayley table has* 4 *rows and* 4 *columns:*

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Definition of subgroup

Definition

Let $G = (M, \circ)$ *be a group.*

A **subgroup** *of the group* G *is a pair* $H = (N, \circ)$ *such that:*

- $N \subseteq M$ *and* $N \neq \emptyset$,
- H *is a group.*

Trivial and proper subgroups

In each group $G = (M, \circ)$, there always exist at least two subgroups (if M contains only one element the two coincide):

- the group containing only the neutral element: $(\{e\}, \circ)$,
- and the group itself $G = (M, \circ)$.

These two groups are the **trivial subgroups**; other subgroups are non-trivial or **proper subgroups**.

Power of an element

Definition

Let $G = (M, \circ)$ be a group with neutral element e. We define for each element $a \in M$ and each positive $n \in \mathbb{N}$ the **n-th power of the element a** *as*

$$\begin{aligned} a^0 &= e \\ a^n &= \underbrace{a \circ a \circ \cdots \circ a}_{n \text{ times}} \\ a^{-n} &= (a^{-1})^n = \underbrace{a^{-1} \circ a^{-1} \circ \cdots \circ a^{-1}}_{n \text{ times}} \end{aligned}$$

Note that $a \circ a \circ \cdots \circ a$ can by written without brackets thanks to associativity (for a non-associative operation the result would depend on the order. . .).

For the additive notation of a group $G = (M, +)$, we define the **n-th multiple of the element a** and we denote it by $n \times a$ (resp. $-n \times a = n \times (-a)$).

Order of a (sub)group

Definition

The **order of a (sub)group** $G = (M, \circ)$*, denoted* $|G|$*, is its number of elements. If* M *is an infinite set, the order is infinite. According to the order we distinguish between* **finite** *and* **infinite groups**.

Example

The group $\mathbb{Z}_{12}^{+}$ *is of order* 12*. It has* 6 *subgroups:*

- *two trivial:* $\{0\}$ *and* $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$*;*
- *and four proper:* $\{0, 6\}$, $\{0, 4, 8\}$, $\{0, 3, 6, 9\}$*, and* $\{0, 2, 4, 6, 8, 10\}$.

of order respectively 1, 2, 3, 4, 6 *and* 12.

Lagrange's Theorem

Let $[G : H]$ denote the **index** of a subgroup H in G (i.e. the number of different cosets of H in G).

Theorem

Let H be a subgroup of a finite group G.
The order of H divides the order of G.
More precisely,

$$|G| = [G : H] \cdot |H|.$$

(Sub)group generated by a set

Definition

Let $G = (M, \circ)$ be a group and $N \subset M$ a nonempty set. The smallest subgroup of G containing N is the **subgroup generated by** N *and is denoted by $\langle N \rangle$.*

In particular, for a singleton $N = \{a\}$ we use the notation $\langle a \rangle = \langle \{a\} \rangle$.

(Sub)group generated by a set

Definition

Let $G = (M, \circ)$ be a group and $N \subset M$ a nonempty set. The smallest subgroup of G containing N is the **subgroup generated by** *N and is denoted by $\langle N \rangle$.*

In particular, for a singleton $N = \{a\}$ we use the notation $\langle a \rangle = \langle \{a\} \rangle$.

Theorem

Let $G = (M, \circ)$ be a group and $N \subset M$ a nonempty set. The following holds:

- *the subgroup $\langle N \rangle$ equals the intersection of all subgroups containing N, i.e.*

$$\langle N \rangle = \bigcap \{H \colon H \text{ is a subgroup of } G \text{ containing } N\}$$

(Sub)group generated by a set

Definition

Let $G = (M, \circ)$ be a group and $N \subset M$ a nonempty set. The smallest subgroup of G containing N is the **subgroup generated by** *N and is denoted by $\langle N \rangle$.*

In particular, for a singleton $N = \{a\}$ we use the notation $\langle a \rangle = \langle \{a\} \rangle$.

Theorem

Let $G = (M, \circ)$ be a group and $N \subset M$ a nonempty set. The following holds:

- *the subgroup $\langle N \rangle$ equals the intersection of all subgroups containing N, i.e.*

$$\langle N \rangle = \bigcap \{H \colon H \text{ is a subgroup of } G \text{ containing } N\}$$

- *all elements belonging to $\langle N \rangle$ can be obtained by means of "group span", i.e.,*

$$\left\{ a_1^{k_1} \circ a_2^{k_2} \circ \cdots a_n^{k_n} \ : \ n \in \mathbb{N}, a_i \in N, k_i \in \mathbb{Z} \right\}.$$

(Sub)group generated by a set

Definition

Let $G = (M, \circ)$ be a group and $N \subset M$ a nonempty set. The smallest subgroup of G containing N is the **subgroup generated by** *N and is denoted by $\langle N \rangle$.*

In particular, for a singleton $N = \{a\}$ we use the notation $\langle a \rangle = \langle \{a\} \rangle$.

Theorem

Let $G = (M, \circ)$ be a group and $N \subset M$ a nonempty set. The following holds:

- *the subgroup $\langle N \rangle$ equals the intersection of all subgroups containing N, i.e.*

$$\langle N \rangle = \bigcap \{H \colon H \text{ is a subgroup of } G \text{ containing } N\}$$

- *all elements belonging to $\langle N \rangle$ can be obtained by means of "group span", i.e.,*

$$\left\{ a_1^{k_1} \circ a_2^{k_2} \circ \cdots a_n^{k_n} \ : \ n \in \mathbb{N}, a_i \in N, k_i \in \mathbb{Z} \right\}.$$

Examples of groups generated by one element

Theorem

An additive group modulo n is equal to $\langle k \rangle$ if and only if k and n are coprime numbers.

The **multiplicative group modulo p**, denoted $\mathbb{Z}_p^\times$, where p is a prime number, is the set $\{1, 2, \ldots, p-1\}$ with the operation of multiplication modulo p.

Example

Is there a one-element set generating the group $\mathbb{Z}_{11}^\times$?

Yes, for example $\langle 2 \rangle = \mathbb{Z}_{11}^\times$.

On the other hand, $\langle 3 \rangle = (\{1, 3, 4, 5, 9\}, \cdot \ (mod\ 11))$.

Definition of cyclic group

Definition

A group $G = (M, \circ)$ is **cyclic** *if there exists an element $a \in M$ such that $\langle a \rangle = G$. This element is a* **generator** *of the cyclic group.*

- $\mathbb{Z}_n^+$ is a cyclic group for every n and its generators are all positive numbers $k \leq n$ coprime with n.
- the infinite group $(\mathbb{Z}, +)$ is cyclic and it has just two generators: 1 and -1.
- $\mathbb{Z}_{11}^\times$ is cyclic, and 2 is a generator.

Fermat's Theorem

Theorem

In a cyclic group $G = (M, \circ)$ of order n, for all elements $a \in M$, it holds that

$$a^n = e$$

Where e is the neutral element of G.

Fermat's Theorem

Theorem

In a cyclic group $G = (M, \circ)$ of order n, for all elements $a \in M$, it holds that

$$a^n = e$$

Where e is the neutral element of G.

$\mathbb{Z}_p^\times$ is always a cyclic group and the order of this group is $p-1$.

Applying the previous statement to $\mathbb{Z}_p^\times$ we obtain the well-known **Fermat's Little Theorem**.

Corollary (Fermat's Little Theorem)

For an arbitrary prime number p and an arbitrary $1 \leq a < p$ we have that

$$a^{p-1} \equiv 1 \,(mod\ p).$$

How to find all generators

Generally, to find all generators is not an easy task (e.g., in groups $\mathbb{Z}_p^\times$ we are not able to do it algorithmically); but if we have one, it is easy to find all the others.

Theorem

If $(G, \circ)$ is a cyclic group of order n and a is one of its generator, then a^k is a generator if and only if k and n are coprime.

Subgroups of cyclic group are cyclic

Theorem

Any subgroup of a cyclic group is again a cyclic group.

The same groups and distinct elements (1/3)

$\mathbb{Z}_5^\times$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

order: 4

subgroups: $\{1\}$, $\{1,4\}$, $\{1,2,3,4\}$

neutral element: 1

inverse elements: $1^{-1} = 1, \quad 2^{-1} = 3,$ $3^{-1} = 2, \quad 4^{-1} = 4.$

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

order: 4

subgroups: $\{0\}$, $\{0,2\}$, $\{0,1,2,3\}$

neutral element: 0

inverse elements: $0^{-1} = 0, \quad 1^{-1} = 3,$ $2^{-1} = 2, \quad 3^{-1} = 1.$

Aren't these two groups in fact the same group differing only in the "names" of their elements?

The same groups and distinct elements (2/3)

$\mathbb{Z}_5^\times$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Let us try to rename the elements of the group $\mathbb{Z}_5^\times$ so to get $\mathbb{Z}_4^+$:

The same groups and distinct elements (2/3)

$\mathbb{Z}_5^\times$	0	2	3	4
0	0	2	3	4
2	2	4	0	3
3	3	0	4	2
4	4	3	2	0

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Let us try to rename the elements of the group $\mathbb{Z}_5^\times$ so to get $\mathbb{Z}_4^+$:

- The neutral element has very special and unique properties: we rename 1 to 0.

The same groups and distinct elements (2/3)

$\mathbb{Z}_5^{\times}$	0	2	3	2
0	0	2	3	2
2	2	2	0	3
3	3	0	2	2
2	2	3	2	0

$\mathbb{Z}_4^{+}$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Let us try to rename the elements of the group $\mathbb{Z}_5^{\times}$ so to get $\mathbb{Z}_4^{+}$:

- The neutral element has very special and unique properties: we rename 1 to 0.
- If the complete structure should be preserved, then the only two-elements subgroup $\{1, 4\}$ (in $\mathbb{Z}_5^{\times}$) must correspond to the subgroup $\{0, 2\}$ (in $\mathbb{Z}_4^{+}$): we map $4 \leftrightarrow 2$.

The same groups and distinct elements (2/3)

$\mathbb{Z}_5^\times$	0	3	1	2
0	0	3	1	2
3	3	2	0	1
1	1	0	2	3
2	2	1	3	0

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Let us try to rename the elements of the group $\mathbb{Z}_5^\times$ so to get $\mathbb{Z}_4^+$:

- The neutral element has very special and unique properties: we rename 1 to 0.
- If the complete structure should be preserved, then the only two-elements subgroup $\{1, 4\}$ (in $\mathbb{Z}_5^\times$) must correspond to the subgroup $\{0, 2\}$ (in $\mathbb{Z}_4^+$): we map $4 \leftrightarrow 2$.
- Now, it remains to rename only 2 and 3; we can check that both remaining possibilities work; we choose $3 \leftrightarrow 1$ and $2 \leftrightarrow 3$.

The same groups and distinct elements (2/3)

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\mathbb{Z}_4^+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Let us try to rename the elements of the group $\mathbb{Z}_5^\times$ so to get $\mathbb{Z}_4^+$:

- The neutral element has very special and unique properties: we rename 1 to 0.
- If the complete structure should be preserved, then the only two-elements subgroup $\{1, 4\}$ (in $\mathbb{Z}_5^\times$) must correspond to the subgroup $\{0, 2\}$ (in $\mathbb{Z}_4^+$): we map $4 \leftrightarrow 2$.
- Now, it remains to rename only 2 and 3; we can check that both remaining possibilities work; we choose $3 \leftrightarrow 1$ and $2 \leftrightarrow 3$.
- It suffices to reorder the rows. . . and we have the Cayley table of $\mathbb{Z}_4^+$.

The same groups and distinct elements (3/3)

The desired property for the bijections is that for all $n, m \in \{1, 2, 3, 4\}$, we have

$$\varphi(n \times_5 m) \; = \; \varphi(n) +_4 \varphi(m)$$

where $\times_5$ denotes the operation in the group $\mathbb{Z}_5^\times$, and $+_4$ the one in the group $\mathbb{Z}_4^+$.

The same groups and distinct elements (3/3)

The desired property for the bijections is that for all $n, m \in \{1, 2, 3, 4\}$, we have

$$\varphi(n \times_5 m) = \varphi(n) +_4 \varphi(m)$$

where $\times_5$ denotes the operation in the group $\mathbb{Z}_5^\times$, and $+_4$ the one in the group $\mathbb{Z}_4^+$.

Homomorphism and isomorphism

Definition

Let $G = (M, \circ_G)$ and $H = (N, \circ_H)$ be two groupoids. The mapping $\varphi : M \to N$ is a **homomorphism** *from G to H if*

$$\text{for all } x, y \in M, \text{ we have } \varphi(x \circ_G y) = \varphi(x) \circ_H \varphi(y).$$

If, moreover, φ is injective (resp. surjective, resp. bijective) we say that φ is a **monomorphism** *(resp.* **epimorphism**, *resp.* **isomorphism***).*

Isomorphic groups

Definition

If there exists an isomorphism between two groups, these groups are **isomorphic***.*

Example

The two groups $\mathbb{Z}_5^{\times}$ *and* $\mathbb{Z}_4^{+}$ *are isomorphic.*

Isomorphic groups have the same order.

Fundamental properties of homomorphisms

Theorem

Let φ *be a homomorphism from a group* $G = (M, \circ_G)$ *to* $H = (N, \circ_H)$*.*
The group $\varphi(G) = (\varphi(M), \circ_H)$ *is a subgroup of* H*.*

Fundamental properties of homomorphisms

Theorem

Let φ be a homomorphism from a group $G = (M, \circ_G)$ to $H = (N, \circ_H)$. The group $\varphi(G) = (\varphi(M), \circ_H)$ is a subgroup of H.

Consequences of the previous theorem

- A homomorphism always maps the neutral element of one group to the neutral element of the other group.
- Inverse elements are preserved as well: $\varphi(x^{-1}) = \varphi(x)^{-1}$.

. . . up to isomorphism (1/3)

If we say that there exists one group with a certain property **up to isomorphism**, it means that all groups with this property are isomorphic to each other.

Theorem

Any two infinite cyclic groups are isomorphic.
For each $n \in \mathbb{N}$, any two cyclic groups of order n are isomorphic.

$(\mathbb{Z}, +)$ and $\mathbb{Z}_n^+$ are the only cyclic groups up to isomorphism.

. . . up to isophormism (2/3)

The **Klein group** is the group $(\mathbb{Z}_2 \times \mathbb{Z}_2, \circ)$, where

$$\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0,0),(0,1),(1,0),(1,1)\}$$

and $\circ$ is the component-wise addition modulo 2: e.g., $(1,0) \circ (1,1) = (0,1)$.

The Klein group is not cyclic and thus cannot be isomorphic to $\mathbb{Z}_4^+$!
It is possible to show this (try it, it is easy):

Theorem

There exists only two groups of order 4 which are not isomorphic.

$\mathbb{Z}_4^+$ and the Klein group are the only two groups of order 4 up to isomorphism.

. . . up to isomorphism (3/3)

The **symmetric group** S_n of the set of all permutations over $\{1, 2, 3, \ldots, n\}$ with the operation of composition.

A (n-)**permutation** $\pi \in S_n$ is a bijection of the set $\{1, 2, 3, \ldots, n\}$ to itselfand can be defined by listing its values:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \pi(1) & \pi(2) & \pi(3) & \cdots & \pi(n) \end{pmatrix}.$$

The first row could by deleted, and so, e.g., $(1\ 2\ 4\ 3\ 5) \in S_5$ is the permutation swapping elements 3 and 4.

The composition of permutations is associative, the permutation $(1\ 2\ 3\ \cdots n)$ is the neutral element, and the inverse element is the inverse permutation.

. . . up to isomorphism (3/3)

The **symmetric group** S_n of the set of all permutations over $\{1, 2, 3, \ldots, n\}$ with the operation of composition.

A (n-)**permutation** $\pi \in S_n$ is a bijection of the set $\{1, 2, 3, \ldots, n\}$ to itselfand can be defined by listing its values:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \pi(1) & \pi(2) & \pi(3) & \cdots & \pi(n) \end{pmatrix}.$$

The first row could by deleted, and so, e.g., $(1\ 2\ 4\ 3\ 5) \in S_5$ is the permutation swapping elements 3 and 4.

The composition of permutations is associative, the permutation $(1\ 2\ 3\ \cdots n)$ is the neutral element, and the inverse element is the inverse permutation.

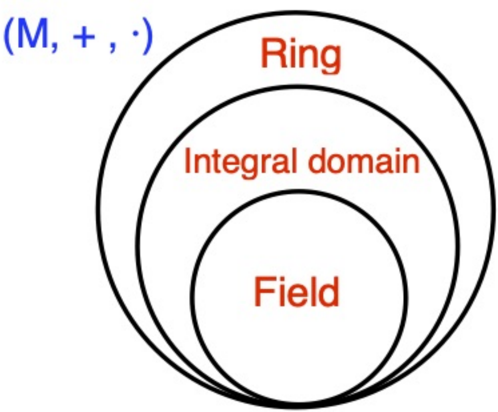
Subgroups of the symmetric group S_n are called **groups of permutations**.

Theorem (Cayley)

Each finite group is isomorphic to some group of permutations.

Sets with two binary operations

For more sophisticated arithmetical operations with numbers we need **both** addition and multiplication.



Definition of a ring

Definition (Ring)

Let M be a nonempty set, and $+$ and $\cdot$ two binary operations. We say that $R = (M, +, \cdot)$ is a **ring** *if the following holds:*

- $(M, +)$ *is an* ***Abelian group****,*
- $(M, \cdot)$ *is a* ***monoid****,*
- *both left and right* ***distributive law*** *hold:*

$$(\forall\ a, b, c \in M) \text{ we have: } \quad a(b+c) = ab + ac \quad \wedge \quad (b+c)a = ba + ca.$$

We respect the standard convention that the multiplication has a higher priority than the addition.

Terminology

Let $R = (M, +, \cdot)$ be a ring.

- If $\cdot$ is commutative, R is a **commutative ring**.
- $(M, +)$ is called the **additive group** of the ring R.
- $(M, \cdot)$ is called the **multiplicative monoid** of the ring R.
- The neutral element of the group $(M, +)$ is called the **zero element** and is denoted by 0; the inverse element to $a \in M$ is denoted as $-a$.
- Inside the ring **we can define subtraction** by

$$a - b := a + (-b).$$

Basic properties of rings

In an arbitrary ring $(M, +, \cdot)$, the following holds.

- Left and right distributive law for subtracting, i.e.,

$$c(b - a) = cb - ca.$$

- Multiplying by the zero element returns the zero element, i.e.,

$$\forall a \in M \qquad a \cdot 0 = 0 \wedge 0 \cdot a = 0.$$

Integral domain

Definition (zero divisors)

Let $R = (M, +, \cdot)$ be a ring. Two arbitrary ***nonzero*** *elements $a, b \in M$ such that*

$$a \cdot b = 0$$

are called **zero divisors**.

Definition (integral domain)

A commutative ring ***without*** *zero divisors is called an* **integral domain**.

Definition of field

Definition (field)

A ring $T = (M, +, \cdot)$ is a **field** *if $(M \setminus \{0\}, \cdot)$ is an Abelian group. This group is called the* **multiplicative group** *of the field T.*

Finite fields

A field with finite number of elements is called **finite**. The number of elements is said to be the **order** of the field.

An example of finite field is the set (of residue classes modulo p)

$$\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$$

with operations modulo a **prime** p.

E.g., for $p = 5$ we obtain the field with following operations:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

and

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Orders of finite fields

We have shown a construction of a finite field of order p with p prime.
Are there fields of any arbitrary order?

Theorem

*Any finite field has **order $\mathbf{p^n}$**, where p is a prime number and n is a positive natural number.*
*The prime number p is called the **characteristic** of the field.*
Furthermore, all fields of order p^n are isomorphic.
Additionally, the multiplicative group of a finite field is cyclic.

Consequence: There are no fields of order 6, 10, 12, 14, ...

If we chose $p = 2$ and $n = 8$, we obtain the field providing us with arithmetic on 1 byte (8 bits)!
The fields with 2^n elements are called **binary fields** and are denoted $GF(2^n)$ (as **Galois Fields**).

Operation in a Galois Field - The wrong way

Consider a field $GF(2^8)$. Each element can be represented as an 8 bit string, e.g., 11010110, 01100011, etc.

Addition: Addition can be defined component-wise modulo 2. i.e.

$$11010110 + 01100011 = (1 + 0 \mod 2)(1 + 1 \mod 2) \cdots$$
$$\cdots (0 + 1 \mod 2) = 10110101.$$

Operation in a Galois Field - The wrong way

Consider a field $GF(2^8)$. Each element can be represented as an 8 bit string, e.g., 11010110, 01100011, etc.

Addition: Addition can be defined component-wise modulo 2. i.e.

$$11010110 + 01100011 = (1 + 0 \mod 2)(1 + 1 \mod 2) \cdots \cdots (0 + 1 \mod 2) = 10110101.$$

The neutral (zero) element is 00000000, and each element is inverse to itself. We have an additive group.

Multiplication: Multiplication cannot be defined component-wise: The neutral element would be 11111111 and the inverse to (e.g.) 11111110 would not exist.

Operation in a Galois Field - The wrong way

Consider a field $GF(2^8)$. Each element can be represented as an 8 bit string, e.g., 11010110, 01100011, etc.

Addition: Addition can be defined component-wise modulo 2. i.e.

$$11010110 + 01100011 = (1+0 \mod 2)(1+1 \mod 2)\cdots \cdots(0+1 \mod 2) = 10110101.$$

The neutral (zero) element is 00000000, and each element is inverse to itself. We have an additive group.

Multiplication: Multiplication cannot be defined component-wise: The neutral element would be 11111111 and the inverse to (e.g.) 11111110 would not exist.

Multiplication must be defined in a different way!

Rings of polynomials over a ring / field

Definition

Let K be a ring. The set of polynomials with coefficients in K together with operations of addition and multiplication defined as

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i;$$

$$\left(\sum_{i=0}^{n} a_i x^i\right) \cdot \left(\sum_{i=0}^{m} b_i x^i\right) = \sum_{i=0}^{n+m} \left(\sum_{j+k=i} a_j b_k\right) x^i$$

is the **commutative ring of polynomials over the ring K**. *This ring is denoted as $K[x]$.*

Irreducible polynomial

Definition

Let K be a field and $P(x) \in K[x]$ be of degree at least 1. We say that $P(x)$ is **irreducible over** K *if, for any two polynomials $A(x)$ and $B(x)$ from $K[x]$, it holds that*

$$A(x) \cdot B(x) = P(x) \quad \Rightarrow \quad \big(\textit{degree of}\, A(x) = 0 \quad \vee \quad \textit{degree of}\, B(x) = 0\big).$$

Irreducible polynomial

Definition

Let K be a field and $P(x) \in K[x]$ be of degree at least 1. We say that $P(x)$ is **irreducible over K** *if, for any two polynomials $A(x)$ and $B(x)$ from $K[x]$, it holds that*

$$A(x) \cdot B(x) = P(x) \quad \Rightarrow \quad \big(\textit{degree of } A(x) = 0 \quad \vee \quad \textit{degree of } B(x) = 0\big).$$

Irreducible polynomials are primes among polynomials!

Their definition is analogous as well as their properties.

Example: Whereas $x^2 + 1$ is irreducible over the field $\mathbb{Q}$, the polynpmial $x^2 - 1 = (x+1)(x-1)$ is not.

Remark: $x^2 + 1$ is irreducible over the field $\mathbb{Q}$, but not over the field $\mathbb{Z}_2$, where the coefficients are added and multiplied modulo 2:

$$x^2 + 1 = (x+1)(x+1) = x^2 + 2x + 1.$$

Irreducible polynomial as a modulus

We define **modulo polynomial** as:

$$A(x) \pmod{P(x)} = \text{the remainder of the division of } A(x) \text{ by } P(x).$$

The result is always a polynomial of degree less than the degree of $P(x)$.

Example: for $A(x) = x^3$ and $P(x) = x^2 + 1$ we have $A(x) = x(x^2 + 1) + (-x)$ and thus

$$x^3 \equiv -x \pmod{x^2 + 1}.$$

Field $GF(2^4)$

The elements $GF(2^4)$ are represented as polynomials of order at most 3 with coefficients h_i from the field $\mathbb{Z}_2$:

$$h_3x^3 + h_2x^2 + h_1x + h_0 \approx (h_3h_2h_1h_0)_2.$$

Addition component-wise modulo 2:

$$(x^3 + x + 1) + (x^2 + x + 1) = x^3 + x^2.$$

Field $GF(2^4)$ – multiplication

Multiplication modulo a chosen irreducible polynomial, e.g., $x^4 + x + 1$.

Example: multiplication $A(x) \cdot B(x)$ for $A(x) = x^3 + x^2 + 1$ and $B(x) = x^2 + x$.

Field $GF(2^4)$ – multiplication

Multiplication modulo a chosen irreducible polynomial, e.g., $x^4 + x + 1$.

Example: multiplication $A(x) \cdot B(x)$ for $A(x) = x^3 + x^2 + 1$ and $B(x) = x^2 + x$.

1. Multiply $A(x) \cdot B(x)$ classically and rewrite coefficients $\mod 2$:

$$A(x) \cdot B(x) = x^5 + 2x^4 + x^3 + x^2 + x = x^5 + x^3 + x^2 + x.$$

2. Find the remainder after division by $P(x)$. Since

$$x^5 = x(x^4 + x + 1) + (x^2 + x), \quad \text{it holds} \quad x^5 \equiv x^2 + x \pmod{x^4 + x + 1},$$

and we have

$$x^5 + x^3 + x^2 + x \equiv (x^2 + x) + (x^3 + x^2 + x) \equiv \mathbf{x^3} \pmod{x^4 + x + 1}.$$

Hence we get that $1101 \cdot 0110 = 1000$.

Construction of a finite field

In general, we construct a finite field $GF(p^k)$ using polynomials as follows.

Let $m(x) \in \mathbb{Z}_p[x]$ be an irreducible polynomial of degree k.

$$GF(p^k) = \left(\{q(x) \in \mathbb{Z}_p[x] \colon \deg(q) < k\}, \ +, \ \times \text{ mod } m(x) \right).$$

Recap

1. Multivariate optimization
   - Critical points
   - Constrained optimization
   - 2-variate function integration
2. General algebra
   - Groupoid, semigroup, monoid, group
   - Subgroups
   - Cyclic groups
   - Homomorphisms
   - Rings and fields
3. Fuzzy Logic
   - Fuzzy sets
   - Operation on fuzzy sets
4. Numerical Mathematics
   - Representation of numbers

Universe and crisp sets

Let U denote the **universe**, that is, our playground containing every set that we may consider.

A set $A \subset U$ can be given by its **characteristic function**:

$$\chi_A : U \to \{0, 1\}, \qquad \chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

There is a bijection between sets and characteristic functions, so we identify each set with its characteristic function.

A is a set in the ordinary sense, sometimes called a **crisp** set.

Fuzzy sets

Fuzzy sets generalize this concept and allow elements to belong to a given set with a certain *degree*.

We replace the characteristic function by a **membership function**

$$\mu_A : U \to [0,1].$$

A **fuzzy subset** A of a set X is a function $\mu_A : X \to [0,1]$.

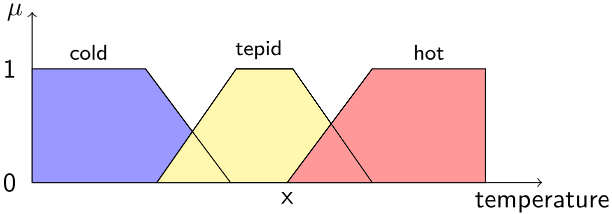
For every element $x \in X$, the **degree of membership** of x to A is given by $\mu_A(x) \in [0,1]$.

Example

Let $X = [0, 100]$ be the set of temperatures of water in our pot.

We consider three fuzzy subsets of X to describe cold, tepid and hot temperatures.

The membership functions may be given as follows:



Operations on crisp sets

Given a set x and its power set $\mathcal{P}(X)$ (the set of all subsets of X), the operations of **union**, **intersection**, and **complement** are given as follows (for the usual sets):

$$A \cup B = \{x : \; x \in A \text{ or } x \in B\}\,,$$
$$A \cap B = \{x : \; x \in A \text{ and } x \in B\}\,,$$
$$A^{\complement} = X \setminus A = \{x \in X : \; x \notin A\}\,.$$

How do these operations translate to characteristic functions?

$$\chi_{A\cup B} = \max\{\chi_A, \chi_B\},$$
$$\chi_{A\cap B} = \min\{\chi_A, \chi_B\},$$
$$\chi_{A^{\complement}} = 1 - \chi_A.$$

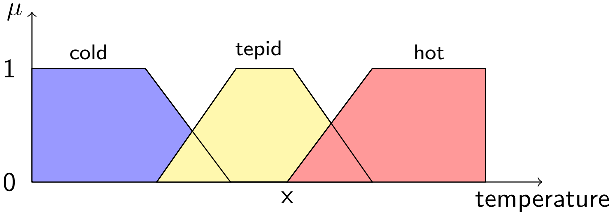
Operations on fuzzy sets

For fuzzy sets, we can define the membership function of a union, intersection, or a complement in the same way.

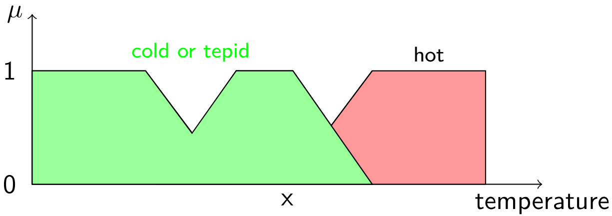
Let A and B be two *fuzzy* subsets of X.
We set

$$\mu_{A \cup B} = \max\{\mu_A, \mu_B\},$$
$$\mu_{A \cap B} = \min\{\mu_A, \mu_B\},$$
$$\mu_{A^{\complement}} = 1 - \mu_A.$$

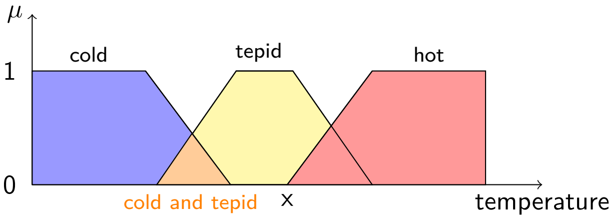
Examples



Examples



Examples



Operations revisited

Our choice for fuzzy set operation was fast.
Let A and B be two subsets of X. We have

$$\begin{aligned}\chi_{A\cap B} &= \min\{\chi_A,\ \chi_B\} \\ &= \chi_A\chi_B \\ &= \max\{0,\ \chi_A(x)+\chi_B(x)-1\}\,.\end{aligned}$$

We can extend the second definition to membership functions and obtain another definition of intersaction (and union) of fuzzy sets.
We shall do this in a more general fashion.

t-norms

We have the following requirements of a mapping

$$\star : [0,1] \times [0,1] \to [0,1]$$

that would interpret intersection of two fuzzy sets.

1. $1 \star x = x$ for all $x \in [0,1]$,
2. $0 \star x = 0$ for all $x \in [0,1]$,
3. $x \star y = y \star x$ for all $x, y \in [0,1]$ (*commutativity*),
4. $(x \star y) \star z = x \star (y \star z)$ for all $x, y, z \in [0,1]$ (*associativity*),
5. $x \leq y$ and $w \leq z$ implies $x \star w \leq y \star z$ (*monotonicity*).

Various t-norms

The following t-norms are usually considered.

Let $x, y \in [0,1]$.

(i) **Gödel** t-norm: $x \star y = \min\{x, y\}$,

(ii) **product** t-norm: $x \star y = x \cdot y$,

(iii) **Łukasiewicz** t-norm: $x \star y = \max\{0, x + y - 1\}$,

(iv) **Hamacher product** t-norm: $x \star y = \begin{cases} 0 & \text{if } x = y = 0 \\ \dfrac{xy}{x + y - xy} & \text{otherwise} \end{cases}$,

(v) ...

The distinct t-norms give us distinct strategies on how to interpret intersection of fuzzy sets.

If we have intersection and complement, we define union by $A \cup B = \left(A^{\complement} \cap B^{\complement}\right)^{\complement}$ (**De Morgan's laws**).

Recap

1. Multivariate optimization
   - Critical points
   - Constrained optimization
   - 2-variate function integration
2. General algebra
   - Groupoid, semigroup, monoid, group
   - Subgroups
   - Cyclic groups
   - Homomorphisms
   - Rings and fields
3. Fuzzy Logic
   - Fuzzy sets
   - Operation on fuzzy sets
4. Numerical Mathematics
   - Representation of numbers

Scientific notation

To store a number in computer we usually use the binary number system.

$$(6)_{10} = (110)_2 \qquad (0.1)_{10} = (0.000110011001100110011001100110011\ldots)_2$$

For non-integers, one can use the **scientific notation**. In the binary base a number x is represented as

$$x = \pm\, m \cdot 2^e.$$

m - **mantissa/significand** having a fixed number of digits / fixed length; these digits are also called **significant digits**.

e - **exponent** having a fixed number of digits / fixed length.

IEEE-754

A number x is represented by its sign s and by the numbers e and m. The standard `IEEE-754` defines the following lengths of e and m and their interpretation.

precision	length of m	d = length of e	b
`binary32` / single precision	23	8	127
`binary64` / double precision	52	11	1023
`binary128` / quadruple precision	112	15	16383

- if $e = 2^d - 1$ and $m \neq 0$, then $x = \texttt{NaN}$ (Not a Number)
- if $e = 2^d - 1$ and $m = 0$ and $s = 0$, then $x = +\texttt{Inf}$
- if $e = 2^d - 1$ and $m = 0$ and $s = 1$, then $x = -\texttt{Inf}$
- if $0 < e < 2^d - 1$, the $x = (-1)^s \cdot (1.m)_2 \cdot 2^{e-b}$ (so-called **normalized numbers**)
- if $e = 0$ and $m \neq 0$, then $x = (-1)^s \cdot (0.m)_2 \cdot 2^{-b+1}$ (so-called **subnormal/unnormalized numbers**)
- if $e = 0$ and $m = 0$ and $s = 0$, then $x = 0$
- if $e = 0$ and $m = 0$ and $s = 1$, then $x = -0$

Machine numbers

The numbers that can be represented as floating point numbers (with selected finite lengths of m and e) are called **machine numbers**.

The set of machine numbers F has the largest and the smallest positive elements as follows:

precision	max. no.	min. pos. normalized	min. pos. subnormal
single	$(2-2^{-23})\cdot 2^{127}$ $\approx 3.4\cdot 10^{38}$	2^{-126} $\approx 1.2\cdot 10^{-38}$	$2^{-126-23}=2^{-149}$ $\approx 1.4\cdot 10^{-45}$
double	$(2-2^{-52})\cdot 2^{1023}$ $\approx 1.8\cdot 10^{308}$	2^{-1022} $\approx 2.2\cdot 10^{-308}$	$2^{-1022-52}=2^{-1074}$ $\approx 4.9\cdot 10^{324}$

Representation of real numbers (1/3)

Let $fl : \mathbb{R} \to F$ be the mapping which assigns to any $x \in \mathbb{R}$ the closest machine number.

The "closest" is given by the method chosen: **rounding** ("ties to even"), **truncation** (rounding towards 0),...

When trying to represent a number which is out of the representable range, an **overflow** or **underflow** is returned.

Definition

Let a number α be an approximate value of a number a.

- *The* **absolute error** *is the value $|\alpha - a|$.*
- *For $a \neq 0$, the* **relative error** *is $\dfrac{|\alpha - a|}{|a|}$.*

Representation of real numbers (2/3)

In single precision, suppose that a number $x \in \mathbb{R}$ lies in the normalized range, i.e.,

$$x = q \cdot 2^{\ell}, \quad \text{where } 1 \leq q < 2 \text{ and } -126 \leq \ell \leq 127.$$

Let's **round towards** 0, i.e., chop off bits which do not fit into the significand (for positive numbers). If

$$x = (1.b_1 b_2 b_3 b_4 \ldots)_2 \cdot 2^{\ell},$$

then

$$fl(x) = (1.b_1 b_2 \ldots b_{23}) \cdot 2^{\ell}.$$

The absolute error is

$$|x - fl(x)| \leq 2^{-23+\ell}$$

and the relative error is

$$\frac{|x - fl(x)|}{|x|} \leq \frac{2^{-23+\ell}}{q \cdot 2^{\ell}} \leq 2^{-23}.$$

Representation of real numbers (3/3)

This threshold of relative error is called the **unit roundoff error** and is denoted by $\mathbf{u}$, i.e., in the single precision with chopping we have $\mathbf{u} = 2^{-23}$.

Proposition

Let $x \in \mathbb{R}$ be greater than the smallest normalized number of F and smaller than the greatest normalized number of F. We have

$$fl(x) = x(1+\delta), \quad \text{where } |\delta| \leq \mathbf{u},$$

Arithmetic operations - errors

Proposition

Let $x, y \in F$ and $\odot$ be the operation of addition, multiplication or division. If there is no overflow or underflow, then we have

$$fl(x \odot y) = (x \odot y)(1+\delta), \quad \text{where } |\delta| \leq \mathbf{u},$$

In general: If we operate with more numbers, it is better to start with the smallest ones.