Assumed to be known:

Gaussian elimination for solving linear systems of equations
matrix operations (including inversion)
determinant

Vector spaces, linear maps and matrices

Examples:

geometrical vectors of $\mathbb{R}^3$,
$\mathbb{R}^n = \{(x_1, \dots, x_n) \,|\, x_i \in \mathbb{R} \ \forall i\}$,
$\mathbb{R}^{n \times m}$: $n \times m$ real matrices,
$\mathbb{R}[x]$: polinomials with real coefficients,
$\mathbb{C}[x]$: polinomials with complex coefficients,
$C[0, 1]$: continuous real functions defined on $[0, 1]$, etc.

V is a **vector space** over the field K

vectors: $\mathbf{u}, \mathbf{v}, \dots \in V$,
scalars: $x, y, \alpha, \beta, \dots, \lambda, \dots \in K$,
operations: $\mathbf{u} + \mathbf{v} \in V$, $\lambda \mathbf{v} \in V$, $\mathbf{0} \in V$
identities:

$$\begin{aligned}
&\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u} && \lambda(\mathbf{u}+\mathbf{v})=\lambda\mathbf{u}+\lambda\mathbf{v}\\
&(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w}) && (\lambda+\mu)\mathbf{v}=\lambda\mathbf{v}+\mu\mathbf{v}\\
&\mathbf{v}+\mathbf{0}=\mathbf{v} && (\lambda\mu)\mathbf{v}=\lambda(\mu\mathbf{v})\\
& && 1\mathbf{v}=\mathbf{v}, \quad 0\mathbf{v}=\mathbf{0}
\end{aligned}$$

K may be $\mathbb{R}$, $\mathbb{C}$, or other subfields of $\mathbb{C}$, or finite fields, e.g. for a prime p

$\mathbb{F}_p = \{0, 1, \dots, p-1\}$, $+, \cdot$ modulo p.

Important: here $\alpha + \ldots + \alpha = n\alpha = 0$, if $p \mid n$,
$(\alpha + \beta)^p = \alpha^p + \beta^p$ (from the binomial theorem)

subspace: nonempty subset of V which is closed under the operations,
notation: $W \leq V$ means that W is a subspace of V

e.g. the subspaces of $\mathbb{R}^3$ are: the origin, lines and planes containing the origin, and the whole $\mathbb{R}^3$
$\mathbb{R}[x] \geq \mathbb{R}[x]_{\leq n}$: real polynomials of degree $\leq n$

spanned subspace: the smallest subspace containing a given subset S

= the intersection of all the subspaces containing S
= the set of linear combinations of the elements of S, i.e.
$\{\sum \lambda_i \mathbf{v}_i \,|\, \mathbf{v}_i \in S,\ \lambda_i \in K\} =: \operatorname{span} S$

spanning set $\mathcal{S}$: spans the whole vector space, i.e. $\forall$ vector can be expressed as a linear combination of some elements of $\mathcal{S}$

linearly independent set $\mathcal{U} = \{\mathbf{u}_i \,|\, i \in I\}$: $\sum \lambda_i \mathbf{u}_i = \mathbf{0} \Rightarrow \lambda_i = 0 \ \forall i$, i.e.

any vector in the spanned subspace can be written uniquely as a linear combination of elements from $\mathcal{U}$

(How do we check if a set of vectors in K^n is a spanning set, or if it is an independent set?)

basis: independent spanning set
= maximal independent set (no new elements can be added)
= minimal spanning set (no elements can be dropped)

$\forall$ independent set can be completed to a basis,
$\forall$ spanning set can be reduced to a basis

dimension the number of elements in a basis (well defined!)
The vector spaces in this course will be finite dimensional.

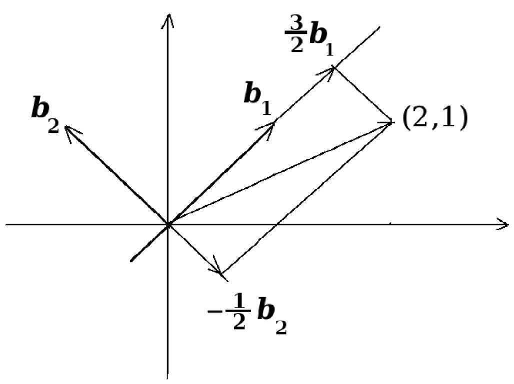
The following are equivalent for a set of vectors $\mathcal{B}$ in an n-dimensional space:
(*i*) $\mathcal{B}$ is a basis
(*ii*) $|\mathcal{B}| = n$, and $\mathcal{B}$ independent
(*iii*) $|\mathcal{B}| = n$, and $\mathcal{B}$ is a spanning set.

Example. A basis (the standard basis) of $\mathbb{R}^{2\times 2}$ is $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$, the standard basis of $\mathbb{C}_{\mathbb{R}}$ is $\{1, i\}$.

In an n-dimensional space with a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ (here the order of the elements is also important!), every vector can be uniquely written in the form $\sum_{i=1}^{n} x_i \mathbf{b}_i$. This defines the coordinatization with respect to $\mathcal{B}$: the **coordinate vector** of $\mathbf{v} = \sum x_i \mathbf{b}_i$ is

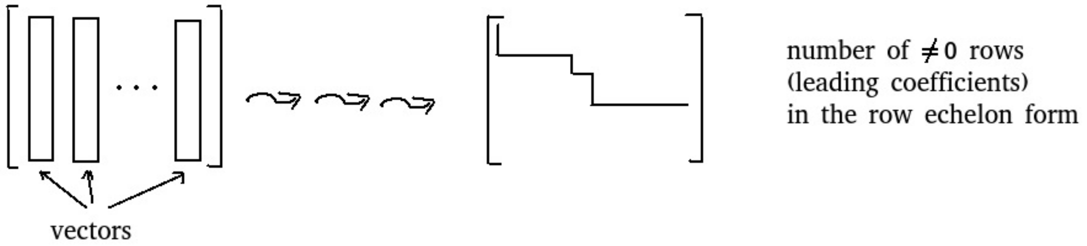
$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (x_1, \ldots, x_n)^T$$

Example. In $\mathbb{R}^2$, what is $[(2,1)]_{\mathcal{B}}$ with respect to the basis $\mathcal{B} = \{(1,1), (-1,1)\}$?



$$[(2,1)]_{\mathcal{B}} = \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix}.$$

rank (of a set of vectors): the dimension of the generated subspace.
calculation using Gauss elimination:



rank of a matrix: the dimension of the column space = the dimension of the row space

linear map: $f : V \to W$ (V and W are vector spaces over K), which satisfies

$$f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$$
$$f(\lambda\mathbf{v}) = \lambda f(\mathbf{v})$$

Example: congruences of $\mathbb{R}^3$ fixing $\mathbf{0}$, differentiation in $\mathbb{R}[x]$.

linear transformation: linear map with $V = W$

matrix of a linear map:

$$f : V \to W$$

bases: $\mathcal{B}$ $\quad$ $\mathcal{C}$

We need a matrix A such that $f : \mathbf{v} \mapsto \mathbf{w}$ if and only if $A \cdot [\mathbf{v}]_{\mathcal{B}} = [\mathbf{w}]_{\mathcal{C}}$.
$\exists!$ such a matrix for $\mathcal{B}$ and $\mathcal{C}$:

$$A = [f]_{\mathcal{B},\mathcal{C}} = \Big[\, [f(\mathbf{b}_1)]_{\mathcal{C}} \;\Big|\; \dots \;\Big|\; [f(\mathbf{b}_n)]_{\mathcal{C}} \,\Big]$$

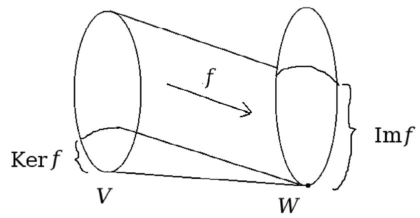
matrix of a linear transformation: usually $\mathcal{C} = \mathcal{B}$, and
$[f]_{\mathcal{B}} := [f]_{\mathcal{B},\mathcal{B}}$

Exercise: Determine the matrix of $z \to \overline{z}$ in $\mathbb{C}_{\mathbb{R}}$ in the basis $\{1, i\}$, or $\{i, 1+i\}$!

Sol.: $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, or $\begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}$, respectively

image: $\operatorname{Im} f = \{ f(\mathbf{v}) \mid \mathbf{v} \in V \} \le W$

kernel: $\operatorname{Ker} f = \{ \mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{0} \} \le V$



Change of basis

Let $\mathcal{B} = \{ \mathbf{b}_1, \dots, \mathbf{b}_n \}$ and $\mathcal{B}' = \{ \mathbf{b}'_1, \dots, \mathbf{b}'_n \}$ be two bases in V. $P := \big[[\mathbf{b}'_1]_{\mathcal{B}} \big| \dots \big| [\mathbf{b}'_n]_{\mathcal{B}}\big]$ is the transition matrix. Then

$[\mathbf{v}]_{\mathcal{B}} = P[\mathbf{v}]_{\mathcal{B}'}$, i.e. $P = [id]_{\mathcal{B}',\mathcal{B}}$, and
$P^{-1}[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}'}$.

Exercise: (a new method for an earlier problem) Determine the coordinate vector of $(2, 1)$ with respect to the basis $\{(1, 1), (-1, 1)\}$. This means that we change the standard basis

$\mathcal{B} = \{(1,0),(0,1)\}$ to the new basis $\mathcal{B}' = \{(1,1),(-1,1)\}$.
The transition matrix is $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

$$[P|I] = \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right] \mapsto \left[\begin{array}{cc|cc} 1 & -1 & 1 & 0 \\ 0 & 2 & -1 & 1 \end{array}\right] \mapsto \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \end{array}\right] = [I|P^{-1}].$$

$$[(2,1)]_{\mathcal{B}'} = P^{-1}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix}$$

The matrix of a linear map with respect to a new pair of bases

Let the transition matrices from $\mathcal{B}$ to $\mathcal{B}'$ and from $\mathcal{C}$ to $\mathcal{C}'$ be P and Q, respectively, $[f]_{\mathcal{B},\mathcal{C}} = A$ and $[f]_{\mathcal{B}',\mathcal{C}'} = A'$.
Then $A' = Q^{-1}AP$:

$$[f(\mathbf{v})]_{\mathcal{C}'} \xleftarrow{Q^{-1}} [f(\mathbf{v})]_{\mathcal{C}} \xleftarrow{A} [\mathbf{v}]_{\mathcal{B}} \xleftarrow{P} [\mathbf{v}]_{\mathcal{B}'}$$

The matrix of a linear transformation with respect to a new basis

$\mathcal{B}$, $\mathcal{B}'$ are two bases of V, $f : V \to V$ a linear transformation, $[f]_{\mathcal{B}} = A$, $[f]_{\mathcal{B}'} = A'$, and P the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$.
Then $A' = P^{-1}AP$.

Exercise: The matrix of the linear transformation $z \mapsto \overline{z}$ of $\mathbb{C}_{\mathbb{R}}$ with respect to the standard basis $\mathcal{B} = \{1, i\}$ is $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. What is the matrix of the transformation with respect to the basis $\mathcal{B}' = \{i, 1+i\}$?
The transition matrix is $P = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$, and the matrix of the transformation with respect to the new basis is $A' = P^{-1}AP = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}$.

Definition. $A, B \in K^{n \times n}$ are **similar** (notation: $A \sim B$), if there is an invertible matrix P such that $B = P^{-1}AP$. In other words: A and B are the matrices of the same linear transformations in two bases (the columns of P give the new basis coordinatized in the old basis).

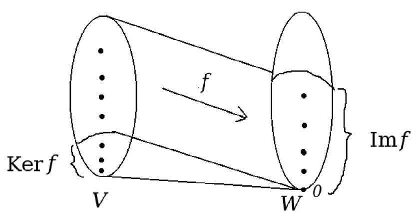
f **injective** if $\operatorname{Ker} f = \{\mathbf{0}\} =: 0$

f **surjective** if $\operatorname{Im} f = W$

f **isomorphism** if f injective and surjective.

Dimension theorem. Let $\dim V = n$ and $f : V \to W$ be linear. Then

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = n$$



Cor.: If $f : V \to V$ and $\dim V = n$ then
f iso. $\Leftrightarrow$ f inj. $\Leftrightarrow$ f surj.

Example: the coordinatization is an isomorphism: for $|\mathcal{B}| = n$
$V \to K^n$
$\mathbf{v} \mapsto [\mathbf{v}]_{\mathcal{B}}$

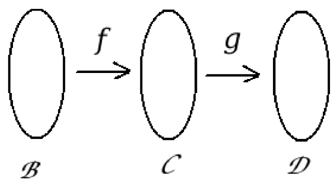
Theorem: Any map from the basis of a vector space to another vector space can be extended uniquely to a linear map.

rank of a linear map: $\operatorname{rank} f = \dim \operatorname{Im} f = \operatorname{rank}[f]_{\mathcal{B},\mathcal{C}}$ for any pair of bases $\mathcal{B}, \mathcal{C}$

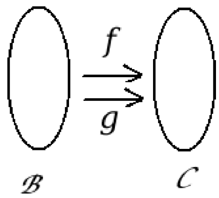
It follows from the Dimension Theorem that $\operatorname{rank} f = \dim V - \dim \operatorname{Ker} f$.
For $\dim V = n$, a linear map $f : V \to V$ is an isomorphism $\Leftrightarrow$ $\operatorname{rank} f = n$.

Matrix operations and linear maps:

$$[g]_{\mathcal{C},\mathcal{D}} \cdot [f]_{\mathcal{B},\mathcal{C}} = [g \circ f]_{\mathcal{B},\mathcal{D}}, \text{ where } (g \circ f)\mathbf{v} := g(f(\mathbf{v}))$$



$$[f]_{\mathcal{B},\mathcal{C}} + [g]_{\mathcal{B},\mathcal{C}} = [f + g]_{\mathcal{B},\mathcal{C}}, \text{ where } (f + g)(\mathbf{v}) := f(\mathbf{v}) + g(\mathbf{v})$$



The **rank** of a matrix A is the rank of the map $\mathbf{x} \mapsto A\mathbf{x}$.

Proposition. *For the matrices A, B*
1) $\operatorname{rank}(AB) \le \min\{\operatorname{rank} A, \operatorname{rank} B\}$
2) $|\operatorname{rank} A - \operatorname{rank} B| \le \operatorname{rank}(A + B) \le \operatorname{rank} A + \operatorname{rank} B$

Proof. Use the linear maps defined by the matrices.

Theorem (The rank of a matrix). *For $A \in K^{m\times n}$ the following are equivalent:*
(i) $\operatorname{rank} A = r$*;*
(ii) *the rank of* $\mathbf{x} \mapsto A\mathbf{x}$ *is r;*
(iii) *the column space of A is r-dimensional;*
(iv) *the row space of A is r-dimensional;*
(v) *in the row echelon form of A there are exactly r nonzero rows (i.e. there are r leading coefficients);*
(vi) *A contains an $r \times r$ submatrix with nonzero determinant but all its $(r+1)\times(r+1)$ submatrices have zero determinant.*

Theorem (Invertible matrices). *For $A \in K^{n\times n}$ the following are equivalent:*

(*i*) *A is invertible;*
(*ii*) *$f : K^n \to K^n$, $f : \mathbf{x} \mapsto A\mathbf{x}$ is an isomorphism;*
(*iii*) *$|A| \neq 0$;*
(*iv*) *the reduced row echelon form of A is I;*
(*v*) *$\operatorname{rank} A = n$;*
(*vi*) *the system of equations $A\mathbf{x} = \mathbf{b}$ has a solution for any $\mathbf{b} \in K^n$;*
(*vii*) *the system of equations $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*

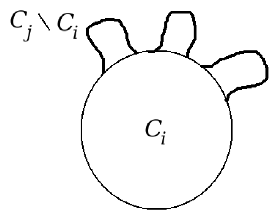
Calculating the inverse by Gaussian elimination:

$$[A|I] \mapsto\mapsto\mapsto [I|A^{-1}].$$

An application: Fisher's inequality

Theorem. Ⓟ *Let $C_1, \dots, C_k \subseteq \{1, \dots, n\}$ be distinct sets. Suppose that there is a $\lambda > 0$ such that $|C_i \cap C_j| = \lambda$ ($\forall i \neq j$). Then $k \leq n$.*

Proof. Case 1: $\exists i$: $|C_i| = \lambda$. Then:



$\Rightarrow n \geq |C_i| + (k-1) \geq k$.

Case 2: $\forall i$ $|C_i| = \lambda + a_i$, $a_i > 0$. The characteristic vector of $X \subseteq \{1, \dots, n\}$ is the n dimensional 0-1-vector, $(x_1, \dots, x_n)$, where $x_i = 1 \Leftrightarrow i \in X$. Let $M \in \mathbb{R}^{k\times n}$ the matrix whose ith row is the characteristic vector of the set C_i. Then

$$A = MM^T = \begin{bmatrix} \lambda + a_1 & \lambda & \lambda & \dots & \lambda \\ \lambda & \lambda + a_2 & \lambda & \dots & \lambda \\ \vdots & & \ddots & & \\ & & & & \lambda + a_n \end{bmatrix}_{k\times k}, \text{ since } \mathbf{x}\cdot\mathbf{y} = |X \cap Y|$$

We know: $\operatorname{rank} A \leq \operatorname{rank} M \leq n$.
We will show: $|A| \neq 0$, so $\operatorname{rank} A = k$.

$$|A| = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & \lambda + a_1 & \lambda & \dots & \lambda \\ 0 & \lambda & \lambda + a_2 & \dots & \lambda \\ \vdots & \vdots & & \ddots & \\ 0 & \lambda & \dots & & \lambda + a_n \end{vmatrix}_{(k+1)\times(k+1)} = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ -\lambda & a_1 & 0 & \dots & 0 \\ -\lambda & 0 & a_2 & \dots & 0 \\ \vdots & & & \ddots & \\ -\lambda & 0 & 0 & \dots & a_n \end{vmatrix} =$$

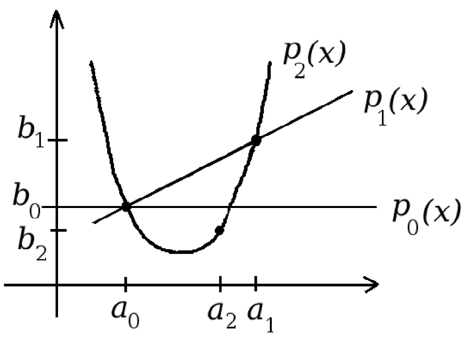
$$\begin{vmatrix} 1+\frac{\lambda}{a_1}+\ldots+\frac{\lambda}{a_n} & 1 & 1 & \ldots & 1 \\ 0 & a_1 & 0 & \ldots & 0 \\ 0 & 0 & a_2 & \ldots & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & 0 & \ldots & a_n \end{vmatrix} = (1+\frac{\lambda}{a_1}+\ldots+\frac{\lambda}{a_n})\cdot a_1\cdots a_n > 0,$$

since $\lambda, a_1, \ldots, a_n > 0$. □

Polynomial interpolation

Ⓟ K is a field, $a_0, \ldots, a_n, b_0, \ldots, b_n \in K$, $a_0, \ldots, a_n$ are pairwise different $\Rightarrow$

$$\exists! p(x) \in K[x]_{\leq n}: \ p(a_i) = b_i \ \forall i.$$



Proof. $f : K[x]_{\leq n} \to K^{n+1}$, $f : p(x) \mapsto \begin{bmatrix} p(a_0) \\ \vdots \\ p(a_n) \end{bmatrix}$ is a linear map. $\operatorname{Ker} f = 0$, since if $p(x) \in \operatorname{Ker} f \Rightarrow p(a_0) = \cdots = p(a_n) = 0 \Rightarrow p(x) = (x-a_0)\cdots(x-a_n)q(x)$, but $\deg p \leq n$, so $p(x) = 0$. $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim K[x]_{\leq n} = n+1$ implies $\dim \operatorname{Im} f = n+1$, that is, f is surjective, and by $\operatorname{Ker} f = 0$ it is also injective, consequently, f is an isomorphism. This means that for any $\mathbf{b} = \begin{bmatrix} b_0 \\ \vdots \\ b_n \end{bmatrix}$ there is exactly one $p(x) \in K[x]_{\leq n}$ such that $f(p(x)) = \mathbf{b}$.

□

Newton's method of interpolation (see also the Lagrange polynomials)

For the given $a_0, \ldots, a_n, b_0, \ldots, b_n$ let $p_i(x) \in K[x]_{\leq i}$ be an interpolating polynomial on $a_0, \ldots, a_i$. Clearly, $p_0(x) \equiv b_0$. If p_i is given, then

$$p_{i+1}(x) = p_i(x) + A \cdot (x-a_0)\cdots(x-a_i)$$

has the same values up to a_i for any $A \in K$, and $\deg p_{i+1}(x) \leq i+1$. Furthermore, A can be chosen so that $p_{i+1}(a_{i+1}) = b_{i+1}$ (if we substitute a_{i+1}, the coefficient of A is not 0, since all the a_j's are different). So in the end we find a suitable $p_n(x)$.

Remark: Using Newton's method, it is easy to improve an interpolation by adding new points, i.e. measuring the value of the function which we wish to approximate by a polynomial at a few more places.

Shamir's secret sharing

We want to share a secret between n people (let the secret be coded by a natural number c) so that any k of the n people together can find out the secret information, but no $k-1$ of them could get closer to the secret if they share their bit of information among them.

Solution: Let $p > c$ be a prime, $q(x) \in \mathbb{F}_p[x]_{<k}$, such that $q(0) = c$ (that is, c is the constant term). The i.'th person is given the value $q(i) \in \mathbb{F}_p$ $(i = 1, \ldots, n)$. Then k people together know k values of the polynomial, so by the interpolation theorem they can determine the polynomial and then also its constant term. But if someone knows only $k-1$ values of the polynomial, then $q(0)$ can still be anything: we can still find such an interpolating polynomial of degree less than k.

Question: Why do we need a polynomial over a finite field $\mathbb{F}_p$? Why do not we choose an integral polynomial? Because in that case it is not true that with given $k-1$ values, $q(0)$ can be anything. It is possible that, though we find an interpolating polynomial over $\mathbb{Q}$, the coefficients of that polynomial are not integers.