ZÁKLADY ALGEBRY
(ÚPRAVA PRO ZS 2016/17)

DAVID STANOVSKÝ

stanovsk@karlin.mff.cuni.cz

Everything should be made as simple as possible, but not one bit simpler.
— Albert Einstein

Date: 19. prosince 2016.

Obsah

Úvod

1. Stručná historie moderní algebry

Význam slova algebra, pocházejícího z arabštiny, se v průběhu staletí měnil, od řešení jednoduchých rovnic a symbolické manipulace s výrazy, až po současné studium abstraktních struktur, jako např. grup, okruhů, či vektorových prostorů. Kořeny algebry sahají, jako u většiny matematických disciplín, do starověku, ve své dnešní podobě se algebra vyvíjí přibližně od počátku 19. století. Na úvod si ukážeme několik problémů z jiných disciplín, které motivovaly vznik abstraktní algebry jako samostatného oboru.

Jedním z takových problémů bylo *hledání kořenů celočíselných polynomů*. Řešení kvadratických rovnic lze vystopovat až do starověku a vzorec v téměř moderní podobě byl podán Al-Chorezmím už v 9. století (ukázku z této knihy najdete na úvodní stránce skript). Návody na řešení rovnic třetího a čtvrtého stupně, tzv. Cardanovy vzorce, pocházejí z doby italské renesance. Hledání vzorců pro řešení rovnic pátého a vyššího stupně pak zaměstnávalo matematiky dalších 300 let, než byla nalezena konečná odpověď: žádné takové neexistují. Práce Ruffiniho, Abela a Galoise lze považovat za první vážné algebraické výsledky. (Podrobněji o tomto tématu viz Sekce **??**.)

Snad ještě starším problémem jsou *konstrukce pravítkem a kružítkem*. Staří Řekové uměli tímto způsobem vyřešit řadu geometrických úloh, některé však řešení vzdorovaly: mezi nejznámější patří rektifikace kružnice, kvadratura kruhu či trisekce úhlu. Vzdorovaly dalších téměř 2000 let. Až v první polovině 19. století Pierre Wantzel dokázal, použitím algebraických metod, že tyto úlohy jsou pravítkem a kružítkem neřešitelné. Šlo o další velký úspěch, který formoval základy algebry. (Podrobněji o tomto tématu viz Sekce 13.)

Posledním prastarým problémem, který zmíníme, je *řešení diofantických rovnic*, tj. polynomiálních rovnic v oboru celých čísel. Nejznámější diofantickou rovnicí je zřejmě Velká Fermatova věta, která říká, že rovnice $x^n + y^n = z^n$ nemá pro $n > 2$ netriviální celočíselné řešení. Jednou z metod, která slavila částečné úspěchy v 19. století, bylo užití aritmetiky v komplexních rozšířeních oboru celých čísel. Tato metoda sice nakonec k cíli nevedla, ale algebra se podílela i na konečném řešení tohoto problému; Velká Fermatova věta byla dokázána Andrew Wilesem na konci 20. století za pomoci algebraické geometrie. (Podrobněji o tomto tématu viz Sekce 9.3.)

Snad jsme v tuto chvíli čtenáře aspoň částečně přesvědčili, že algebra je užitečná věc a nachází uplatnění v mnoha matematických disciplínách; dokladem budiž obory jako algebraická teorie čísel, algebraická geometrie, algebraická topologie atd. Aplikace algebry však sahají daleko za hranice matematiky. Lieova algebra stojí za nemalou částí teoretické fyziky, algebraické metody využívá teoretická informatika (např. teorie pologrup je základním nástrojem teorie automatů) a v poslední

době se objevily např. aplikace výpočetní algebry v biologii. V neposlední řadě, algebra je důležitým nástrojem moderní kryptografie a stěžejní ingrediencí teorie samoopravných kódů.

Vývoj algebry lze zhruba rozdělit na dvě fáze: *klasická algebra*, cca od začátku 19. století do poloviny 20. století, se zabývala počítáním v konkrétních oborech (číselné obory, vektory, matice, permutace atd.). Motivací bylo zpravidla řešení problémů jiných matematických disciplín, které vedlo na výpočty v netradičních oborech. Na přelomu století, pod vlivem Davida Hilberta a řady dalších matematiků, docházelo ke změně pohledu na celou matematiku, algebru nevyjímaje. Docházelo k formalizaci jednotlivých disciplín a důsledkem byl mj. abstraktní, axiomatický, přístup k algebře, kdy konkrétní obory byly nahrazeny abstraktními strukturami (grupy, okruhy, vektorové prostory atd.). Za přelom lze označit učebnici Bartela van der Waerdena z roku 1930 nazvanou *Moderne Algebra*, která zafixovala současnou terminologii a postulovala základy toho, co dnes nazýváme moderní algebrou. Osnova většiny učebnic algebry se od van der Waerdenových dob výrazně nezměnila, a ani tato skripta nebudou výjimkou. Co se mění je styl výkladu této látky.

2. Uspořádání a jiné pomocné pojmy

Cíl. *Připomeneme několik pojmů, které by měly být známy z úvodních matematických kurzů, především uspořádané množiny.*

2.1. Uspořádané množiny.

Definice. Relaci $\leq$ na množině X nazýváme *částečné uspořádání*, pokud je

(1) *reflexivní*, tj. $x \leq x$ pro všechna $x \in X$,
(2) *tranzitivní*, tj. $x \leq y$ a $y \leq z$ implikuje $x \leq z$,
(3) a *antisymetrická*, tj. $x \leq y$ a $y \leq x$ implikuje $x = y$.

Alternativně říkáme, že $(X, \leq)$ je *uspořádaná množina*. Uspořádání se nazývá *lineární*, pokud navíc pro každé x, y nastane $x \leq y$ nebo $y \leq x$. *Intervalem* rozumíme množinu

$$[a, b] = \{x \in X : \ a \leq x \leq b\}.$$

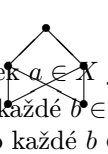
Pokud $x \leq y$ a $x \neq y$, píšeme $x < y$.

Příklady.

- Na množině přirozených čísel uvažujme obvyklé uspořádání $1 < 2 < 3 < \ldots$; uspořádaná množina $(\mathbb{N}, \leq)$ je lineární.
- Na množině přirozených čísel uvažujme uspořádání dělitelnosti, tj. „a je menší než b pokud $a \mid b$“; uspořádaná množina $(\mathbb{N}, |)$ *není* lineární: např. čísla 2, 3 jsou neporovnatelné.
- Na množině $P(X)$ všech podmnožin dané množiny X uvažujme uspořádání inkluzí, tj. „A je menší než B pokud $A \subset B$“; je-li $|X| > 1$, pak uspořádaná množina $(P(X), \subseteq)$ není lineární: např. dvě různé jednoprvkové množiny jsou neporovnatelné.

Konečné uspořádané množiny se často zadávají pomocí tzv. *Hasseova diagramu*. Jde o graf relace $\leq$, přičemž nekreslíme smyčky (reflexivita), vynecháváme všechny hrany, jejichž existence je zaručena tranzitivitou, a místo šipek kreslíme neorientované hrany tak, aby větší prvky byly výše. Např.

Definice. Řekneme, že prvek $a \in X$ je v $(X, \leq)$



- *největší*, pokud pro každé $b \in X$ platí $b \leq a$;
- *nejmenší*, pokud pro každé $b \in X$ platí $b \geq a$;
- *maximální*, pokud neexistuje žádné $b \in X$ takové, že $b > a$;
- *minimální*, pokud neexistuje žádné $b \in X$ takové, že $b < a$.

Příklady.

- Uspořádaná množina **A** má jeden největší prvek, jeden maximální (ten samý), žádný nejmenší a dva minimální prvky.
- Uspořádaná množina **B** má jeden největší (a zároveň maximální) a jeden nejmenší (a zároveň minimální) prvek. Je to lineární uspořádání.
- Uspořádaná množina $(\mathbb{N}, \leq)$ má nejmenší prvek 1, ale žádný maximální prvek.
- Uspořádaná množina $(\mathbb{N}, |)$ přirozených čísel s relací dělitelnosti má nejmenší prvek 1, ale žádný maximální prvek. Uspořádaná množina $(\mathbb{N} \setminus \{1\}, |)$ má za minimální prvky právě všechna prvočísla.

Definice. Nechť $Y \subseteq X$. Řekneme, že prvek $a \in X$ je v $(X, \leq)$

- *horní mez* množiny Y, pokud $a \geq y$ pro každý prvek $y \in Y$;
- *supremum* množiny Y, pokud to je nejmenší horní mez Y; značíme jej $a = \sup Y$;
- *dolní mez* množiny Y, pokud $a \leq y$ pro každý prvek $y \in Y$;
- *infimum* množiny Y, pokud to je největší dolní mez Y; značí se $a = \inf Y$.

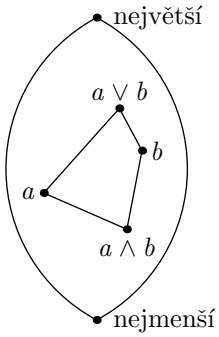
Jinými slovy, supremum množiny Y je nejmenší prvek množiny X, který je větší nebo rovný než všechny prvky Y. Podobně, infimum množiny Y je největší prvek množiny X, který je menší nebo rovný než všechny prvky Y.

Příklady.

- V uspořádané množině **A** podmnožina sestávající z obou minimálních prvků nemá supremum ani infimum. Infimum proto, že nemá ani žádnou dolní mez. Horní meze sice tato podmnožina má tři, avšak žádná z nich není nejmenší.
- V uspořádané množině **B** má každá neprázdná podmnožina supremum i infimum. Obecně, v každé lineárně uspořádané množině má každá neprázdná *konečná* podmnožina supremum i infimum, přičemž $\sup Y = \max Y$ a $\inf Y = \min Y$. Pozor, pro nekonečné to obecně nefunguje: např. v $(\mathbb{N}, \leq)$ neexistuje $\sup \mathbb{N}$.
- V uspořádané množině $(P(X), \subseteq)$ má každá podmnožina infimum i supremum, přičemž $\inf Y$ je rovno průniku všech množin z Y a $\sup Y$ je rovno sjednocení všech množin z Y.
- V uspořádané množině $(\mathbb{N}, |)$ má každá konečná podmnožina infimum i supremum. Přitom $\inf Y$ je rovno NSD všech čísel z Y a $\sup Y$ je rovno NSN všech čísel z Y. Na druhou stranu, např. $\sup\{p : p \text{ prvočíslo}\}$ neexistuje.

Uvědomte si, že $\sup \emptyset$ je rovno nejmenšímu prvku, pokud takový v $(X, \leq)$ existuje; podobně, $\inf \emptyset$ je rovno největšímu prvku, pokud takový existuje.

Definice. Uspořádanou množinu nazveme *svazově uspořádanou*, pokud v ní existují suprema a infima všech *dvouprvkových* podmnožin (pak také zřejmě existují



Obrázek 1. Průsek a spojení ve svazově uspořádané množině

suprema a infima všech *neprázdných konečných* podmnožin). Nazveme ji *úplně svazově uspořádanou*, pokud existují suprema a infima všech podmnožin. Ve svazově uspořádaných množinách obvykle značíme zkráceně

$$a \vee b = \sup\{a, b\} \quad \text{a} \quad a \wedge b = \inf\{a, b\},$$

symboly $\vee, \wedge$ čteme jako *spojení* a *průsek*.

Z definice plyne, že v úplně svazově uspořádané množině existuje nejmenší i největší prvek, jsou jimi $\sup \emptyset$, resp. $\inf \emptyset$.

Příklady.

- Uspořádaná množina **A** není svazově uspořádaná.
- Lineárně uspořádáná množina je vždy svazově uspořádaná, přičemž $a \vee b = \max(a, b)$, $a \wedge b = \min(a, b)$.
- $(\mathbb{N}, \leq)$ je svazově uspořádaná množina, ale ne úplně: např. $\sup \mathbb{N}$ neexistuje. $(\mathbb{N} \cup \{\infty\}, \leq)$ je úplně svazově uspořádaná množina.
- $(P(X), \subseteq)$ je úplně svazově uspořádaná množina: pro $U \subseteq P(X)$, $\sup U = \bigcup_{A \in U} A$, $\inf U = \bigcap_{A \in U} A$.
- $(\mathbb{N}, |)$ je svazově uspořádaná množina (ne úplně): $a \vee b = \mathrm{NSN}(a, b)$, $a \wedge b = \mathrm{NSD}(a, b)$.

Definici úplně svazově uspořádané množiny lze zjednodušit: pokud v dané uspořádané množině existují všechna infima, pak existují i všechna suprema, a naopak.

Tvrzení 2.1. *Uspořádaná množina, ve které existují infima všech podmnožin, je úplně svazově uspořádaná.*

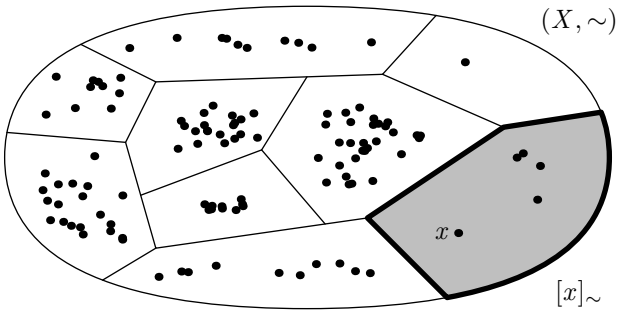
Analogicky lze předpokládat pouze existenci suprem.

Důkaz. Označme danou uspořádanou množinu $(X, \leq)$. Stačí si uvědomit, že

$$\sup Y = \inf\{a \in X : \ a \geq y \text{ pro každé } y \in Y\},$$

tedy že suprema lze definovat pomocí infim. □

V závěrečné kapitole budeme na jednom místě potřebovat *Zornovo lemma*. Jde o jednu z forem tzv. *axiomu výběru*, jednoho ze základních axiomů teorie množin. Zornovo lemma se proto nedokazuje, nýbrž postuluje. *Řetězcem* v částečně uspořádané množině rozumíme podmnožinu, která je lineárně uspořádaná.



Obrázek 2. Ekvivalence $\sim$ na množině X a její blok $[x]_\sim$.

Lemma 2.2 (Zornovo lemma)**.** *Buď $(X, \leq)$ neprázdná částečně uspořádaná množina. Předpokládejme, že každý řetězec má horní mez. Pak $(X, \leq)$ obsahuje aspoň jeden maximální prvek.*

2.2. **Ekvivalence.**

Relací ρ na množině X rozumíme libovolnou podmnožinu kartézského součinu $X \times X$; tedy prvky relace ρ jsou některé dvojice prvků množiny X. Místo $(a, b) \in \rho$ píšeme často $a \,\rho\, b$, zejména pokud relaci označíme symbolem typu $\sim$, $\leq$ apod. Na relace je užitečné nahlížet jako na orientované grafy.

Definice. Relaci $\sim$ na množině X nazýváme *ekvivalence*, pokud je

(1) *reflexivní*, tj. $x \sim x$ pro všechna $x \in X$,
(2) *tranzitivní*, tj. $x \sim y$ a $y \sim z$ implikuje $x \sim z$,
(3) a *symetrická*, tj. $x \sim y$ implikuje $y \sim x$.

Blokem (nebo *třídou*) ekvivalence $\sim$ příslušnou prvku $x \in X$ rozumíme množinu

$$[x]_\sim = \{y \in X : \ x \sim y\}.$$

Pro daná x, y jsou příslušné bloky $[x]_\sim, [y]_\sim$ buď stejné (pokud $x \sim y$), nebo disjunktní; tvoří tedy *rozklad* množiny X. Množinu všech bloků ekvivalence $\sim$ značíme $X/\!\sim$, tj. $X/\!\sim = \{[x]_\sim : \ x \in X\}$.

Naopak, každému disjunktnímu rozkladu $X = \bigcup_{B \in \mathcal{B}} B$ přísluší ekvivalence definovaná předpisem „$x \sim y \ \Leftrightarrow x, y$ leží ve stejném bloku“.

Příklady.

- Na množině přirozených čísel $\mathbb{N}$ zavedeme relaci definovanou předpisem „$a \sim b \ \Leftrightarrow \ a + b$ je sudé číslo“. Je to ekvivalence s dvěma bloky: jeden blok je tvořen sudými čísly, druhý lichými.
- Na množině všech přímek v rovině zavedeme relaci definovanou předpisem „$p_1 \parallel p_2 \ \Leftrightarrow$ přímky p_1 a p_2 jsou rovnoběžné“. Blok $[p]_\parallel$ obsahuje právě všechny přímky rovnoběžné s p.
- Na množině všech trojúhelníků v rovině zavedeme relaci definovanou předpisem „$T_1 \simeq T_2 \ \Leftrightarrow$ trojúhelníky T_1 a T_2 jsou shodné“. Blok $[T]_\simeq$ obsahuje právě všechny trojúhelníky shodné s T.
- Na množině vrcholů daného grafu zavedeme relaci definovanou předpisem „$x \sim y \ \Leftrightarrow$ existuje cesta z x do y“. Bloky této ekvivalence jsou komponenty souvislosti daného grafu.

Na závěr úvodní kapitoly zformulujeme jedno pozorování o konečných množinách, které nijak nesouvisí s uspořádánými množinami, avšak bude se nám v budoucnu párkrát hodit.

Lemma 2.3. *Buď $f : X \to Y$ zobrazení mezi stejně velkými konečnými množinami. Je-li f prosté, pak je bijektivní.*

Důkaz. Nechť $n = |X| = |Y|$. Každému z n prvků množiny X přiřadí f nějakou hodnotu, přičemž tyto hodnoty jsou navzájem různé; obor hodnot zobrazení f tedy musí mít n prvků. Takže to musí být celé Y. □

3. Elementární teorie čísel

> **Cíl.** *Nejprve stručně nastíníme, jak se formálně definují přirozená čísla, a hned poté se pustíme do základních poznatků o dělitelnosti: existence a jednoznačnost rozkladu na prvočísla (základní věta aritmetiky); Eukleidův algoritmus a Bézoutova rovnost; čínská věta o zbytcích; Eulerova funkce a Eulerova věta. Naučíme se pracovat s šikovným značením pomocí kongruencí $\equiv \pmod n$.*

KECY: algebra v kontextu teorie cisel - zaklady uvedeme ted, casem se nektere veci prirozene vysvetli algebraick

3.1. Přirozená a celá čísla.

Přirozenými čísly intuitivně rozumíme množinu $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$. Formálně vzato však tento zápis nedává valný smysl ze dvou důvodů: jednak nekonečnou množinu nemůžeme definovat výčtem prvků, a pak také není jasné, co vlastně symboly 1,2,3 atd. znamenají. V tomto odstavci nastíníme, jak lze přirozená čísla zavést formálně. Protože však u čtenáře nepředpokládáme žádnou znalost matematické logiky, nebudeme se pouštět do detailů a některé pojmy z logiky budeme používat bez dalšího vysvětlení na intuitivní úrovni. Z jistých důvodů se v logice zavádějí přirozená čísla i s nulou, čehož se v tomto odstavci přidržíme.

Jeden ze způsobů, jak přirozená čísla zavést, je zformulovat sadu *axiomů*, z nichž se budou všechna tvrzení o přirozených číslech dokazovat. Standardním přístupem je tzv. *Peanova axiomatika*. Přirozená čísla s nulou zavedeme jako teorii, v níž máme konstantu 0, unární funkční symbol s a následující axiomy:

(1) pro každé a existuje právě jedno b takové, že $s(a) = b$;
(2) pro každé a je $s(a) \neq 0$;
(3) pro každé $a \neq b$ platí $s(a) \neq s(b)$;
(4) je-li V vlastnost taková že
   (a) 0 má vlastnost V;
   (b) pro každé a platí následující: jestliže má a vlastnost V, pak $s(a)$ má také vlastnost V;

   pak má každé a vlastnost V.

Číslovky pak můžeme zavést jako $1 = s(0)$, $2 = s(1)$, atd. Interpretace symbolu s je taková, že „číslu“ přiřadí „číslo o jedna větší“. První tři axiomy říkají, že s je prostá funkce, v jejímž oboru hodnot není 0. Poslednímu axiomu se říká *matematická indukce*.

Na základě těchto axiomů můžeme induktivně definovat standardní operace: sčítání předpisy $a + 0 = a$ a $a + s(b) = s(a + b)$ (tj. umíme-li spočítat $a + b$, definujeme

na jeho základě $a + s(b)$), násobení předpisy $a \cdot 0 = 0$ a $a \cdot s(b) = a \cdot b + a$, atd. Uspořádání definujeme předpisem $a \leq b \Leftrightarrow \exists c\ a + c = b$ a podobně lze postupovat pro další známé pojmy a vlastnosti.

Z Peanových axiomů lze logicky odvodit všechna tvrzení o přirozených číslech, na která si vzpomenete – i když zpravidla nejde vůbec o jednoduchou práci (zkuste např. dokázat, že sčítání je komutativní!). Přesto má tato metoda své limity: slavná Gödelova věta o neúplnosti říká, že existují tvrzení, jež z těchto axiomů nelze dokázat ani vyvrátit. A ještě hůře: dokonce neexistuje žádná „hezká“ sada axiomů, která by tuto nepříjemnou vlastnost neměla. Naštěstí se ukazuje, že taková tvrzení jsou poměrně obskurní, Gödelovou větou se tedy nemusíme příliš trápit.

Druhým přístupem, který uvedeme, je vybudování *modelu* přirozených čísel (s nulou) v rámci nějaké dobře známé teorie, např. teorie množin. Standardním modelem v teorii množin jsou tzv. *von Neumannova čísla*, definovaná jako nejmenší množina ω splňující

(1) $\emptyset \in \omega$;
(2) jestliže $A \in \omega$, pak $A \cup \{A\} \in \omega$.

Tedy ω obsahuje postupně množiny

$$\emptyset, \quad \{\emptyset\}, \quad \{\emptyset, \{\emptyset\}\}, \quad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \quad \dots$$

Tímto způsobem můžeme definovat číslovky $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\emptyset, \{\emptyset\}\}$ atd. Všimněte si, že v tomto značení je $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, atd. Pokud iterpretujeme symbol s jako $s(A) = A \cup \{A\}$, pro von Neumannova čísla budou platit Peanovy axiomy.

Máme-li vybudována přirozená čísla, není těžké definovat, co jsou to čísla celá, racionální, reálná či komplexní. Celá čísla lze formálně definovat jako sjednocení dvou kopií přirozených čísel, z nichž jednu budeme označovat symboly $-a$, kde a je přirozené číslo (obě nuly ztotožníme). Operace a uspořádání se pak definují zřejmým způsobem. Celá čísla s operacemi sčítání, odčítání a násobení tvoří strukturu, které se říká *obor integrity* (viz sekce 4). Racionální čísla se pak definují jako *podílové těleso* tohoto oboru (viz Tvrzení 4.5). Způsobů, jak formálně zavést čísla reálná je celá řada, jeden příklad za všechny: jde o tzv. *zúplnění* uspořádaného tělesa racionálních čísel – doplníme suprema a infima všech omezených podmožin a pomocí limit na ně přeneseme operace (detaily konstrukce patří spíše do topologie). Na komplexní čísla pak lze nahlížet třeba jako na *algebraický uzávěr* čísel reálných (viz Věta **??**).

3.2. **Základní věta aritmetiky.**

Výklad teorie čísel začneme větou o prvočíselném rozkladu. Všechna fakta z tohoto odstavce byla známa již starořeckým matematikům a v moderní podobě byly formulovány Carlem Friedrichem Gaussem v jeho slavné knize *Disquisitiones Arithmeticae* z roku 1801, která položila základ moderní teorie čísel.

Buď a, b celá čísla. Řekneme, že číslo b *dělí* číslo a, píšeme $b \mid a$, pokud existuje celé číslo q splňující $a = b \cdot q$. Pro každé a platí $\pm 1 \mid a$ a $\pm a \mid a$; tito dělitelé se nazývají *nevlastní*.

Jak známo, pro každou dvojici celých čísel a, b, kde $b \neq 0$, existuje právě jedna dvojice celých čísel q, r splňující

$$a = q \cdot b + r \quad \text{a} \quad 0 \leq r < |b|.$$

Číslo q se nazývá *celočíselný podíl* čísel a, b, značí se a div b, a číslo r se nazývá *zbytek* po dělení, značí se a mod b. Existence podílu a zbytku plyne ze známého algoritmu celočíselného dělení, ale jak to je s jednoznačností? Kdyby $a = q_1 b + r_1 = q_2 b + r_2$, pak $b(q_1 - q_2) = r_2 - r_1$, tedy $b \mid r_2 - r_1$, avšak $0 \leq |r_2 - r_1| < |b|$, takže jedinou možností je případ $r_2 - r_1 = 0$. Z toho plyne $r_1 = r_2$ i $q_1 = q_2$.

Přirozené číslo $p \neq 1$, které má pouze nevlastní dělitele, se nazývá *prvočíslo*; ostatní přirozená čísla se nazývají *složená*. Zcela základním poznatkem teorie čísel je fakt, že každé číslo lze jednoznačně vyjádřit jako součin prvočísel.

Věta 3.1 (základní věta aritmetiky)**.** *Pro každé přirozené číslo $a \neq 1$ existují po dvou různá prvočísla $p_1, p_2, \ldots, p_n$ a přirozená čísla $k_1, k_2, \ldots, k_n$ splňující*

$$a = p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_n^{k_n}$$

(tomuto vyjádření se říká prvočíselný rozklad*). Tento zápis je jednoznačný až na pořadí činitelů.*

Přiznejme si však na tomto místě: kdo z nás umí takovou „samozřejmost“, jakou je existence a jednoznačnost prvočíselného rozkladu, dokázat?

Tedy existenci rozkladu lze dokázat poměrně snadno indukcí: je-li a prvočíslo, rozklad zřejmě existuje; budeme tedy předpokládat, že a je složené a že rozklad existuje pro všechna menší čísla. Napíšeme $a = b \cdot c$ pro nějaká $1 < b, c < a$. Podle indukčního předpokladu existuje prvočíselný rozklad jak pro b, tak pro c. Jejich složením získáme rozklad čísla a.

Jednoduchým důsledkem je např. fakt, že existuje nekonečně mnoho prvočísel. Kdyby jich bylo jenom konečně mnoho, označme je $p_1, \ldots, p_n$, uvažujme číslo $p_1 \cdot p_2 \cdot \ldots \cdot p_n + 1$. Toto číslo není dělitelné žádným prvočíslem, přitom musí mít nějaký prvočíselný rozklad. Spor.

S důkazem jednoznačnosti je to složitější.

Největší společný dělitel celých čísel a a b je největší přirozené číslo c splňující zároveň $c \mid a$ a $c \mid b$. Toto číslo značíme $\mathrm{NSD}(a, b)$. Podobně, *nejmenší společný násobek* čísel a a b je nejmenší číslo c splňující zároveň $a \mid c$ a $b \mid c$. Toto číslo značíme $\mathrm{NSN}(a, b)$. Je snadné nahlédnout, že

$$\mathrm{NSN}(a, b) = \frac{a \cdot b}{\mathrm{NSD}(a, b)}.$$

Jedna možnost jak počítat NSD je pomocí (jednoznačných) prvočíselných rozkladů: např. $168 = 2^3 \cdot 3 \cdot 7$ a $396 = 2^2 \cdot 3^2 \cdot 11$, a tak vidíme, že $\mathrm{NSD}(168, 396) = 2^2 \cdot 3 = 12$. Problém je, že kdybychom neměli jednoznačnost rozkladů, kdyby se např. číslo 396 rozkládalo na součin úplně jiných prvočísel než $2, 3, 11$, dostali bychom z jiného rozkladu jiný NSD, což je absurdní. Nejsme tedy schopni dokázat správnost této metody, a tímto způsobem ani existenci NSD, aniž bychom dokončili důkaz základní věty aritmetiky. (Skutečným protipříkladem, že tato metoda v obecnosti nefunguje, je např. následující situace v oboru $\mathbb{Z}[\sqrt{5}]$: zde $4 = 2 \cdot 2 = (\sqrt{5} - 1)(\sqrt{5} + 1)$. Z prvního rozkladu bychom vydedukovali $\mathrm{NSD}(2, 4) = 2$, z druhého $\mathrm{NSD}(2, 4) = 1$. Viz sekce 7.)

Druhou možností výpočtu NSD je *Eukleidův algoritmus*. Ten funguje následujícím způsobem: začneme s danými dvěma nezápornými čísly (evidentně $\mathrm{NSD}(a, b) = \mathrm{NSD}(-a, b)$, můžeme tedy na vstupu předpokládat nezáporná čísla) a budujeme posloupnost tak, že vždy vezmeme zbytek po dělení předposledního čísla posledním.

Odpovědí je poslední nenulová hodnota. Formálně, je-li na vstupu $a \geq b \geq 0$, pak inicializujeme $a_0 = a$, $a_1 = b$ a budujeme posloupnost předpisem

$$a_{i+1} = a_{i-1} \bmod a_i.$$

Pokud vyjde $a_{i+1} = 0$, odpovědí je a_i. Např. pro $\mathrm{NSD}(168, 396)$ dostáváme posloupnost 396, 168, 60, 48, 12, 0, a tedy $\mathrm{NSD}(168, 396) = 12$. Správnost algoritmu plyne z následujícího pozorování, nezávislého na základní větě aritmetiky. Eukleidův algoritmus tak prokazuje existenci NSD pro každou dvojici přirozených čísel.

Lemma 3.2. *Pro libovolná celá čísla a, b platí*

$$\mathrm{NSD}(a, b) = \mathrm{NSD}(b, a \bmod b).$$

Důkaz. Zopakujme, že

$$a = b \cdot (a \operatorname{div} b) + (a \bmod b).$$

Tedy dané číslo c dělí obě čísla a, b právě tehdy, když c dělí obě čísla $b, a \bmod b$ (což je snadno vidět, pokud si rozpíšeme $a \bmod b = a - b \cdot (a \operatorname{div} b)$). Protože tyto dvě dvojice mají stejné společné dělitele, mají stejného i toho největšího. □

Tedy $\mathrm{NSD}(a, b) = \mathrm{NSD}(a_0, a_1) = \mathrm{NSD}(a_1, a_2) = \ldots = \mathrm{NSD}(a_k, 0) = a_k$, což je poslední nenulová hodnota v posloupnosti. Pomocí Eukleidova algoritmu lze dokázat také následující větu.

Věta 3.3 (Bézoutova rovnost)**.** *Pro každou dvojici celých čísel a, b existují celá čísla u, v (tzv.* Bézoutovy koeficienty*) splňující*

$$\mathrm{NSD}(a, b) = u \cdot a + v \cdot b.$$

Eukleidův algoritmus rozšíříme tak, že v každém kroku spočte u_i, v_i taková, že $a_i = u_i \cdot a + v_i \cdot b$. Inicializujeme $(u_0, v_0) = (1, 0)$ a $(u_1, v_1) = (0, 1)$. Protože $a_{i+1} = a_{i-1} \bmod a_i = a_{i-1} - q_i a_i$, položíme

$$(u_{i+1}, v_{i+1}) = (u_{i-1}, v_{i-1}) - q_i(u_i, v_i),$$

kde $q_i = a_{i-1} \operatorname{div} a_i$. Pokud $a_{i+1} = 0$, pak u_i, v_i jsou zřejmě Bézoutovy koeficienty pro $a_i = \mathrm{NSD}(a, b)$.

Příklad. Pro $\mathrm{NSD}(168, 396)$ dostáváme posloupnosti

a_i	u_i	v_i
396	1	0
168	0	1
60	1	−2
48	−2	5
12	3	−7
0		

Tedy $\mathrm{NSD}(168, 396) = 3 \cdot 396 - 7 \cdot 168$.

Pomocí Bézoutovy rovnosti dokážeme jedno pomocné tvrzeníčko.

Lemma 3.4. *Buď p prvočíslo a $a, b \in \mathbb{Z}$. Platí-li $p \mid a \cdot b$, pak $p \mid a$ nebo $p \mid b$.*

(Opět, kdybychom měli v ruce jednoznačnost prvočíselných rozkladů, bylo by tvrzení očividné. Avšak v oboru $\mathbb{Z}[\sqrt{5}]$ platí $2 \mid (\sqrt{5}+1)(\sqrt{5}-1)$, přesto $2 \nmid \sqrt{5} \pm 1$; jednoznačnost tedy hraje roli.)

Důkaz. Předpokládejme, že $p \nmid a$. Pak $\mathrm{NSD}(a,p) = 1$, protože je p prvočíslo, a tedy podle Věty 3.3 existují čísla u, v splňující $au + pv = 1$. Vynásobením obou stran rovnosti číslem b dostaneme $abu + pvb = b$. Jelikož p dělí oba sčítance na levé straně, dělí i b. □

Indukcí snadno odvodíme následující důsledek:

Lemma 3.5. *Buď p prvočíslo a $a_1, \ldots, a_n$ celá čísla. Platí-li $p \mid a_1 \cdot \ldots \cdot a_n$, pak $p \mid a_i$ pro alespoň jedno i.*

Nyní můžeme přistoupit k důkazu jednoznačnosti prvočíselných rozkladů. Buď a nejmenší přirozené číslo s nejednoznačným prvočíselným rozkladem a uvažujme dva různé rozklady

$$a = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m} = q_1^{l_1} \cdot \ldots \cdot q_n^{l_n}.$$

Protože $p_1 \mid a = q_1^{l_1} \cdot \ldots \cdot q_n^{l_n}$, musí existovat i takové, že $p_1 \mid q_i$. Ovšem q_i je prvočíslo, tedy $p_1 = q_i$. Pak ale uvažujme číslo $b = \frac{a}{p_1}$: to má také dva různé rozklady

$$b = p_1^{k_1 - 1} \cdot p_2^{k_2} \cdot \ldots \cdot p_m^{k_m} = q_1^{l_1} \cdot \ldots \cdot q_i^{l_i - 1} \cdot \ldots \cdot q_n^{k_n},$$

ale přitom $b < a$, což je spor s minimalitou a. Věta 3.1 je dokázána.

3.3. **Kongruence.**

Zápis pomocí kongruencí, zavedený Gaussem ve zmiňované knize *Disquisitiones Arithmeticae* (1801), značně usnadňuje počítání modulo dané číslo.

Definice. Buď a, b, m celá čísla, $m \neq 0$. Řekneme, že *a je kongruentní s b modulo m*, a zapisujeme

$$a \equiv b \pmod{m},$$

pokud $m \mid a - b$, tj. pokud a a b dávají stejný zbytek po dělení m.

Je zřejmé, že relace „býti kongruentní modulo m“ je ekvivalence, tj. že pro všechna $a, b, c \in \mathbb{Z}$ platí

- $a \equiv a \pmod{m}$;
- pokud $a \equiv b \pmod{m}$, pak $b \equiv a \pmod{m}$;
- pokud $a \equiv b \pmod{m}$, a $b \equiv c \pmod{m}$, pak $a \equiv c \pmod{m}$.

Uvedené pozorování lze interpretovat tak, že symbol $\equiv$ můžeme používat podobně jako rovnítko: reflexivita říká, že z $a = b$ plyne $a \equiv b$, symetrie říká, že zápis je platný zleva doprava i zprava doleva, a tranzitivita říká, že máme-li sérii po sobě jdoucích kongruencí, pak výraz zcela vlevo je kongruentní výrazu zcela vpravo.

Druhou základní vlastností je invariance vůči základním operacím.

Tvrzení 3.6. *Nechť $a \equiv b \pmod{m}$ a $c \equiv d \pmod{m}$. Pak platí*

$$a + c \equiv b + d \pmod{m}, \quad a - c \equiv b - d \pmod{m}, \quad a \cdot c \equiv b \cdot d \pmod{m}$$

a pro každé přirozené k platí

$$a^k \equiv b^k \pmod{m}.$$

Důkaz. Podle předpokladu $m \mid a - b$ a $m \mid c - d$. Tedy $m \mid (a-b) + (c-d) = (a+c) - (b+d)$ a podobně pro operaci $-$. Dále $m \mid (a-b) \cdot c$ a $m \mid (c-d) \cdot b$, a tedy $m \mid (a-b) \cdot c + (c-d) \cdot b = ac - bd$. Poslední tvrzení se snadno dokáže z předchozího vzorce indukcí: $a^2 = a \cdot a \equiv b \cdot b = b^2 \pmod{m}$, $a^3 = a^2 \cdot a \equiv b^2 \cdot b = b^3 \pmod{m}$ atd. □

Na následující jednoduché úloze si ukážeme přehledný zápis výpočtu s využitím kongruence.

Úloha. Spočtěte $77^{123} + 66^{321} \bmod 6$.

Řešení. Protože $66 \equiv 0$ a $77 \equiv -1 \pmod 6$, z Tvrzení 3.6 plyne

$$77^{123} + 66^{321} \equiv (-1)^{123} + 0^{321} = -1 + 0 \equiv 5 \pmod 6.$$

Uvedený výraz tedy dává zbytek 5. □

Další důležitou vlastností je, že v kongruenci smíme krátit číslem, které je nesoudělné s modulem m. Naopak, jsou-li všechna tři čísla v kongruenci soudělná, celý výraz můžeme zjednodušit tím, že společný faktor vykrátíme na obou stranách *i v modulu*. Formálně tyto vlastnosti vyjadřuje následující tvrzení.

Tvrzení 3.7. *Buď a, b, c, m celá čísla, $c, m \neq 0$. Pak*

(1) $a \equiv b \pmod m \;\Leftrightarrow\; ca \equiv cb \pmod{cm}$*;*

(2) *jsou-li c, m nesoudělná, pak* $a \equiv b \pmod m \;\Leftrightarrow\; ca \equiv cb \pmod m$.

Důkaz. (1) Tvrzení říká, že $m \mid a - b \;\Leftrightarrow\; cm \mid ca - cb = c(a - b)$, což je zřejmé.

(2) Protože $m \mid ca - cb = c(a - b)$ a čísla c, m jsou nesoudělná, musí platit $m \mid a - b$. Opačná implikace plyne z Tvrzení 3.6. □

Poznámka. Obecně platí

$$ca \equiv cb \pmod m \;\Leftrightarrow\; a \equiv b \pmod{m/\operatorname{NSD}(c, m)},$$

z čehož Tvrzení 3.7 snadno plyne. Proveďte důkaz jako cvičení!

Úloha. Najděte všechna x splňující a) $6x \equiv 9 \pmod{21}$, b) $10x \equiv 5 \pmod{21}$.

Řešení. Použijeme několikrát Tvrzení 3.7.

a) Užitím (1) dostaneme ekvivalentní podmínku $2x \equiv 3 \pmod 7$, a po přenásobení obou stran číslem 4, díky (2), ekvivalentní podmínku $x \equiv 5 \pmod 7$. Řešením jsou všechna $x = 5 + 7k$, $k \in \mathbb{Z}$.

b) Užitím (2) dostaneme ekvivalentní podmínku $2x \equiv 1 \pmod{21}$, a po přenásobení obou stran číslem 11, díky (2), ekvivalentní podmínku $x \equiv 11 \pmod{21}$. Řešením jsou všechna $x = 11 + 21k$, $k \in \mathbb{Z}$. □

3.4. **Eulerova věta.**

Pro motivaci připomeňme úlohu uvedenou za základními vlastnostmi kongruencí: řešení bylo snadné především proto, že $66 \equiv 0$ a $77 \equiv -1$, přičemž tato čísla se snadno umocňují. Zamyslete se nad následující úlohou.

Úloha. Zjistěte poslední cifru čísla 77^{123}.

Řešení. Jinými slovy, spočtěte $77^{123} \bmod 10$. Můžeme psát $77^{123} \equiv 7^{123} \pmod{10}$. Nemáme-li však k dispozici lepší teorii, nezbývá, než zkoušet mocnit sedmičku. Záhy si všimneme, že se poslední cifry opakují s periodou 4, a protože $123 \bmod 4 = 3$, dostáváme $7^{123} \equiv 7^3 = 3 \pmod{10}$. □

To, že zbytky modulo dané číslo vykazují periodu jako v předchozí úloze, není náhoda, nýbrž pravidlo, které se nazývá *Eulerova věta*. Délku periody udává tzv. Eulerova funkce.

Definice. *Eulerova funkce* $\varphi(n)$ značí pro přirozené číslo n počet čísel $k \in \{1, \ldots, n\}$ nesoudělných s číslem n, tj. splňujících $\operatorname{NSD}(k, n) = 1$.

Např. $\varphi(10) = 4$, neboť s desítkou nesoudělná jsou právě čísla $1, 3, 7, 9$. Pro libovolné prvočíslo p platí $\varphi(p) = p-1$, protože nesoudělná jsou s ním právě všechna menší čísla.

Výpočet Eulerovy funkce pouze z definice by byl pro větší než malá čísla poněkud pracný. Naštěstí existuje vzorec, pomocí něhož je snadné spočítat hodnotu $\varphi(n)$, pokud známe prvočíselný rozklad čísla n.

Tvrzení 3.8. *Je-li $n = p_1^{k_1} \cdot \ldots \cdot p_m^{k_m}$ prvočíselný rozklad čísla $n > 1$, pak*

$$\varphi(n) = p_1^{k_1-1}(p_1 - 1) \cdot \ldots \cdot p_m^{k_m-1}(p_m - 1).$$

Příklad. $\varphi(4056) = \varphi(2^3 \cdot 3^1 \cdot 13^2) = 2^2 \cdot 1 \cdot 3^0 \cdot 2 \cdot 13^1 \cdot 12 = 1248$.

Důkaz správnosti vzorce není úplně jednoduchý, necháme si jej na později. Teď se podíváme na samotnou Eulerovu větu.

Věta 3.9 (Eulerova věta)**.** *Buďte a, m nesoudělná přirozená čísla. Pak*

$$a^{\varphi(m)} \equiv 1 \pmod{m}.$$

K důkazu se nám bude hodit jedno pomocné lemma. Označme

$$m^* = \{k \in \{1, \ldots, m-1\} : \ \mathrm{NSD}(k, m) = 1\}.$$

Eulerovu funkci pak můžeme zapsat jako $\varphi(m) = |m^*|$.

Lemma 3.10. *Buďte a, m nesoudělná přirozená čísla a definujme*

$$\begin{aligned} f_a : m^* &\to m^* \\ x &\mapsto ax \bmod m. \end{aligned}$$

Pak je zobrazení f_a bijekce.

Důkaz. Předně vzniká otázka: je vůbec $ax \bmod m$ vždy prvek m^*? Ovšemže ano: jsou-li obě čísla a, x nesoudělná s m, pak je s m nesoudělné i číslo ax a tudíž podle Lemmatu 3.2 také $ax \bmod m$.

Dokážeme, že zobrazení f_a je bijekce. Protože jde o zobrazení na konečné množině, stačí díky Lemmatu 2.3 ověřit prostost. Uvažujme tedy $x, y \in m^*$ taková, že $f_a(x) = f_a(y)$, tj. $ax \equiv ay \pmod{m}$. Podle Tvrzení 3.7 je $x \equiv y \pmod{m}$, tedy x i y dávají stejný zbytek po dělení m. Ovšem obě čísla jsou menší než m, takže musí být stejná. □

Důkaz Eulerovy věty. Uvažujme následující výpočet, kde f_a je zobrazení definované v předchozím lemmatu:

$$\prod_{b \in m^*} b \ = \ \prod_{b \in m^*} f_a(b) \ = \ \prod_{b \in m^*} ab \bmod m \ \equiv \ \prod_{b \in m^*} ab \ = \ a^{\varphi(m)} \cdot \prod_{b \in m^*} b \pmod{m}.$$

První rovnost platí díky tomu, že v obou případech násobíme přes všechny prvky množiny m^*, pouze v různém pořadí. Označíme-li

$$c = \prod_{b \in m^*} b,$$

právě jsme dokázali, že

$$c \equiv a^{\varphi(m)} \cdot c \pmod{m}.$$

Číslo c je nesoudělné s m (protože je součinem čísel nesoudělných s m), takže jím můžeme podle Tvrzení 3.7 krátit a dostáváme $1 \equiv a^{\varphi(m)} \pmod{m}$. □

Leonhard Euler publikoval tuto větu v roce 1736. Speciální případ pro m prvočíslo bývá připisován Pierre de Fermatovi (objevuje se v jednom z jeho dopisů z roku 1640), a někdy se nazývá Malá Fermatova věta.

Důsledek 3.11 (Malá Fermatova věta)**.** *Je-li p prvočíslo a $p \nmid a$, pak*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Poznámka. Malou Fermatovu větu je velmi snadné dokázat přímo: indukcí podle a dokažte ekvivalentní tvrzení $a^p \equiv a \pmod{p}$. Alternativní důkaz Eulerovy věty bychom pak mohli dostat takto: nejprve pomocí Fermatovy věty dokážeme indukcí podle k Eulerovu větu pro všechna $n = p^k$, a poté využijeme následující vlastnosti kongruencí: jsou-li m, n nesoudělné, pak $u \equiv v \pmod{mn}$ právě tehdy, když $u \equiv v \pmod{m}$ a $u \equiv v \pmod{n}$.

Úloha. Zjistěte poslední cifru čísla 77^{123}.

Řešení. Použijeme Eulerovu větu: protože $\varphi(10) = 4$ a $\mathrm{NSD}(77, 10) = 1$, platí

$$77^{123} \equiv 7^{123} = 7^{4 \cdot 30 + 3} \equiv (7^4)^{30} \cdot 7^3 \equiv 1^{30} \cdot 3 = 3 \pmod{10}.$$

(Z didaktických důvodů jsme vše detailně rozepsali, v praxi samozřejmě provedete většinu úvah zpaměti a budete psát rovnou $7^{333} \equiv 7^3 = 3$.) □

Úloha. Spočtěte $8^{7^6} \bmod 21$.

Řešení. Opět použijeme Eulerovu větu: protože $\varphi(21) = 12$ a $\mathrm{NSD}(8, 21) = 1$, stačí zjistit zbytek po dělení 7^6 číslem 12. Tedy řešíme úlohu $7^6 \bmod 12$ a ještě jednou použijeme Eulerovu větu: protože $\varphi(12) = 4$ a $\mathrm{NSD}(7, 12) = 1$, stačí zjistit zbytek po dělení exponentu 6 číslem 4, což je 2. Tedy $7^6 \equiv 7^2 = 49 \equiv 1 \pmod{12}$ a $8^{7^6} \equiv 8^1 = 8 \pmod{21}$. □

Úloha. Řešte $x^6 + x + xy \equiv 1 \pmod{7}$.

Řešení. Pokud $7 \mid x$, pak 7 dělí levou stranu, a tedy $x^6 + x + xy$ nedává zbytek 1 po dělení 7. Takže budeme předpokládat, že 7 nedělí x a použijeme malou Fermatovu větu, která říká, že $x^6 \equiv 1 \pmod{7}$. Zadaná rovnice je tak ekvivalentní rovnici $1 + x + xy \equiv 1 \pmod{7}$, tj. $7 \mid x(y + 1)$. Protože předpokládáme, že $7 \nmid x$, musí 7 dělit $y + 1$, tj. $y \equiv -1 \pmod{7}$. Řešením je tedy množina

$$\big\{(x, y) : \ 7 \nmid x, \ y \equiv -1 \pmod{7}\big\}.$$

□

Poznámka. Podle Lemmatu 3.10 pro každé a nesoudělné s m existuje právě jedno $b \in \{1, \ldots, m-1\}$ takové, že

$$ab \equiv 1 \pmod{m}.$$

Toto b lze podle Eulerovy věty spočítat jako $b = a^{\varphi(m)-1}$. Jiný, efektivnější, postup dává Eukleidův algoritmus: pokud zjistíme Bézoutovy koeficienty $1 = \mathrm{NSD}(a, m) = ua + vm$, pak $1 \equiv ua \pmod{m}$ a odpovědí je $b = u \bmod m$. Toto pozorování nachází aplikaci např. při výpočtu inverzních prvků v tělese $\mathbb{Z}_p$, viz kapitola o tělesech.

3.5. **Čínská věta o zbytcích.**

Čínská věta o zbytcích hovoří o řešeních soustav lineárních kongruencí. Byla známa již starověkým Číňanům (je uvedena v knize matematika Sun-c' ze 4. století) a o něco málo později i ve staré Indii.

Věta 3.12 (Čínská věta o zbytcích)**.** *Nechť $m_1, \ldots, m_n$ jsou po dvou nesoudělná přirozená čísla, označme $M = m_1 \cdot \ldots \cdot m_n$. Pak pro libovolná celá čísla $u_1, \ldots, u_n$ existuje právě jedno $x \in \{0, \ldots, M-1\}$, které řeší soustavu kongruencí*

$$x \equiv u_1 \pmod{m_1}, \quad \ldots, \quad x \equiv u_n \pmod{m_n}.$$

Důkaz. Nejprve dokážeme jednoznačnost řešení. Předpokládejme, že soustava má dvě řešení $x, y \in \{0, \ldots, M-1\}$, tj. pro každé i platí

$$x \equiv y \equiv u_i \pmod{m_i}.$$

Pak pro každé i

$$m_i \mid x - y$$

a protože jsou čísla m_i navzájem nesoudělná, dostáváme

$$M = m_1 \cdot \ldots \cdot m_n \mid x - y.$$

Ovšem $|x - y| < M$ (protože x, y volíme z intervalu $0, \ldots, M-1$), takže $x - y = 0$, tj. $x = y$.

Nyní dokážeme, že nějaké řešení vůbec existuje. Uvažujme zobrazení

$$f : \{0, \ldots, M-1\} \to \{0, \ldots, m_1 - 1\} \times \cdots \times \{0, \ldots, m_n - 1\}$$
$$x \mapsto (x \bmod m_1, \ldots, x \bmod m_n).$$

V předchozím odstavci jsme vlastně ukázali, že zobrazení f je prosté. Přitom definiční obor i obor hodnot této funkce mají stejnou velikost M (velikost kartézského součinu je součin velikostí činitelů), takže zobrazení f musí být podle Lemmatu 2.3 i na. Tedy ke každé n-tici $(u_1, \ldots, u_n)$ existuje právě jedno x, které se na něj zobrazuje; a to je hledané řešení soustavy. □

Uvedený důkaz je zvláštní tím, že nedává žádný návod, jak řešení dané soustavy spočítat. Obecný postup (a tím i alternativní důkaz čínské věty o zbytcích) lze vypozorovat z řešení následující úlohy.

Úloha. Najděte všechna řešení soustavy kongruencí

$$x \equiv 2 \pmod{3}, \quad x \equiv 1 \pmod{4}, \quad x \equiv 3 \pmod{5}.$$

Řešení. Z první rovnice vidíme, že $x = 3k + 2$, $k \in \mathbb{Z}$. Dosadíme do druhé rovnice a dostaneme $3k + 2 \equiv 1 \pmod{4}$, tedy $k \equiv 1 \pmod{4}$ a vidíme, že $k = 4l + 1$ a $x = 12l + 5$, $l \in \mathbb{Z}$. Dosadíme do třetí rovnice a dostaneme $12l + 5 \equiv 3 \pmod{5}$, tedy $l \equiv 4 \pmod{5}$, takže $l = 5m + 4$ a $x = 60m + 53$, $m \in \mathbb{Z}$. □

Traduje se, že motivací čínské věty o zbytcích byl způsob, jakým jistý čínský generál počítal své vojáky — ostatně, sám Sun-c' je znám spíše jako vojevůdce a jeho stěžejním dílem je kniha *Umění války*. Generál věděl, že před bitvou měl 1000 vojáků, a chtěl je spočítat po bitvě. Nechal je tedy řadit do trojstupů, čtyřstupů, atd., a zjišťoval, kolik mu jich zbyde mimo řady. Jinými slovy, zjistil, kolik je počet vojáků modulo 3, modulo 4, atd. Z čínské věty o zbytcích plyne, že pokud zvolil dostatek nesoudělných čísel (součin > 1000), může jednoznačně určit celkový počet svých vojáků.

Na závěr pomocí čínské věty o zbytcích dokážeme vzorec na výpočet Eulerovy funkce, tj. vztah

$$\varphi(p_1^{k_1} \cdot \ldots \cdot p_m^{k_m}) = p_1^{k_1-1}(p_1-1) \cdot \ldots \cdot p_m^{k_m-1}(p_m-1).$$

Důkaz Tvrzení 3.8. Dokážeme následující dvě vlastnosti:

(1) pro každé prvočíslo p platí $\varphi(p^k) = p^{k-1}(p-1)$;
(2) pro každá dvě nesoudělná čísla a, b platí $\varphi(ab) = \varphi(a) \cdot \varphi(b)$.

Uvedený vzorec snadno plyne z těchto dvou tvrzení: číslo n rozložíme na součin m po dvou nesoudělných mocnin $p_i^{k_i}$ a dostaneme

$$\varphi(n) \overset{(2)}{=} \varphi(p_1^{k_1}) \cdot \ldots \cdot \varphi(p_m^{k_m}) \overset{(1)}{=} p_1^{k_1-1}(p_1-1) \cdot \ldots \cdot p_m^{k_m-1}(p_m-1).$$

(1) V tomto speciálním případě je snadné spočítat *soudělná* čísla: jsou to právě čísla $p, 2p, 3p, \ldots, p^{k-1} \cdot p$. Vidíme, že jich je p^{k-1}. Všechna zbylá čísla jsou nesoudělná, takže $\varphi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1)$.

(2) Uvažujme zobrazení

$$f : \{0, \ldots, ab-1\} \to \{0, \ldots, a-1\} \times \{0, \ldots, b-1\}$$
$$x \mapsto (x \bmod a, x \bmod b).$$

Podle čínské věty o zbytcích je f bijekce. Dále uvažujme pouze restrikci f na množinu $(ab)^*$. To je prosté zobrazení, jehož definiční obor je množina $(ab)^*$ velikosti $\varphi(ab)$. Stačí tedy dokázat, že jeho oborem hodnot je množina $a^* \times b^*$ – pak, díky prostosti, bude $\varphi(ab) = |(ab)^*| = |a^* \times b^*| = |a^*| \cdot |b^*| = \varphi(a) \cdot \varphi(b)$, což chceme dokázat. Potřebujeme tedy ověřit, že

(a) f zobrazuje množinu $(ab)^*$ do množiny $a^* \times b^*$, tj. že $\mathrm{NSD}(x, ab) = 1$ implikuje $\mathrm{NSD}(x \bmod a, a) = \mathrm{NSD}(x \bmod b, b) = 1$;
(b) f zobrazuje množinu $(ab)^*$ <u>na</u> tuto množinu, tj. že pokud $\mathrm{NSD}(u, a) = 1$ a $\mathrm{NSD}(v, b) = 1$, pak to jediné x, které se zobrazuje na dvojici (u, v), splňuje $\mathrm{NSD}(x, ab) = 1$.

Pro důkaz (a) si stačí uvědomit, že $\mathrm{NSD}(x \bmod a, a) = \mathrm{NSD}(x, a)$, a kdyby tato čísla byla soudělná, tím spíše by byla soudělná čísla x, ab. Podobně pro b.

Pro důkaz (b) uvažujme (to jediné) x zobrazující se na (u, v), tj. $u = x \bmod a$ a $v = x \bmod b$. Dosazením za u, v plyne $\mathrm{NSD}(x, a) = \mathrm{NSD}(x \bmod a, a) = 1$ a $\mathrm{NSD}(x, b) = \mathrm{NSD}(x \bmod b, b) = 1$. Kdyby byla čísla x, ab soudělná, pak by existovalo prvočíslo p, které dělí zároveň x i ab, tedy podle Lemmatu 3.4 by p dělilo a nebo b, a tudíž by x, a nebo x, b byly soudělné, což je spor. □

Teorie dělitelnosti

4. Obory integrity

Cíl. *Zavedeme pojem oboru integrity, který abstraktně vymezuje prostředí, ve kterém lze studovat dělitelnost. Jako hlavní příklady představíme obor celých čísel a jeho rozšíření, a dále obory polynomů a formálních mocninných řad. Na závěr uvedeme konstrukci podílových těles, která slouží k formalizaci pojmu zlomek.*

4.1. Základní vlastnosti.

Celá čísla sdílí z hlediska dělitelnosti řadu vlastností s dalšími obory. Dělitelnost lze studovat pro polynomy, ale také třeba pro různá rozšíření celých čísel, např. *Gaussova celá čísla*, tj. komplexní čísla s celočíselnými koeficienty, a další struktury. V různých oborech pak platí různě silná tvrzení: např. analogie základní věty aritmetiky 3.1 platí pro celočíselné i racionální polynomy i pro Gaussova celá čísla. Polynomy nad tělesem i Gaussova čísla lze dělit se zbytkem a platí pro ně Bézoutova rovnost, což však není pravda např. pro celočíselné polynomy nebo pro polynomy více proměnných. A pro některá rozšíření $\mathbb{Z}$ neplatí ani základní věta aritmetiky. V následujících čtyřech sekcích se budeme snažit udělat v uvedených vlastnostech a příkladech pořádek.

Abychom mohli studovat všechny zmíněné obory naráz, zavádí se obecná struktura nazývaná *obor integrity*, jejíž axiomy vystihují základní aritmetické vlastnosti. Jde o stejný princip, který vedl v lineární algebře k abstraktnímu pojmu tělesa a vektorového prostoru.

Definice. *Komutativním okruhem* $\mathbf{R}$ rozumíme pětici $(R, +, -, \cdot, 0)$, kde R je množina, na které jsou definovány binární operace $+, \cdot$, unární operace $-$ a konstanta 0 splňující pro každé $a, b, c \in R$ následující podmínky:

$$a + (b + c) = (a + b) + c, \qquad a + b = b + a, \qquad a + 0 = a,$$
$$a + (-a) = 0,$$
$$a \cdot (b \cdot c) = (a \cdot b) \cdot c, \qquad a \cdot b = b \cdot a,$$
$$a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

Komutativním okruhem s jednotkou pak rozumíme komutativní okruh, ve kterém existuje prvek $1 \in R$ splňující $a \cdot 1 = a$ pro každé $a \in R$.

- Platí-li navíc podmínka

$$\text{pokud } a, b \neq 0, \text{ pak } a \cdot b \neq 0,$$

nazýváme $\mathbf{R}$ *obor integrity.*

- Platí-li navíc podmínka

  $$\text{pro každé } a \neq 0 \text{ existuje } b \text{ splňující } a \cdot b = 1,$$

  nazýváme $\mathbf{R}$ *těleso*. Podle Tvrzení 4.1(5) je takové b jednoznačně určeno a značíme jej a^{-1}.

V zápise zpravidla vynecháváme závorky, násobení má vyšší prioritu než sčítání. Místo $a + (-b)$ píšeme $a - b$. Formálně odlišujeme mezi množinou R, tzv. *nosnou množinou*, a pěticí $\mathbf{R} = (R, +, -, \cdot, 0)$, která navíc obsahuje informaci o algebraické struktuře definované na R. Nebude-li výslovně uvedeno jinak, zápisem $\mathbf{R}$ rozumíme takto označenou pětici.

V matematice obecně je zvykem uvádět množinu axiomů tak krátkou, jak je to jen možné; spousta užitečných vlastností se tak do ní nevejde. Následující tvrzení ukazuje několik aritmetických pravidel, které z definice snadno plynou a v dalším textu je budeme zcela automaticky používat.

Tvrzení 4.1. *Buď $\mathbf{R}$ komutativní obor s jednotkou, $a, b, c \in R$. Pak*

(1) *pokud $a + c = b + c$, pak $a = b$;*
(2) $a \cdot 0 = 0$*;*
(3) $-(-a) = a, \quad -(a + b) = -a - b$*;*
(4) $-(a \cdot b) = (-a) \cdot b = a \cdot (-b), \quad (-a) \cdot (-b) = ab$*;*
(5) *je-li $\mathbf{R}$ oborem integrity, pokud $a \cdot c = b \cdot c$ a $c \neq 0$, pak $a = b$.*

Důkaz. (1) Je-li $a+c = b+c$, pak také $(a+c)+(-c) = (b+c)+(-c)$. Použitím axiomů dostaneme $(a+c)+(-c) = a+(c+(-c)) = a+0 = a$ a podobně $(b+c)+(-c) = b$, tedy $a = b$.

(2) Pomocí distributivity spočteme $0 + a \cdot 0 = a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$ a z (1) dostáváme $a \cdot 0 = 0$.

(3) Protože $0 = a + (-a) = -(-a) + (-a)$, z (1) dostáváme $a = -(-a)$. Protože $0 = (a + b) + (-(a + b))$ a zároveň $0 = a + (-a) + b + (-b) = (a + b) + (-a - b)$, z (1) dostáváme $-(a + b) = -a - b$.

(4) Protože $a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0 \cdot b = 0 = a \cdot b + (-(a \cdot b))$, z (1) dostáváme $-(a \cdot b) = (-a) \cdot b$. Druhou rovnost dokážeme analogicky a užitím předchozího $(-a) \cdot (-b) = -(a \cdot (-b)) = -(-(a \cdot b)) = a \cdot b$.

(5) Protože $a \cdot c = b \cdot c$, platí $0 = a \cdot c - b \cdot c = (a - b) \cdot c$. Tedy aspoň jeden z prvků c, $a - b$ musí být 0. Protože předpokládáme $c \neq 0$, musí být $a - b = 0$, tedy $a = b$. □

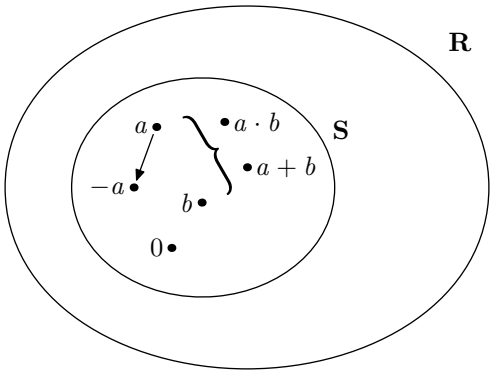
Buď $\mathbf{R}$ komutativní okruh s jednotkou. Jeho *charakteristikou* rozumíme nejmenší přirozené číslo n, pro které

$$\underbrace{1 + 1 + \ldots + 1}_{n} = 0.$$

Pokud takové n neexistuje, charakteristiku definujeme 0.

Tvrzení 4.2. *Obor integrity má charakteristiku 0 nebo prvočíslo.*

Důkaz. Označme prvek $\underbrace{1 + 1 + \ldots + 1}_{n}$ symbolem $\underline{n}$. Kdyby byla charakteristika $n = a \cdot b$, $a, b \neq 1$, pak bychom měli $0 = \underline{n} = \underline{a} \cdot \underline{b}$, a tedy $\underline{a} = 0$ nebo $\underline{b} = 0$, což je spor s minimalitou n. □



Obrázek 3. Podobor **S** oboru **R**.

4.2. **Příklady oborů integrity.**

Příklad. Základními příklady komutativních okruhů s jednotkou jsou standardní *číselné obory*:

- $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ se standardními operacemi jsou příklady těles charakteristiky 0;
- $\mathbb{Z}$ se standardními operacemi je příkladem oboru integrity charakteristiky 0;
- $\mathbb{Z}_n = (\{0, \dots, n-1\}, +_{\bmod n}, -_{\bmod n}, \cdot_{\bmod n}, 0, 1)$ s operacemi modulo n je příkladem komutativního okruhu charakteristiky n. Pro n složené tedy nemůže jít o obor integrity.

Bez jednotky je např. komutativní okruh všech sudých celých čísel.

Tvrzení 4.3. *Každé těleso je oborem integrity.*

Důkaz. Kdyby existovaly $a, b \neq 0$ takové, že $a \cdot b = 0$, pak $b = (a^{-1} \cdot a) \cdot b = = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0$, což je spor. □

Tvrzení 4.4. *Konečný obor integrity je tělesem.*

Důkaz. Označme tento obor **R**. Pro nenulové $a \in R$ uvažujme zobrazení

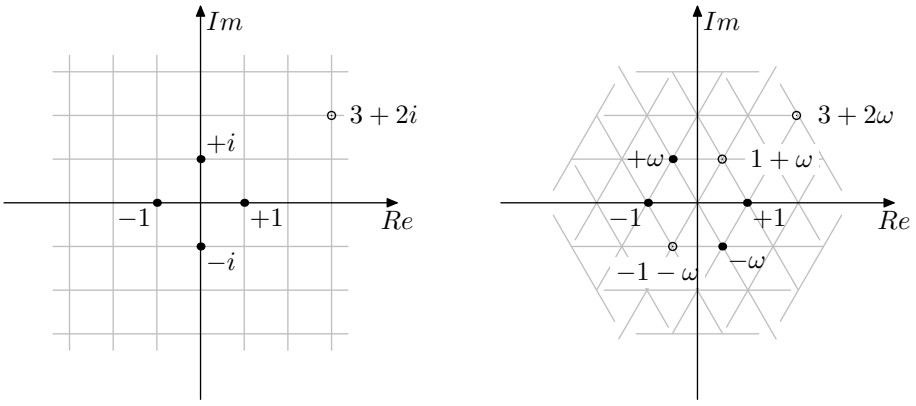
$$f_a : R \to R, \quad x \mapsto a \cdot x.$$

Podle Tvrzení 4.1(5) je toto zobrazení prosté, a protože jde o zobrazení na konečné množině, podle Lemmatu 2.3 je to bijekce. Inverzním prvkem k prvku a je tedy $f_a^{-1}(1)$. □

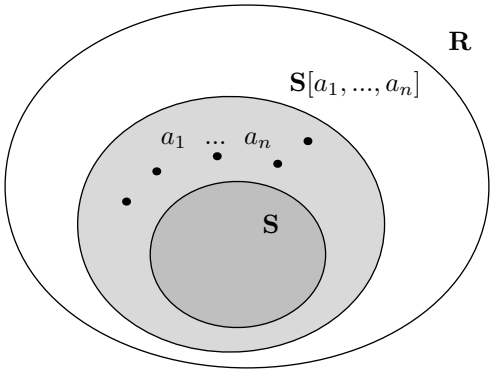
Více o konečných tělesech se dozvíme v Sekci **??**.

Další příklady oborů integrity můžeme dostat z již známých oborů pomocí různých konstrukcí. Jedním z nejdůležitějších příkladů jsou *obory polynomů*, které budou formálně zavedeny v Sekci 5. Druhým základním pojmem je *podobor* a *rozšíření* daného oboru.

Definice. Buď **R** obor integrity a $S \subseteq R$ podmnožina taková, že $0, 1 \in S$ a kdykoliv $a, b \in S$, pak také $-a \in S$, $a+b \in S$ a $a \cdot b \in S$. Vezmeme-li na této množině restrikce operací oboru **R**, dostaneme také obor integrity, který označíme **S** (jsou-li všechny axiomy splněny na větší množině R, pak jistě i na její podmnožině S). Takové obory se nazývají *podobory* oboru **R**.



Obrázek 4. Gaussova a Eisensteinova celá čísla



Obrázek 5. Rozšíření $\mathbf{S}[a_1, \ldots, a_n]$ oboru $\mathbf{S}$.

Příklady.

- Obor $\mathbb{Z}$ je podoborem oboru $\mathbb{Q}$, který je podoborem oboru $\mathbb{R}$, který je podoborem oboru $\mathbb{C}$.
- Množina $\{a + bi : \ a, b \in \mathbb{Z}\}$ tvoří podobor oboru $\mathbb{C}$. Tento obor se nazývá *Gaussova celá čísla*.
- Množina $\{a + b\omega : \ a, b \in \mathbb{Z}\}$, kde $\omega = e^{2\pi i/3}$ je komplexní třetí odmocnina z jedné, tvoří podobor oboru $\mathbb{C}$. Tento obor se nazývá *Eisensteinova celá čísla*.
- Množina $\{a \in \mathbb{C} : \ a$ algebraické číslo$\}$ tvoří podobor oboru $\mathbb{C}$, viz Věta **??**.

Definice. Buď $\mathbf{S}$ podobor oboru $\mathbf{R}$ a $a_1, \ldots, a_n \in R$. Definujeme $\mathbf{S}[a_1, \ldots, a_n]$ jako nejmenší podobor oboru $\mathbf{R}$ obsahující množinu S i prvky $a_1, \ldots, a_n$. Tomuto oboru se říká *rozšíření* $\mathbf{S}$ *o prvky* $a_1, \ldots, a_n$.

Více o podoborech a rozšířeních se dozvíte v kapitole o okruzích a tělesech. Zatím si ukážeme nějaké příklady.

Příklady.

- $\mathbb{Z}[i]$ jsou Gaussova celá čísla, $\mathbb{R}[i] = \mathbb{C}$.

- Obecněji,
$$\mathbb{Z}[\sqrt{s}] = \{a + b\sqrt{s} : \ a, b \in \mathbb{Z}\} \subset \mathbb{C}$$
je oborem integrity pro libovolné celé číslo s (rozumí se $\sqrt{-1} = i$).
- Můžeme uvažovat i komplikovanější obory, jako např.
$$\mathbb{Z}[\sqrt{2}, \sqrt{3}] = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : \ a, b, c, d \in \mathbb{Z}\}$$
nebo
$$\mathbb{Z}[\sqrt[3]{s}] = \{a + b\sqrt[3]{s} + c\sqrt[3]{s^2} : \ a, b, c \in \mathbb{Z}\}.$$
- Obecně
$$\mathbf{S}[u] = \{a_0 + a_1 u + \ldots + a_n u^n : \ n \in \mathbb{N}, a_0, \ldots, a_n \in S\}.$$
Pokud např. $\mathbf{S} = \mathbb{Z}$ a $u = \pi$ (nebo libovolné jiné transcendentní číslo), pak jsou tyto prvky pro různé koeficienty různé.

Rozšíření oboru celých čísel se objevují v řadě aplikací, především v teorii čísel. Obecnou teorii budeme často ilustrovat na kvadratických rozšířeních celých čísel. Základní početní nástroj, *normu*, představíme v sekci 6. V Sekci 9.3 si pak ukážeme, jak tuto teorii aplikovat při řešení jistého typu diofantických rovnic.

4.3. **Podílová tělesa.**

Tak jako lze obor celých čísel rozšířit do tělesa racionálních čísel, každý obor integrity $\mathbf{R}$ lze rozšířit na tzv. *podílové těleso*, které lze zkonstruovat jako „těleso zlomků“, jejichž čitatel i jmenovatel jsou prvky daného oboru. Podílová tělesa hrají v komutativní algebře důležitou roli, jak uvidíme například v Sekci **??**, kde nám budou nástrojem k důkazu Gaussovy věty.

Konstrukce probíhá následujícím způsobem. Definujeme relaci $\sim$ na množině $R \times (R \setminus \{0\})$ předpisem
$$(a, b) \sim (c, d) \quad \Longleftrightarrow \quad ad = bc.$$
Není těžké nahlédnout, že jde o ekvivalenci: reflexivita je zřejmá, symetrie plyne z komutativity násobení a tranzitivitu získáme následujícím výpočtem: je-li $(a, b) \sim$ $\sim (c, d) \sim (e, f)$, tedy $ad = bc$ a $cf = de$. Pak ale $adf = bcf = bde$, a tedy $af = be$, protože $d \neq 0$ (ke krácení potřebujeme předpoklad, že $\mathbf{R}$ je obor integrity!).

Pro jednoduchost vyjadřování budeme značit blok $[(a, b)]_\sim$ této ekvivalence jako zlomek $\frac{a}{b}$. Uvažujme množinu Q všech bloků této ekvivalence (tj. všech zlomků) a definujme na ní operace
$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad -\frac{a}{b} = \frac{-a}{b}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad 0 = \frac{0}{1}, \qquad 1 = \frac{1}{1}.$$
Je třeba dokázat, že tyto operace jsou dobře definované. Předně, aby jmenovatel součtu a součinu zůstal nenulový, potřebujeme předpoklad, že $\mathbf{R}$ je obor integrity. A dále musíme dokázat, že pokud zvolíme jiné reprezentanty zlomků, výsledek operace zůstane stejný. Formálně, pokud $\frac{a}{b} = \frac{a'}{b'}$ a $\frac{c}{d} = \frac{c'}{d'}$, potřebujeme dokázat, že $\frac{a}{b} + \frac{c}{d} = \frac{a'}{b'} + \frac{c'}{d'}$, a podobně pro odčítání a násobení. Důkaz provedeme pro sčítání: chceme ověřit, že $\frac{ad+bc}{bd} = \frac{a'd'+b'c'}{b'd'}$, tedy že $(ad + bc)(b'd') = (a'd' + b'c')(bd)$. Roznásobíme a využijeme faktu, že $ab' = a'b$ a $cd' = c'd$. Označme $\mathbf{Q}$ množinu Q s operacemi $+, -, \cdot$ a konstantami 0,1.

Tvrzení 4.5. *Buď* $\mathbf{R}$ *obor integrity a* $\mathbf{Q}$ *výsledek právě popsané konstrukce. Pak* $\mathbf{Q}$ *je těleso a obor* $\mathbf{R}$ *je podoborem tělesa* $\mathbf{Q}$*, pokud ztotožníme prvek* $a \in R$ *s prvkem* $\frac{a}{1} \in Q$.

Těleso $\mathbf{Q}$ se nazývá *podílové těleso* oboru $\mathbf{R}$.

Důkaz. Ověříme postupně všechny axiomy:

- Asociativita sčítání: $\frac{a}{b}+(\frac{c}{d}+\frac{e}{f}) = \frac{a}{b}+\frac{cf+de}{df} = \frac{adf+b(cf+de)}{bdf} = \frac{adf+bcf+bde}{bdf} = = \frac{ad+bc}{bd} + \frac{e}{f} = (\frac{a}{b} + \frac{c}{d}) + \frac{e}{f}$.
- Komutativita sčítání: $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd} = \frac{cb+da}{db} = \frac{c}{d} + \frac{a}{b}$.
- Nula: $\frac{a}{b} + \frac{0}{1} = \frac{a\cdot 1+b\cdot 0}{b\cdot 1} = \frac{a}{b}$.
- Odčítání: $\frac{a}{b} + \frac{-a}{b} = \frac{ab+(-ab)}{b^2} = \frac{0}{b^2} = 0$.
- Asociativita a komutativita násobení plyne okamžitě z týchž vlastností oboru $\mathbf{R}$.
- Jednotka: $\frac{a}{a} \cdot \frac{1}{1} = \frac{a\cdot 1}{a\cdot 1} = \frac{a}{a}$.
- Distributivita: $\frac{a}{b} \cdot (\frac{c}{d} + \frac{e}{f}) = \frac{acf+ade}{bdf} = \frac{abcf+abde}{b^2df} = \frac{ac}{bd} + \frac{ae}{bf}$.
- $0 = \frac{0}{1} \neq 1 = \frac{1}{1}$, protože $0 \cdot 1 \neq 1 \cdot 1$.

Navíc $\frac{a}{b} \cdot \frac{b}{a} = \frac{ab}{ba} = \frac{1}{1}$ pro každé $\frac{a}{b} \neq 0$, čili $\mathbf{Q}$ je těleso. Zbývá dokázat, že prvky tvaru $\frac{a}{1}$ tvoří podobor $\mathbb{Q}$, což je snadné. □

Příklad. Těleso racionálních čísel $\mathbb{Q}$ je *definováno* jako podílové těleso oboru $\mathbb{Z}$.

Příklad. Je-li $\mathbf{T}$ těleso, pak jeho podílové těleso je, při výše uvedeném ztotožnění $a = \frac{a}{1}$, rovno $\mathbf{T}$, protože $\frac{a}{b} = \frac{ab^{-1}}{1}$ pro každé $a, b \in T$, $b \neq 0$.

Příklad. Podílové těleso oboru $\mathbb{Z}[i]$ je, formálně vzato, těleso zlomků tvaru $\frac{a+bi}{c+di}$, kde $a, b, c, d \in \mathbb{Z}$. Pokud ztotožníme tento zlomek se číslem $\frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i$, dostaneme těleso $\mathbb{Q}[i]$. (Formálně bychom řekli, že podílové těleso oboru $\mathbb{Z}[i]$ je *izomorfní* s tělesem $\mathbb{Q}[i]$, viz Sekce **??**).

5. Polynomy

Cíl. *Nejprve zformulujeme, co je to přesně polynom (a formální mocninná řada) a jak se definují základní operace, a poté se bude věnovat nejdůležitějším vlastnostem polynomů: dělení se zbytkem, souvislost kořenů s dělitelností, ukážeme, že za rozumných předpokladů má polynom stupně n nejvýše n kořenů, podíváme se, jak souvisí násobnost kořene daného polynomu s kořeny jeho derivací a na závěr zmíníme větu o interpolaci.*

5.1. Základní operace s polynomy.

Definice. *Polynomem proměnné x* nad oborem integrity $\mathbf{R}$ rozumíme formální výraz

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n,$$

nebo zkráceně

$$\sum_{i=0}^{n} a_ix^i,$$

kde $a_0, \ldots, a_n \in R$ a $a_n \neq 0$. Prvky $a_0, \ldots, a_n$ nazýváme *koeficienty* a symbol x *proměnná*. (Implicitně se rozumí se $a_m = 0$ pro všechna $m > n$.) Číslo n nazýváme *stupeň polynomu*, značíme $\deg f$. Prvek a_n se nazývá *vedoucí koeficient* a a_0 *absolutní člen*. Polynom se nazývá *monický*, pokud je vedoucí člen 1. Je třeba speciálně dodefinovat *nulový polynom*; pro něj položíme $\deg 0 = -1$.

Na množině všech polynomů definujeme operace předpisy

$$\sum_{i=0}^{m} a_i x^i \;+\; \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{\max(m,n)} (a_i + b_i)x^i, \qquad -\sum_{i=0}^{m} a_i x^i = \sum_{i=0}^{m} (-a_i)x^i,$$

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

Jak si za chvíli dokážeme, dostaneme obor integrity; značíme jej $\mathbf{R}[x]$.

Definice. *Formální mocninnou řadou proměnné x* nad oborem integrity $\mathbf{R}$ rozumíme formální výraz

$$\sum_{i=0}^{\infty} a_i x^i,$$

kde $a_0, a_1, \ldots \in R$; používáme obdobnou terminologii. Tedy polynom je mocninná řada, v níž je jen konečně mnoho nenulových koeficientů. Speciálně $0 = \sum_{i=0}^{\infty} 0x^i$. (Jde o *formální výrazy*, nikoliv o funkce nebo součty. Otázky typu konvergence nás nezajímají.)

Na množině všech formálních mocninných řad definujeme analogicky operace

$$\sum_{i=0}^{\infty} a_i x^i \;+\; \sum_{i=0}^{\infty} b_i x^i = \sum_{i=0}^{\infty} (a_i + b_i)x^i, \qquad -\sum_{i=0}^{\infty} a_i x^i = \sum (-a_i)x^i,$$

$$(\sum_{i=0}^{\infty} a_i x^i) \cdot (\sum_{i=0}^{\infty} b_i x^i) = \sum_{i=0}^{\infty} \big(\sum_{j+k=i} a_j b_k \big) x^i.$$

Jak si nyní dokážeme, dostaneme obor integrity; značíme jej $\mathbf{R}[[x]]$. Polynomy zřejmě tvoří jeho podobor, protože součet i součin dvou polynomů je opět polynom.

Tvrzení 5.1. *Je-li $\mathbf{R}$ obor integrity, pak $\mathbf{R}[x]$ i $\mathbf{R}[[x]]$ jsou také obory integrity.*

Důkaz. Důkaz stačí provést pro formální mocninné řady, protože polynomy jsou jejich speciálním případem (obecněji, podokruh oboru integrity je vždy oborem integrity).

Ověření rovností z definice komutativního okruhu je mechanická práce, ukážeme pouze hlavní myšlenky. Rovnosti pro sčítání jsou očividné, komutativita násobení také. Pro jednotku, součin $(\sum a_i x^i) \cdot (1 + 0 + 0 + \ldots)$ dává řadu $\sum (\sum_{j+k=i} a_j b_k) x^i$, kde všechny b_i kromě b_0 jsou nulové, takže výsledkem je opět $\sum a_i x^i$. Asociativita je obtížnější: z jedné strany $(\sum a_i x^i) \cdot ((\sum b_i x^i) \cdot (\sum c_i x^i)) = (\sum a_i x^i) \cdot ((\sum (\sum_{k+l=i} b_k c_l) x^i) = \sum (\sum_{j+k+l=i} a_j b_k c_l) x^i)$, a je vidět, že stejně vyjde i analogický výpočet součinu $((\sum a_i x^i) \cdot (\sum b_i x^i)) \cdot (\sum c_i x^i)$. Distributivita se prověří podobně.

Zajímavější je důkaz, že pro $f, g \neq 0$ je $f \cdot g \neq 0$. Buď $f = \sum a_i x^i$ a $g = \sum b_i x^i$ dva nenulové prvky $\mathbf{R}[[x]]$ a označme m, n nejmenší indexy takové, že $a_m, b_n \neq 0$. Uvažujeme-li v součinu $f \cdot g$ koeficient u x^{m+n}, dostáváme vyjádření

$$\sum_{j+k=m+n} a_j b_k = \underbrace{a_0 b_{m+n} + \ldots + a_{m-1} b_{n+1}}_{0} + \underbrace{a_m b_n}_{\neq 0} + \underbrace{a_{m+1} b_{n-1} + \ldots + a_{m+n} b_0}_{0}.$$

Protože je $\mathbf{R}$ obor integrity a $a_m, b_n \neq 0$, tak také $a_m b_n \neq 0$ a tento koeficient je nenulový. □

Obory *polynomů a mocninných řad více proměnných* se definují induktivně, předpisy

$$\mathbf{R}[x_1,\dots,x_n] = (\mathbf{R}[x_1,\dots,x_{n-1}])[x_n],$$
$$\mathbf{R}[[x_1,\dots,x_n]] = (\mathbf{R}[[x_1,\dots,x_{n-1}]])[[x_n]].$$

Polynom f z $\mathbf{R}[x_1,\dots,x_n]$ je výraz tvaru $f = \sum_{i=0}^{\infty} f_i x_n^i$, kde f_i jsou polynomy z $\mathbf{R}[x_1,\dots,x_{n-1}]$. Za pomocí distributivity jej můžeme přepsat (právě jedním způsobem) do standardního tvaru

$$f = \sum_{k_1,\dots,k_n=0}^{N} a_{k_1,\dots,k_n} x_1^{k_1}\cdots x_n^{k_n}$$

s koeficienty $a_{k_1,\dots,k_n} \in R$. Podobně pro mocninné řady. Z Tvrzení 5.1 za pomoci indukce ihned plyne, že jde o obory integrity.

5.2. **Hodnota polynomu v bodě.**

Je třeba striktně rozlišovat mezi polynomem jako *formálním výrazem* a jeho *hodnotou po dosazení* nějakého prvku. Formálně, buď $R \le S$ obory integrity. Polynom $f \in R[x]$ je formální výraz

$$f = a_0 + a_1 x + \ldots + a_n x^n$$

(tento se bude zapisovat výhradně f, bez uvedení proměnné). Jeho hodnotou po dosazení prvku $u \in S$ rozumíme prvek

$$f(u) = a_0 + a_1 u + \ldots + a_n u^n \in S,$$

přičemž v uvedeném zápise provádíme všechny operace (mocnění, násobení i sčítání) v oboru $\mathbf{S}$. Např. pro $\mathbf{R} = \mathbf{S} = \mathbb{Z}_p$ a $f = x^p + 1$ platí $f(0) = 1$, $f(1) = 2$, $f(2) = 3$ atd., viz malá Fermatova věta.

Pro daný polynom $f \in R[x]$ a obor $S \ge R$ můžeme uvažovat tzv. *polynomiální zobrazení* $S \to S$, které každému prvku $u \in S$ přiřadí hodnotu $f(u)$. Různé polynomy mohou dávat stejná polynomiální zobrazení, např. výše uvedený polynom určuje na $\mathbb{Z}_p$ stejné zobrazení jako polynom $g = x + 1$. (Jinak to pro konečné obory být ani nemůže, protože existuje nekonečně mnoho polynomů, ale pouze konečně mnoho zobrazení na konečné množině.)

Pojem hodnoty mocninné řady nemá v algebře smysl uvažovat. Bez další geometrické struktury není možné říci, co se rozumí nekonečným součtem, v řadě oborů (třeba konečných) se smysluplný pojem konvergence ani nedá vybudovat.

5.3. **Dělení polynomů se zbytkem.**

Buď f, g polynomy z $\mathbf{R}[x]$. Řekneme, že g *dělí* f, píšeme $g \mid f$, pokud existuje polynom $h \in R[x]$ takový, že $f = gh$. Všimněte si, že pokud $g \mid f$ a $f \neq 0$, pak $\deg g \le \deg f$. (Každý polynom dělí nulový polynom, přitom stupeň nulového polynomu je -1.) Pokud g nedělí f, má smysl se ptát po zbytku po dělení.

Tvrzení 5.2. *Buď* $\mathbf{R}$ *obor integrity,* $\mathbf{Q}$ *jeho podílové těleso,* $f, g \in R[x]$, $g \neq 0$. *Pak existuje právě jedna dvojice* $q, r \in Q[x]$ *splňující* $f = gq + r$ *a* $\deg r < \deg g$. *Navíc, je-li* g *monický, pak* $q, r \in R[x]$.

Díky jednoznačnosti můžeme definovat $f \operatorname{div} g = q$ a $f \bmod g = r$. Je vidět, že $g \mid f$ právě tehdy, když $f \bmod g = 0$.

Důkaz. Podíl a zbytek dvou polynomů se počítá podobně jako pro celá čísla. Algoritmus lze formulovat takto: inicializujeme $q_0 = 0$, $r_0 = f$, a poté definujeme rekurzivně

$$q_{i+1} = q_i + \frac{l(r_i)}{l(g)} \cdot x^{\deg r_i - \deg g}, \quad r_{i+1} = r_i - \frac{l(r_i)}{l(g)} \cdot x^{\deg r_i - \deg g} \cdot g,$$

kde $l(u)$ značí vedoucí koeficient polynomu u. Rekurzí pokračujeme do té doby, než bude $\deg r_i$ menší než $\deg g$. To jistě někdy nastane, protože je vždy $\deg r_{i+1} < \deg r_i$. Přitom evidentně platí $f = gq_i + r_i$ pro všechna i, a tedy poslední dvojice q_i, r_i je hledaným podílem a zbytkem.

Z algoritmu je vidět, že je-li g monický, žádné zlomky se neobjeví a výsledkem budou polynomy z $\mathbf{R}[x]$

Jednoznačnost se dokáže podobně jako pro celá čísla. Kdyby $f = gq_1 + r_1 = gq_2 + r_2$, pak $g(q_1 - q_2) = r_2 - r_1$, tedy $g \mid r_2 - r_1$. Přitom $\deg(r_2 - r_1) < \deg g$, tedy $r_2 - r_1 = 0$, čili $r_1 = r_2$. Z toho ihned plyne $q_1 - q_2 = 0$, tj. $q_1 = q_2$, protože $g \neq 0$ a jsme v oboru integrity. □

5.4. **Kořeny a dělitelnost.**

Buď $R \leq S$ obory integrity, $f \in R[x]$ a $a \in R$. Řekneme, že a je *kořen* polynomu f, pokud $f(a) = 0$. Ukážeme si, jak existence kořene souvisí s děliteli daného polynomu.

Tvrzení 5.3. *Buď* $\mathbf{R}$ *obor integrity,* $f \in R[x]$ *a* $a \in R$*. Pak* a *je kořen polynomu* f *právě tehdy, když* $x - a \mid f$*.*

Důkaz. ($\Leftarrow$) Předpokládejme, že $x - a \mid f$. Pak $f = (x - a) \cdot g$ pro nějaké $g \in R[x]$ a dosadíme-li do f prvek a, dostaneme

$$f(a) = (a - a) \cdot g(a) = 0 \cdot g(a) = 0.$$

($\Rightarrow$) Buďte q, r podíl a zbytek při dělení polynomu f polynomem $x - a$ (ty existují, neboť dělíme monickým polynomem). Tedy $f = (x - a) \cdot q + r$ a r je konstantní polynom (zbytek musí mít menší stupeň než dělitel). Dosadíme-li prvek a, dostaneme

$$0 = f(a) = (a - a) \cdot q(a) + r(a) = 0 \cdot q(a) + r = r,$$

takže $r = 0$ a $x - a \mid f$. □

Z důkazu plyne jedno důležité pozorování: je-li $f \in R[x]$ a $a \in R$, pak

$$f \bmod (x - a) = f(a).$$

Věta 5.4. *Buď* $\mathbf{R}$ *obor integrity,* $0 \neq f \in R[x]$ *a* $\deg f = n$*. Pak má polynom* f *nejvýše* n *kořenů.*

Důkaz. Budeme postupovat indukcí podle stupně polynomu f. Je-li $\deg f = 0$, tj. f je nenulový konstantní polynom, pak žádné kořeny nemá. Nyní předpokládejme, že tvrzení platí pro všechny polynomy stupně nejvýše n. Je-li $\deg f = n + 1$, pak jsou dvě možnosti. Buď polynom f nemá žádný kořen, v tom případě tvrzení platí. Nebo má polynom f nějaký kořen a a v tom případě jej lze podle předchozího lemmatu napsat jako $f = (x - a) \cdot g$ pro nějaký polynom g stupně n. Je-li b nějaký jiný kořen, tj. $f(b) = (b - a) \cdot g(b) = 0$, pak, protože jde o obor integrity, musí být buď $b = a$ nebo $g(b) = 0$. Protože má polynom g nejvýše n kořenů, má polynom f nejvýše $n + 1$ kořenů. □

Příklad. Počet kořenů polynomu f samozřejmě může být menší než $\deg f$: např. polynom x^2+1 nemá nad $\mathbb{Z}$ žádný kořen a nad $\mathbb{Z}_2$ má jeden.

Poznámka. Věta 5.4 neplatí, není-li $\mathbf{R}$ oborem integrity, ale např. jen komutativním okruhem s jednotkou. Předpoklad jsme použili v poslední fázi důkazu, když z $f(b) = (b-a)\cdot g(b) = 0$ plynulo $b-a=0$ nebo $g(b)=0$. Uvažte např. polynom $2x \in \mathbb{Z}_4[x]$ nebo $x^2+x \in \mathbb{Z}_6[x]$. První z nich má kořeny $0, 2$, druhý $0, 2, 3, 5$.

Poznámka. Věta 5.4 neplatí, není-li $\mathbf{R}$ oborem integrity, ale např. jen nekomutativním tělesem – celá teorie dělitelnosti funguje jinak. Příkladem je polynom x^4-1 nad okruhem kvaternionů, jeho kořeny jsou ± 1, $\pm i$, $\pm j$, $\pm k$.

5.5. **Derivace a vícenásobné kořeny.**

Matematická analýza zavádí pojem *derivace* reálné funkce, tedy speciálně také polynomu nad reálnými čísly. V oboru reálných čísel má derivace jistý geometrický význam (tečna grafu) a tak se také definuje (pomocí limit). Pro polynomy se z této definice odvodí jistý vzorec, ve kterém figurují koeficienty původního polynomu. V diskrétních oborech se geometrická představa ztrácí (co je tečna grafu funkce na celých číslech?), ale přesto má smysl derivaci zavést, a to tak, že postulujeme základní vlastnosti, které derivace splňuje.

Definice. Definujeme *derivaci* v $\mathbf{R}[x]$ jako zobrazení $D : R[x] \to R[x]$ splňující následující podmínky pro všechny polynomy $f, g \in R[x]$:

(1) $D(f+g) = D(f) + D(g)$;
(2) $D(fg) = gD(f) + fD(g)$;
(3) $D(x) = 1$, $D(c) = 0$ pro každý konstantní polynom c.

Derivaci polynomu zpravidla značíme zkráceně $f' = D(f)$. Dále definujeme induktivně derivace vyšších řádů jako

$$f^{(0)} = f \quad \text{a} \quad f^{(k+1)} = (f^{(k)})'.$$

Než ukážeme vzorec na výpočet derivace, musíme si ujasnit, co značí v obecném oboru $\mathbf{R}$ přirozená čísla. Pod přirozeným číslem n budeme rozumět prvek

$$\underbrace{1+1+\ldots+1}_{n} \in R.$$

Připomeňme, že *charakteristikou* oboru $\mathbf{R}$ rozumíme nejmenší přirozené číslo n takové, že v $\mathbf{R}$ platí $n=0$, pokud takové n existuje, resp. 0 v opačném případě.

Lemma 5.5. *Pro každý obor integrity* $\mathbf{R}$ *existuje právě jedna derivace na* $\mathbf{R}[x]$ *a platí*

$$\left(\sum_{i=0}^{n} a_i x^i\right)' = \sum_{i=0}^{n-1} (i+1) a_{i+1} x^i.$$

Důkaz. Nejprve si všimněte, že z (2) plyne $(cf)' = cf' + fc' = cf'$ pro každý polynom f a každý konstantní polynom c. Dále indukcí dokážeme, že $(x^n)' = nx^{n-1}$. Případ $n=1$ je pokryt vlastností (3) a dále, pomocí (2) a indukčního předpokladu, $(x^n)' = x(x^{n-1})' + x^{n-1}x' = x(n-1)x^{n-2} + x^{n-1} = nx^{n-1}$. Na závěr použijeme (1) a vidíme, že $(\sum_{i=0}^n a_i x^i)' = \sum_{i=0}^n (a_i x^i)' = \sum_{i=0}^n a_i (x^i)' = \sum_{i=0}^{n-1} (i+1) a_{i+1} x^i$. □

Lemma 5.6. *Buď* $\mathbf{R}$ *obor integrity,* $f, g \in R[x]$ *a* $n \in \mathbb{N}$. *Pak*

(1) $(f+g)^{(n)} = f^{(n)} + g^{(n)}$;

(2) $(f \cdot g)^{(n)} = \sum_{i=0}^{n} \binom{n}{i} \cdot f^{(i)} \cdot g^{(n-i)}$ [Leibnitzova formule]*;*
(3) $(f^n)' = n \cdot f^{n-1} \cdot f'$.

Důkaz je pouze technický výpočet a doporučujeme čtenáři jej provést samostatně. Níže je uveden stručný návod.

Princip důkazu. (1) Indukcí podle n. Pro $n = 1$ viz definice. Indukční krok plyne z výpočtu $(f+g)^{(n)} = ((f+g)^{(n-1)})' = (f^{(n-1)} + g^{(n-1)})' = (f^{(n-1)})' + (g^{(n-1)})' = f^{(n)} + g^{(n)}$.

(2) Indukcí podle n. Pro $n = 1$ viz definice. V indukčním kroku využijte známý vzorec $\binom{n}{i} + \binom{n}{i+1} = \binom{n+1}{i+1}$.

(3) se dokáže snadno indukcí podle n pomocí (2). □

Tvrzení 5.3 umožňuje definovat násobnost kořene daného polynomu.

Definice. Řekneme, že $a \in R$ je *n-násobný kořen* polynomu $f \in R[x]$, pokud

$$(x-a)^n \mid f \quad \text{a} \quad (x-a)^{n+1} \nmid f.$$

Násobnost kořene daného polynomu úzce souvisí s kořeny derivací tohoto polynomu. Vztah popisuje následující věta. Její hlavní význam spočívá v tom, že umožňuje výpočetně podchytit pojem násobnosti kořene.

Věta 5.7. *Buď $\mathbf{R}$ obor integrity, $0 \neq f \in R[x]$, $a \in R$ a $n \in \mathbb{N}$. Předpokládejme, že charakteristika oboru $\mathbf{R}$ je buď 0, nebo $\geq n$. Pak jsou následující tvrzení ekvivalentní:*

(1) *a je alespoň n-násobný kořen polynomu f;*
(2) $f^{(0)}(a) = f^{(1)}(a) = \ldots = f^{(n-1)}(a) = 0$.

Důkaz. (1) ⇒ (2) Je-li a alespoň n-násobný kořen polynomu f, můžeme napsat

$$f = (x-a)^n \cdot g$$

pro nějaký polynom $g \in R[x]$. Pomocí Leibnitzovy formule spočítáme k-tou derivaci polynomu f pro $k < n$:

$$\begin{aligned} f^{(k)} &= \sum_{i=0}^{k} \binom{k}{i} \cdot ((x-a)^n)^{(i)} \cdot g^{(k-i)} \\ &= \sum_{i=0}^{k} \binom{k}{i} \cdot n(n-1) \cdot \ldots \cdot (n-i+1) \cdot (x-a)^{n-i} \cdot g^{(k-i)}. \end{aligned}$$

Protože $k < n$, v každém členu součtu je člen $x - a$ v nenulové mocnině, a tak dostáváme

$$f^{(k)}(a) = \sum_{i=0}^{k} 0 = 0.$$

(2) ⇒ (1) Protože $f^{(0)}(a) = f(a) = 0$, prvek a je kořenem polynomu f. Buď m jeho násobnost a pro spor předpokládejme, že $m < n$. Napišme

$$f = (x-a)^m \cdot g$$

pro nějaký polynom $g \in R[x]$ splňující $g(a) \neq 0$. Pomocí Leibnitzovy formule spočítáme m-tou derivaci polynomu f:

$$\begin{aligned} f^{(m)} &= \sum_{i=0}^{m} \binom{m}{i} \cdot ((x-a)^m)^{(i)} \cdot g^{(m-i)} \\ &= \binom{m}{m} \cdot m! \cdot g^{(0)} + \sum_{i=0}^{m-1} \binom{m}{i} \cdot m(m-1) \cdot \ldots \cdot (m-i+1) \cdot (x-a)^{m-i} \cdot g^{(m-i)} \end{aligned}$$

a po dosazení dostaneme

$$f^{(m)}(a) = 1 \cdot m! \cdot g(a) + \sum_{i=0}^{m-1} 0 = m! \cdot g(a).$$

Podle předpokladu $f^{(m)}(a) = 0$. Protože pracujeme v oboru integrity, musí platit $m! = 0$ nebo $g(a) = 0$. Druhá možnost je ve sporu s volbou g, takže $m! = m \cdot (m-1) \cdot \ldots \cdot 1 = 0$. Tedy některý z prvků $1, \ldots, m$ musí být roven nule, což je ve sporu s předpokladem na charakteristiku oboru $\mathbf{R}$. □

Z věty ihned plyne následující kritérium pro určení přesné násobnosti. Je-li charakteristika 0 nebo $> n$, pak a je (přesně) n-násobným kořenem polynomu f právě tehdy, když $f^{(0)}(a) = f^{(1)}(a) = \ldots = f^{(n-1)}(a) = 0$ a navíc $f^{(n)}(a) \neq 0$. Důvod je jednoduchý: kdyby $f^{(n)}(a) = 0$, šlo by o alespoň $(n+1)$-násobný kořen. Všimněte si však, že k tomuto argumentu pořebujeme charakteristiku ostře větší než n.

Příklad. Větu 5.7 lze použít k detekci vícenásobných kořenů i nad tělesem $\mathbb{Z}_2$, zatímco právě uvedené kritérium nikoliv: polynom $f \in \mathbb{Z}_2[x]$ má vícenásobný (tj. alespoň 2-násobný) kořen a právě tehdy, když $f(a) = f'(a) = 0$. Pro přesně dvojnásobné kořeny však nemusí být pravda, že $f''(a) \neq 0$: např. pro $f = x^3+x = x(x+1)^2$ je 1 dvojnásobným kořenem, avšak $f' = x^2 + 1$, $f'' = 0$, tedy $f''(1) = 0$.

Úloha. Spočtěte násobnost kořene 1 polynomu $f = x^4+x^3+x^2+x+1$ nad tělesem $\mathbb{Z}_5$.

Řešení. Nejprve si uvědomíme, že můžeme použít Větu 5.7, protože f má stupeň 4, tedy 1 bude nejvýše 4-násobným kořenem, což je méně než charakteristika oboru $\mathbb{Z}_5$. Postupně spočteme $f(1) = 0$; $f' = 4x^3+3x^2+2x+1$, tedy $f'(1) = 0$; $f'' = 2x^2+x+2$, tedy $f''(1) = 0$; $f''' = 4x+1$, tedy $f'''(1) = 0$; a nakonec $f'''' = 4$. Čili 1 je 4-násobný kořen. (Roznásobením snadno ověříme, že $(x-1)^4 = f$, což nás mohlo, ale nemuselo napadnout hned na začátku.) □

Z Věty 5.7 plyne důležité kritérium existence vícenásobného kořene daného polynomu. Je-li $\mathbf{R}$ obor integrity a f polynom nad $\mathbf{R}$, pak prvek $a \in R$ je vícenásobným kořenem polynomu f právě tehdy, když $f(a) = f'(a) = 0$. Tedy, má-li f vícenásobný kořen, oba polynomy f, f' jsou dělitelné nějakým monočlenem $x-a$, a pokud existuje jejich největší společný dělitel, pak jej $x-a$ dělí také. Tím dostáváme algoritmicky snadno ověřitelné kritérium: jsou-li f, f' nesoudělné, polynom f zaručeně žádný vícenásobný kořen nemá.

5.6. **Věta o interpolaci.**

S kořeny polynomů souvisí tzv. *interpolace*: předepíšeme-li hodnoty v n bodech, existuje právě jeden polynom stupně $< n$, který v těchto bodech nabývá daných hodnot.

Věta 5.8 (o interpolaci)**.** *Buď* $\mathbf{T}$ *těleso. Mějme po dvou různé body* $a_1, \ldots, a_n \in T$ *a libovolné hodnoty* $u_1, \ldots, u_n \in T$*. Pak existuje právě jeden polynom* $f \in T[x]$ *stupně* $< n$ *splňující* $f(a_i) = u_i$ *pro všechna* $i = 1, \ldots, n$*.*

Není těžké nahlédnout, že řešením je polynom

$$f = \sum_{i=1}^{n} \left(u_i \cdot \prod_{j \neq i} \frac{x - a_j}{a_i - a_j} \right),$$

říká se mu někdy *Lagrangeův interpolační polynom.*

Důkaz. Dosazením do uvedeného vzorce snadno zjistíme, že

$$f(a_k) = 0 + \ldots + 0 + u_k \cdot \prod_{j \neq k} \frac{a_k - a_j}{a_k - a_j} + 0 + \ldots + 0 = u_k.$$

Zbývá dokázat jednoznačnost. Uvažujme dva polynomy f, g stupně $< n$ splňující $f(a_i) = g(a_i) = u_i$ pro všechna $i = 1, \ldots, n$. Polynom $h = f - g$ je také stupně $< n$, avšak $h(a_i) = f(a_i) - g(a_i) = 0$ pro všechna i, tedy h má aspoň n kořenů. To je spor s Větou 5.4. □

Důkaz věty o interpolaci nápadně připomíná důkaz čínské věty o zbytcích. Ve skutečnosti je velmi podobné i znění věty: podmínku $f(a_i) = u_i$ lze napsat ekvivalentně jako $f \equiv u_i \pmod{x - a_i}$, takže vlastně řešíme soustavu kongruencí vzhledem k polynomům $x - a_1, \ldots, x - a_n$. Řešení je určeno jedoznačně mezi polynomy omezeného stupně. Věta o interpolaci a čínská věta o zbytcích mají společné zobecnění, které je předmětem Sekce **??**.

Důsledek 5.9. *Buď* $\mathbf{T}$ *konečné těleso. Pak pro každé zobrazení* $\varphi : T \to T$ *existuje právě jeden polynom* $f \in T[x]$ *stupně* $< |T|$ *takový, že* $f(a) = \varphi(a)$ *pro každé* $a \in T$*.*

Důkaz. Interpolujme v bodě a hodnotou $\varphi(a)$ pro každé $a \in T$. □

Pro *nekonečná* tělesa samozřejmě nic takového platit nemůže, přesto polynomy hrají důležitou roli i v reálné analýze: např. *Weierstrassova věta* říká, že každou spojitou reálnou funkci na omezeném uzavřeném intervalu lze polynomem libovolně přesně aproximovat, tj. pro každou spojitou $f : [u, v] \to \mathbb{R}$ a každé $\varepsilon > 0$ existuje polynom $g \in \mathbb{R}[x]$ takový, že $|f(a) - g(a)| < \varepsilon$ pro každé $a \in [u, v]$.

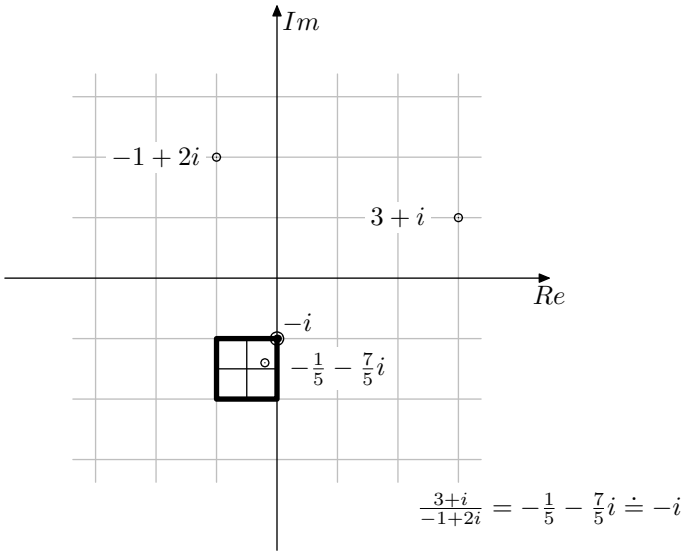
6. Kvadratická rozšíření celých čísel

Cíl. *Ukážeme si základní triky pro počítání v oborech* $\mathbb{Z}[\sqrt{s}]$*. Zvláštní pozornost bude věnována Gaussovým celým číslům.*

Mezi nejdůležitější rozšíření oboru celých čísel patří tzv. *kvadratická rozšíření.* Zde se soustředíme na obory $\mathbb{Z}[\sqrt{s}]$, pro obecnější teorii doporučujeme libovolnou knihu o algebraické teorii čísel. Buď s číslo, jež není dělitelné druhou mocninou žádného prvočísla, a definujme zobrazení

$$\nu : \mathbb{Z}[\sqrt{s}] \to \mathbb{N} \cup \{0\}, \qquad a + b\sqrt{s} \mapsto |a^2 - sb^2|.$$

Je dobré mít na paměti, že pro $s < 0$ je $\nu(u) = |u|^2$, čtverec obyčejné absolutní hodnoty komplexního čísla, díky čemuž se dá často aplikovat geometrický náhled na situaci. Zobrazení ν nazýváme *normou.* Základním pozorováním je fakt, že norma se chová hezky vzhledem k dělitelnosti.



Obrázek 6. Dělení se zbytkem v $\mathbb{Z}[i]$.

Tvrzení 6.1. *Pro každá $u, v \in \mathbb{Z}[\sqrt{s}]$ platí*

(1) $\nu(u \cdot v) = \nu(u) \cdot \nu(v)$,

(2) $\nu(u) = 1 \Leftrightarrow$ *u je* invertibilní, *tj. existuje $w \in \mathbb{Z}[\sqrt{s}]$ takové, že $uw = 1$.*

Důkaz. (1) Označme $u = a + b\sqrt{s}$ a $v = c + d\sqrt{s}$. Pak

$$\begin{aligned}\nu(u \cdot v) &= \nu((ac + sbd) + (ad + bc)\sqrt{s}) \\ &= |a^2c^2 + 2sabcd + s^2b^2d^2 - s(a^2d^2 + 2abcd + b^2c^2)| \\ &= |a^2c^2 + s^2b^2d^2 - sa^2d^2 - sb^2c^2)| \\ &= |a^2 - sb^2| \cdot |c^2 - sd^2| = \nu(u) \cdot \nu(v).\end{aligned}$$

(2) Pokud $\nu(u) = \nu(a+b\sqrt{s}) = |a^2 - sb^2| = 1$, pak $a^2 - sb^2 = (a+b\sqrt{s})(a-b\sqrt{s}) = \pm 1$, a tedy $w = \pm(a - b\sqrt{s})$. Opačná implikace plyne z (1): je-li $uw = 1$, pak $1 = \nu(1) = \nu(uw) = \nu(u)\nu(w)$, a tedy $\nu(u) = \nu(w) = 1$. □

Pro některé obory $\mathbb{Z}[\sqrt{s}]$ umožňuje norma definovat *dělení se zbytkem*. Fakt, že zbytek by měl být „menší“ než dělitel, formalizujeme pomocí normy. Dělení se zbytkem funguje např. pro obory $\mathbb{Z}[i]$, $\mathbb{Z}[i\sqrt{2}]$ nebo $\mathbb{Z}[\sqrt{2}]$, pro jiné podíl a zbytek v tomto smyslu neexistuje, např. pro $\mathbb{Z}[i\sqrt{3}]$ nebo $\mathbb{Z}[\sqrt{5}]$. Důkaz provedeme pro Gaussova celá čísla.

Tvrzení 6.2. *Pro každá $u, v \in \mathbb{Z}[i]$, $v \neq 0$, existují $q, r \in \mathbb{Z}[i]$ splňující podmínky $u = vq + r$ a $\nu(r) < \nu(v)$.*

Důkaz. Položme

$$z = \frac{u}{v} \in \mathbb{C}$$

(přesný podíl v $\mathbb{C}$). Buď q nejbližší prvek $\mathbb{Z}[i]$ k prvku z (tj. takový, pro který je $|z - q|$ minimální); je-li takových více, zvolme libovolný z nich. Položme

$$r = u - vq.$$

Pak zřejmě $vq + r = u$ a zbývá dokázat, že $\nu(r) < \nu(v)$. Jaká je vzdálenost q a z? V nejhorším případě je z uprostřed čtverce s celočíselnými vrcholy, tedy určitě $|z - q| \leq \frac{\sqrt{2}}{2} < 1$. Proto

$$\nu(r) = |r|^2 = |u - vq|^2 = |v|^2 \cdot |\frac{u}{v} - q|^2 = |v|^2 \cdot |z - q|^2 < |v|^2 = \nu(v).$$

□

Na rozdíl od situace v $\mathbb{Z}$ či pro polynomy (viz Tvrzení 5.2), podíl a zbytek q, r není určen jednoznačně: např. $z = \frac{1}{2} + \frac{1}{2}i$ lze zaokrouhlit čtyřmi způsoby, každý z nich bude splňovat uvedené podmínky.

Pro obory $\mathbb{Z}[i\sqrt{2}]$ či $\mathbb{Z}[e^{2\pi i/3}]$ lze důkaz provést zcela analogicky, protože i zde platí $\nu(u) = |u|^2$ a jediný rozdíl tak je v odhadu $|z - q|$. Pro $\mathbb{Z}[i\sqrt{3}]$ už důkaz neprojde, protože střed obdélníka má vzdálenost od vrcholu rovnou 1. (Ve skutečnosti v tomto oboru není možné dělit se zbytkem žádným způsobem. Tato teorie je předmětem následujících dvou sekcí.) Pro obory $\mathbb{Z}[\sqrt{s}]$ s kladným s schází geometrická představa, nicméně pro $s = 2, 3$ funguje podobný algoritmus dělení, stačí zaokrouhlit koeficienty přesného podílu. Důkaz odhadu normy zbytku je však o něco komplikovanější.

7. Základní pojmy teorie dělitelnosti

Cíl. *Ujasníme si, které prvky jsou z hlediska dělitelnosti nerozlišitelné (relace asociovanosti, souvislost s invertibilními prvky), což nám umožní na relaci dělitelnosti pohlížet jako na uspořádání. Zavedeme největší společný dělitel a definujeme analogii k pojmu prvočísla, tzv. ireducibilní prvky.*

V celé sekci budeme uvažovat nějaký pevně daný obor integrity $\mathbf{R}$.

7.1. Invertibilní prvky.

Definice. Řekneme, že a *dělí* b v oboru $\mathbf{R}$ (píšeme $a \mid b$), pokud existuje $c \in R$ takové, že $b = ac$. Řekneme, že prvky a a b jsou *asociované* (píšeme $a \parallel b$), pokud $a \mid b$ a $b \mid a$. Prvek a se nazývá *invertibilní*, pokud $a \parallel 1$, tj. existuje b takové, že $ab = 1$; toto b obvykle značíme a^{-1}. Dělitel prvku a se nazývá *vlastní*, jestliže není asociovaný ani s 1, ani s a.

Tvrzení 7.1. *Dva prvky a, b jsou asociované právě tehdy, když existuje invertibilní prvek q takový, že $a = bq$.*

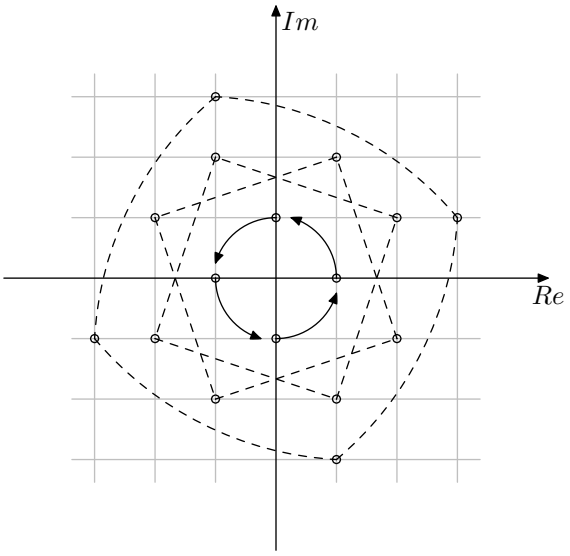
Důkaz. ($\Leftarrow$) Protože $a = bq$, platí $b \mid a$. Protože taky $b = aq^{-1}$, platí $a \mid b$.

($\Rightarrow$) Protože $b \mid a$, můžeme psát $a = bu$, a protože $a \mid b$, můžeme psát $b = av$, pro nějaká u, v. Tedy $a = bu = avu$ a krácením dostáváme $uv = 1$, čili $u, v \parallel 1$. □

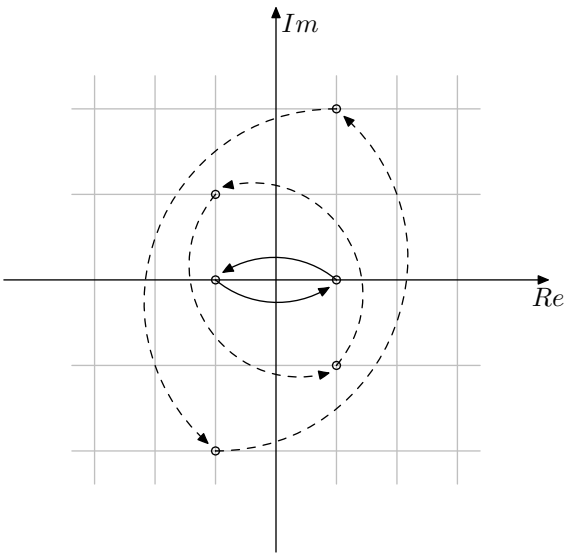
Příklady.

- V tělese je každý nenulový prvek invertibilní. Tedy $a \parallel b$ pro každé $a, b \neq 0$.
- V oboru $\mathbb{Z}$ jsou invertibilní pouze prvky ± 1. Tedy $a \parallel b$ právě tehdy, když $a = \pm b$.
- V oboru $\mathbf{R}[x]$ jsou invertibilní právě polynomy stupně 0, jejichž člen je invertibilní v oboru $\mathbf{R}$.

Příklad. Pozor na následující záludnost!



Obrázek 7. Asociovanost v $\mathbb{Z}[i]$.



Obrázek 8. Asociovanost v $\mathbb{Z}[i\sqrt{2}]$.

- $3x+6 \parallel x+2$ v oboru $\mathbb{Q}[x]$, protože $3x+6 = 3\cdot(x+2)$ a $x+2 = \frac{1}{3}\cdot(3x+6)$;
- $3x+6 \nparallel x+2$ v oboru $\mathbb{Z}[x]$, protože $\frac{1}{3} \notin \mathbb{Z}[x]$.

Při použití symbolu pro dělitelnost, asociovanost, apod. musíme vždy uvést, v jakém oboru pracujeme (není-li to zřejmé z kontextu).

K určení invertibilních prvků v oborech $\mathbb{Z}[\sqrt{s}]$ lze použít Tvrzení 6.1.

Příklady. Podmínku (2) lze s úspěchem využít pro hledání *invertibilních* prvků.

- V oboru $\mathbb{Z}[i]$ máme $\nu(a+bi) = a^2+b^2$, tedy
$$\nu(u) = 1 \;\Leftrightarrow\; u = \pm 1, u = \pm i.$$
Tedy $u \parallel v$ právě tehdy, když $u = \pm v$ nebo $u = \pm iv$.
- V oboru $\mathbb{Z}[i\sqrt{2}]$ máme $\nu(a+bi\sqrt{2}) = a^2+2b^2$, tedy
$$\nu(u) = 1 \;\Leftrightarrow\; u = \pm 1.$$
Tedy $u \parallel v$ právě tehdy, když $u = \pm v$.
- V oboru $\mathbb{Z}[\sqrt{2}]$ máme $\nu(a+b\sqrt{2}) = |a^2-2b^2|$. Řešením rovnice $\nu(u) = 1$ je např. ± 1, ale také $\pm 1 \pm \sqrt{2}$, $\pm 3 \pm 2\sqrt{2}$, atd. Všimněte si, že je-li u invertibilní, pak u^k je také invertibilní, pro libovolné $k \in \mathbb{N}$, inverzním prvkem bude $(u^{-1})^k$. Tedy obor $\mathbb{Z}[\sqrt{2}]$ obsahuje nekonečně mnoho invertibilních prvků $(1+\sqrt{2})^k$ pro libovolné k.

7.2. **Dělitelnost jako uspořádání.**

Uvažujme na množině R relaci dělitelnosti. Je reflexivní: $a \mid a$, protože $a = a \cdot 1$. Je tranzitivní, protože pokud $a \mid b$ a $b \mid c$, tj. $b = ax$ a $c = by$, pak $c = a(xy)$, tedy $a \mid c$. Z toho ihned plyne následující pozorování:

Pozorování 7.2. *Relace $\parallel$ je ekvivalence na množině R.*

K tomu, aby byla relace $\mid$ uspořádání, chybí antisymetrie. Ta téměř nikdy splněna není, neboť v každém oboru platí $1 \mid -1$ a zároveň $-1 \mid 1$. (Výjimkou jsou obory charakteristiky 2, kde $1 = -1$, jako např. v oboru $\mathbb{Z}_2[x]$.) Tuto vadu lze napravit tak, že z každého bloku ekvivalence $\parallel$ na množině R vybereme po jednom zástupci. Označíme-li množinu takto vybraných prvků $\bar{R}$, pak $(\bar{R}, |)$ je uspořádanou množinou.

Volbu množiny $\bar{R}$ můžeme provést mnoha způsoby. V některých oborech však existuje přirozený výběr, proto se zavádějí následující konvence:

Příklady.

- V tělese $\mathbf{T}$ má ekvivalence $\parallel$ pouze dva bloky: $\{0\}$ a $T \setminus \{0\}$. Proto můžeme zvolit např. $\bar{T} = \{0, 1\}$.
- V oboru $\mathbb{Z}$ z dvou asociovaných čísel vybereme to nezáporné, tj. $\bar{\mathbb{Z}} = \mathbb{N} \cup \{0\}$.
- V oboru $\mathbb{Z}[i]$ ze čtyřech asociovaných čísel vybereme to $a+bi$, kde $a > 0$, $b \geq 0$ (resp. nulu ve svém bloku).
- V oboru $\mathbb{Z}[x]$ z dvou asociovaných polynomů vybereme ten s nezáporným vedoucím koeficientem (resp. nulový polynom ve svém bloku).
- V oboru $\mathbf{T}[x]$, $\mathbf{T}$ těleso, volíme z navzájem asociovaných polynomů ten monický (resp. nulový polynom ve svém bloku).

7.3. **Největší společný dělitel.**

Definice. Řekneme, že $c = \mathrm{NSD}(a, b)$ (*největší společný dělitel*), pokud

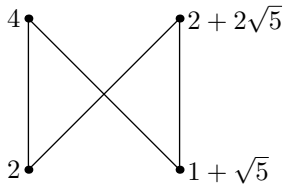
(1) $c \mid a$ a $c \mid b$ (tj. c je společný dělitel);
(2) kdykoliv $d \mid a$ a $d \mid b$, pak $d \mid c$ (tj. c je největší takový).

Řekneme, že $c = \mathrm{NSN}(a, b)$ (*nejmenší společný násobek*), pokud

$$c \cdot \mathrm{NSD}(a, b) = a \cdot b.$$

Prvky a, b nazýváme *nesoudělné*, pokud $\mathrm{NSD}(a, b) = 1$.

NSD a NSN není určen jednoznačně (pokud vůbec existuje). Například,



Obrázek 9. V $\mathbb{Z}[\sqrt{5}]$ neexistuje $\mathrm{NSD}(4, 2+2\sqrt{5})$.

- v oboru $\mathbb{Z}$ platí $\mathrm{NSD}(4,10) = 2$, ale také $\mathrm{NSD}(4,10) = -2$,
- v oboru $\mathbb{Q}[x]$ platí $\mathrm{NSD}(x^2-2x+1, x^2-1) = x-1$, ale zrovna tak třeba $\mathrm{NSD}(x^2-2x+1, x^2-1) = -5x+5$.

Na jednu stranu stranu, pokud $\mathrm{NSD}(a,b) = c$ a $\mathrm{NSD}(a,b) = d$, pak c i d jsou společní dělitelé a, b, a tedy $c \mid d$ a zároveň $d \mid c$, tedy $c \parallel d$. Na druhou stranu, pokud $\mathrm{NSD}(a,b) = c$ a $c \parallel d$, pak d jistě také splňuje podmínky největšího společného dělitele. Čili NSD a NSN jsou určené *jednoznačně až na asociovanost.*

Operátory NSD a NSN se obvykle používají ve významu funkce dvou parametrů. Jednoznačnosti lze dosáhnout trikem popsaným v předchozím odstavci: máme-li dánu množinu $\bar{R}$, pak definujeme hodnotu $\mathrm{NSD}(a,b)$ jako to jediné $c \in \bar{R}$ splňující $\mathrm{NSD}(a,b) = c$. Pro představu je šikovné mít na paměti, že

$$\mathrm{NSD}(a,b) = \inf\{a,b\} \qquad \text{a} \qquad \mathrm{NSN}(a,b) = \sup\{a,b\},$$

kde sup a inf se rozumí v uspořádané množině $(\bar{R}, |)$.

Pozor! V některých oborech nemusí NSD a NSN pro danou dvojici prvků vůbec existovat. Uvažujme obor $\mathbb{Z}[\sqrt{5}]$ a prvky $x = 4$ a $y = 2+2\sqrt{5}$, viz obrázek 7.3. Čísla 2 a $1+\sqrt{5}$ jsou určitě společnými děliteli, protože $4 = 2 \cdot 2 = (1+\sqrt{5})(1-\sqrt{5})$. Dokážeme si, že jsou to *maximální* vlastní dělitele (ve smyslu uspořádání dělitelností), přitom jsou navzájem neasociované, a tedy žádný *největší* společný dělitel neexistuje. Z Tvrzení 6.1 snadno plynou následující dvě fakta pro $u, v \in \mathbb{Z}[\sqrt{s}]$:

(1) Pokud $u \mid v$, pak $\nu(u) \mid \nu(v)$.
(2) Je-li u vlastní dělitel prvku v, pak $1 < \nu(u) < \nu(v)$.

Uvažujme největšího společného dělitele $z = \mathrm{NSD}(x,y)$. Z (1) plyne, že $\nu(z) \mid \nu(x) = \nu(y) = 16$. Protože $2 \mid z$ a $1+\sqrt{5} \mid z$, musí platit $\nu(2) = \nu(1+\sqrt{5}) = 4 \mid \nu(z)$. Protože $x \nparallel y$ a $2 \nparallel 1+\sqrt{5}$, ve všech uvedených případech musí jít o vlastní dělitele, díky (2) musí platit $\nu(z) = 8$. Napišme si $z = 2w$ pro nějaké w, a spočítejme normu: $8 = \nu(z) = \nu(2)\nu(w) = 4\nu(w)$, a tedy $\nu(w) = 2$. Ale žádný prvek normy 2 v $\mathbb{Z}[\sqrt{5}]$ neexistuje: označíme-li $w = a+b\sqrt{5}$, jednoduchý rozbor parity ukazuje, že $\nu(w) = |a^2 - 5b^2|$ je buď liché číslo, nebo číslo dělitelné čtyřmi.

7.4. **Ireducibilní prvky.**

Definice. Prvek a se nazývá *ireducibilní*, pokud $a \neq 0$, $a \nparallel 1$ a a nemá vlastní dělitele. Jinými slovy, pokud pro každý rozklad $a = bc$ platí $b \parallel 1$ nebo $c \parallel 1$.

Příklady.

- V tělesech žádné ireducibilní prvky nejsou.
- V oboru $\mathbb{Z}$ jsou ireducibilní právě čísla $\pm p$, kde p je prvočíslo.
- V oboru $\mathbb{C}[x]$ jsou ireducibilní právě polynomy stupně 1.

- V oboru $\mathbb{R}[x]$ jsou ireducibilní právě polynomy stupně 1 a ty polynomy stupně 2, které nemají reálný kořen.
- V oboru $\mathbb{Q}[x]$ jsou ireducibilní i některé polynomy vyšších stupňů, např. všechny polynomy $x^n - 2$, $n \geq 2$, jak plyne z Eisensteinova kritéria 8.9.
- V oboru $\mathbb{Z}[x]$ jsou naopak ireducibilní i některé polynomy stupně 0, konkrétně ty, které jsou ireducibilní jako prvky $\mathbb{Z}$.

Příklad. V tabulce jsou uvedeny rozklady polynomů na součin ireducibilních v různých oborech:

	x^2+1	$2x^2+2$	x^2-2	x^4+2x^2+1
$\mathbb{Z}[x]$	ireducibilní	$2\cdot(x^2+1)$	ireducibilní	$(x^2+1)^2$
$\mathbb{Q}[x]$	ireducibilní	ireducibilní	ireducibilní	$(x^2+1)^2$
$\mathbb{R}[x]$	ireducibilní	ireducibilní	$(x-\sqrt{2})(x+\sqrt{2})$	$(x^2+1)^2$
$\mathbb{C}[x]$	$(x-i)(x+i)$	$(2x-2i)(x+i)$	$(x-\sqrt{2})(x+\sqrt{2})$	$(x-i)^2(x+i)^2$
$(\mathbb{Z}[i])[x]$	$(x-i)(x+i)$	*	ireducibilní	$(x-i)^2(x+i)^2$

* Chybějícím polynomem je $(1-i)(1+i)(x-i)(x+i)$ – pozor na rozklad dvojky, která není v $\mathbb{Z}[i]$ ireducibilní!

K určení ireducibilních prvků v oborech $\mathbb{Z}[\sqrt{s}]$ lze použít Tvrzení 6.1. Podmínka (1) říká, že pokud $u \mid v$, pak $\nu(u) \mid \nu(v)$. Navíc, pokud je u vlastní dělitel, pak $1 \neq \nu(u) \neq \nu(v)$. Speciálně, je-li $\nu(u)$ prvočíslo, pak je u zaručeně ireducibilní. Opačná implikace neplatí, např. v $\mathbb{Z}[i]$ je prvek 3 ireducibilní, ačkoliv má normu 9: pokud by existoval vlastní dělitel $u \mid 3$, pak $\nu(u) = 3$. Označíme-li $u = a + bi$, hledáme a, b splňující $a^2 + b^2 = 3$, ale taková nejsou.

Příklad. V oboru $\mathbb{Z}[i]$ jsou ireducibilní následující prvky:

- $a + 0i$ a $0 + ai$ právě tehdy, když je $|a|$ prvočíslo a $|a| \equiv 3 \pmod 4$;
- $a + bi$, $b \neq 0$, právě tehdy, když $a^2 + b^2$ je prvočíslo.

Důkaz toho, že uvedené prvky jsou ireducibilní, lze snadno provést pomocí Tvrzení 6.1. Důkaz, že ostatní prvky ireducibilní nejsou, je těžší.

8. Gaussovské obory

Cíl. *Budeme zkoumat obory, ve kterých platí analogie základní věty aritmetiky. Ukážeme, jak tato vlastnost souvisí s existencí největších společných dělitelů.*

8.1. Rozklady na ireducibilní činitele.

Definice. Obor integrity se nazývá *gaussovský*, pokud má každý neinvertibilní nenulový prvek jednoznačný rozklad na ireducibilní činitele.

Rozkladem prvku a na ireducibilní činitele rozumíme zápis $a \parallel p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_n^{k_n}$, kde $p_1, \ldots, p_n$ jsou ireducibilní prvky, $p_i \nparallel p_j$ pro $i \neq j$, a $k_1, \ldots, k_n$ jsou přirozená čísla. *Jednoznačností rozkladu* prvku a pak rozumíme jednoznačnost *až na pořadí a asociovanost*, neboli následující vlastnost: jsou-li $a \parallel p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_n^{k_n} \parallel q_1^{l_1} \cdot q_2^{l_2} \cdot \ldots \cdot q_m^{l_m}$ dva ireducibilní rozklady prvku a, pak $m = n$ a existuje permutace indexů π taková, že $p_i \parallel q_{\pi(i)}$ a $k_i = l_{\pi(i)}$ pro každé i.

Definice jednoznačnosti je motivována následujícím pozorováním: v oboru $\mathbb{Z}$ můžeme psát $6 = 2 \cdot 3 = 3 \cdot 2 = (-2) \cdot (-3)$. Formálně vzato, jde o tři různé

rozklady. Přesto je rozumné je považovat za „stejné“: liší se pouze pořadím a volbou z navzájem asociovaných prvků.

Svůj význam má v definici rozkladu také znaménko asociovanosti: prvek -4 v $\mathbb{Z}$ nelze vyjádřit jako druhá mocnina ireducibilního prvku, nicméně přesto má rozklad, $-4 \parallel 2^2$.

Příklad. Řada oborů integrity je gaussovských:

- Tělesa jsou gaussovské obory; podmínka z definice je prázdná.
- Obor $\mathbb{Z}$ je gaussovský, jak říká základní věta aritmetiky 3.1.
- Gaussova věta 8.13 říká: *Je-li* $\mathbf{R}$ *gaussovský obor, pak je* $\mathbf{R}[x_1, \ldots, x_n]$ *také gaussovský obor.*
- Některé obory $\mathbb{Z}[\sqrt{s}]$ jsou gaussovské, např. pro $s = -1, \pm 2, 3$, některé ne, např. pro $s = -3, 5$. Důkaz pro obor $\mathbb{Z}[i]$ uvidíme v Sekci 9.

Příklad. Obor $\mathbb{Z}[\sqrt{5}]$ *není* gaussovský: prvek 4 má dva různé rozklady na ireducibilní činitele

$$4 = 2^2 = (\sqrt{5} - 1)(\sqrt{5} + 1).$$

Je zřejmé, že $2 \nmid \sqrt{5} \pm 1$, protože všechny prvky dělitelné 2 mají sudé koeficienty. Tyto prvky jsou ireducibilní, neboť mají normu 4. Vlastní dělitel by musel mít normu 2, ale jak jsme si ukázali v Sekci 7.3, takové prvky v oboru $\mathbb{Z}[\sqrt{5}]$ nejsou.

To, že protipříkladem na existenci NSD i jednoznačnost rozkladů byl v obou případech obor $\mathbb{Z}[\sqrt{5}]$, není náhoda. V tomto oboru neplatí např. ani analogie Lemmatu 3.4: prvek 2 ireducibilní, $2 \mid (\sqrt{5} - 1)(\sqrt{5} + 1)$, ale $2 \nmid (\sqrt{5} \pm 1)$. Vzájemná souvislost těchto vlastností je hlavním tématem této sekce.

Pro dělitelnost v gaussovských oborech je stěžejní následující pozorování o tom, jak vypadají dělitelé daného prvku.

Tvrzení 8.1. *Buď* $\mathbf{R}$ *gaussovský obor,* $a, b \in R$ *a uvažujme rozklad* a *na ireducibilní činitele*

$$a \parallel p_1^{k_1} \cdot \ldots \cdot p_n^{k_n}.$$

Pak $b \mid a$ *právě tehdy, když*

$$b \parallel p_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$$

pro nějaká $0 \leq l_i \leq k_i$.

Důkaz. Zpětná implikace je snadná: pokud $a = q p_1^{k_1} \cdot \ldots \cdot p_n^{k_n}$ a $b = r p_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$ pro nějaké invertibilní prvky $q, r \in R$, definujme $c = q r^{-1} p_1^{k_1 - l_1} \cdot \ldots \cdot p_n^{k_n - l_n}$ a vidíme, že $a = bc$, tedy $b \mid a$.

Pro opačnou implikaci uvažujme prvek c takový, že $a = b \cdot c$ a označme

$$b \parallel b_1^{r_1} \cdot \ldots \cdot b_u^{r_u} \quad \text{a} \quad c \parallel c_1^{s_1} \cdot \ldots \cdot c_v^{s_v}$$

ireducibilní rozklady prvků b, c. Pak

$$a \parallel p_1^{k_1} \cdot \ldots \cdot p_n^{k_n} = b_1^{l_1} \cdot \ldots \cdot b_u^{r_u} \cdot c_1^{s_1} \cdot \ldots \cdot c_v^{s_v}$$

jsou dva rozklady prvku a, a tedy z jednoznačnosti plyne, že ke každému $i = 1, \ldots, r$ existuje j takové, že $b_i \parallel p_j$, přičemž pro každé $j = 1, \ldots, n$ existuje nejvýše k_j indexů i takových, že $b_i \parallel p_j$. Z toho vyplývá, že $b \parallel p_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$ pro nějaká $0 \leq l_i \leq k_i$. □

Snadným důsledkem je, že v gaussovských oborech platí analogie Lemmatu 3.4, které tvořilo klíčový krok důkazu základní věty aritmetiky.

Tvrzení 8.2. *Buď* $\mathbf{R}$ *gaussovský obor a* $p \in R$ *ireducibilní prvek. Platí-li* $p \mid a \cdot b$, *pak* $p \mid a$ *nebo* $p \mid b$.

Ideu důkazu předvedeme na příkladě: pokud $p \mid 14 \cdot 12 = 2 \cdot 7 \cdot 2^2 \cdot 3$, pak p je buď 2 (pak $p \mid 14$ i $p \mid 12$), nebo 3 (pak $p \mid 12$), nebo 7 (pak $p \mid 14$).

Důkaz. Označme $a = a_1^{k_1} \cdot \ldots \cdot a_m^{k_m}$ a $b = b_1^{l_1} \cdot \ldots \cdot b_n^{l_n}$ ireducibilní rozklady prvků a, b. Protože

$$p \mid a_1^{k_1} \cdot \ldots \cdot a_m^{k_m} \cdot b_1^{l_1} \cdot \ldots \cdot b_n^{l_n},$$

podle Tvrzení 8.1 musí p mít rozklad, který obsahuje některé z prvků $a_1, \ldots a_m$, $b_1, \ldots, b_n$. Protože je p ireducibilní, musí být $p \parallel a_i$ nebo $p \parallel b_i$ pro nějaké i. V prvním případě $p \mid a$, v druhém $p \mid b$. □

Poznámka. Prvek p splňující implikaci

$$p \mid a \cdot b \ \Rightarrow p \mid a \text{ nebo } p \mid b$$

se nazývá *prvočinitel*. Právě jsme dokázali, že v gaussovských oborech jsou ireducibilní prvky prvočinitelé. Obecně to však neplatí, viz výše uvedený protipříklad v $\mathbb{Z}[\sqrt{5}]$.

Na druhou stranu, prvočinitelé jsou vždy ireducibilní: kdybychom měli rozklad $p = ab$, pak $p \mid ab$, tedy $p \mid a$ nebo $p \mid b$, z čehož plyne $p \parallel a$ nebo $p \parallel b$, čili jde o triviální rozklad. (Tedy v gaussovských oborech oba pojmy splývají.)

Jiným snadným důsledkem Tvrzení 8.1 je existence největších společných dělitelů.

Tvrzení 8.3. *V gaussovských oborech existuje NSD všech dvojic prvků.*

Ideu důkazu předvedeme na příkladě: $\mathrm{NSD}(540, 336) = \mathrm{NSD}(2^2 \cdot 3^3 \cdot 5, 2^4 \cdot 3 \cdot 7) =$
$= \mathrm{NSD}(2^2 \cdot 3^3 \cdot 5^1 \cdot 7^0, 2^4 \cdot 3^1 \cdot 5^0 \cdot 7^1) = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 = 12.$

Důkaz. Buď

$$a \parallel p_1^{k_1} \cdot \ldots \cdot p_n^{k_n} \quad \text{a} \quad b \parallel p_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$$

rozklady prvků a, b, přičemž předpokládáme p_i ireducibilní, $p_i \nparallel p_j$ pro $i \neq j$ a $k_i, l_i \geq 0$. (Uvědomte si, že rozklady můžeme zvolit v této speciální formě, tj. se stejnými ireducibilními prvky: do rozkladu případně doplníme činitele v nulté mocnině.) Položme

$$c = p_1^{\min(k_1, l_1)} \cdot \ldots \cdot p_n^{\min(k_n, l_n)}$$

a ukažme, že $\mathrm{NSD}(a, b) = c$. Z Tvrzení 8.1 plyne, že d je společný dělitel a, b právě tehdy, když $d \parallel p_1^{r_1} \cdot \ldots \cdot p_n^{r_n}$ pro nějaká $r_1, \ldots, r_n \geq 0$ splňující zároveň $r_i \leq k_i$ a $r_i \leq l_i$ pro všechna i. Je zřejmé, že největší (vzhledem k dělitelnosti) je takové d, kde $r_i = \min(k_i, l_i)$. □

Poslední vlastnost gaussovských oborů, kterou vypícheneme, říká, že žádný prvek „nelze dělit do nekonečna“, tj., že neexistuje nekonečná posloupnost vlastních dělitelů. Myšlenka důkazu je založena na intuici, že každý prvek má svoji „velikost“, danou tím, kolik ireducibilních prvků se vyskytuje v jeho rozkladu (včetně násobnosti).

Tvrzení 8.4. *Buď* $\mathbf{R}$ *gaussovský obor. Pak neexistuje posloupnost* $a_1, a_2, a_3, \ldots \in R$ *taková, že* $a_{i+1} \mid a_i$ *a* $a_{i+1} \nparallel a_i$.

Důkaz. Buď a nenulový neinvertibilní prvek oboru $\mathbf{R}$. Tento prvek má, až na pořadí a volbu ireducibilních prvků v základu mocnin, jednoznačný rozklad $a \parallel p_1^{k_1} \cdot \ldots \cdot p_n^{k_n}$. Označme $\nu(a) = k_1 + \ldots + k_n$ a dodefinujme $\nu(a) = 0$ pro všechny invertibilní prvky $a \in R$. Z jednoznačnosti rozkladů plyne, že číslo $\nu(a)$ je nezávislé na volbě rozkladu. Z Tvrzení 8.1 plyne, že pokud $u \mid v$ a $v \nmid u$, pak $\nu(u) < \nu(v)$.

Pro spor předpokládejme existenci takové posloupnosti $a_1, a_2, a_3 \ldots$ Z úvah v předešlém odstavci plyne, že $\nu(a_1) > \nu(a_2) > \nu(a_3) > \ldots$ je nekonečná klesající posloupnost nezáporných celých čísel, spor. □

Na závěr uveďme, že gaussovské obory nesdílejí všechny hezké vlastnosti oboru celých čísel: například obecně nelze dělit se zbytkem a neplatí analogie Bézoutovy rovnosti. Např. v oboru $\mathbb{Z}[x]$ je $\mathrm{NSD}(x+1, x-1) = 1$, ale neexistují $u, v \in \mathbb{Z}[x]$ taková, že $u \cdot (x+1) + v \cdot (x-1) = 1$.

8.2. **Analogie základní věty aritmetiky.**

Dostáváme se k slibované souvislosti ireducibilních rozkladů a existence NSD. Už víme, že v gaussovských oborech NSD existují. K důkazu základní věty aritmetiky jsme ale potřebovali dvě stěžejní ingredience: kromě existence NSD také matematickou indukci v poměrně silné formě. Obecné obory integrity nemusí být uspořádatelné, klasická indukce nám tedy nepomůže. Ale pomůže z Tvrzení 8.4 vycházející z faktu, že ka6d7 prvek má svoji „velikost“ danou ireducibilním rozkladem.

Věta 8.5. *Buď* $\mathbf{R}$ *obor integrity. Pak* $\mathbf{R}$ *je gaussovský právě tehdy, když*

(1) *existuje NSD všech dvojic prvků;*
(2) *neexistuje posloupnost* $a_1, a_2, a_3, \ldots \in R$ *taková, že* $a_{i+1} \mid a_i$ *a* $a_{i+1} \nparallel a_i$.

Přímou implikaci jsme dokázali v Tvrzeních 8.3 a 8.4. K důkazu opačné implikace se nám bude hodit ještě jedna analogie Lemmatu 3.4, tentokrát dokázaná za předpokladu existence NSD. Protože obecně nemáme k dispozici Bézoutovu rovnost, budeme muset postupovat obezřetněji než v důkaze zmíněného lemmatu v Sekci 3.

Lemma 8.6. *Buď* $\mathbf{R}$ *obor integrity a* $a, b, c \in R$ *takové, že existuje* $\mathrm{NSD}(a, b)$ *i* $\mathrm{NSD}(ac, bc)$. *Pak*

$$\mathrm{NSD}(ac, bc) = c \cdot \mathrm{NSD}(a, b).$$

Důkaz. Vzhledem k tomu, že NSD je definován *až na asociovanost*, stačí dokázat, že levá strana rovnosti dělí pravou a naopak. Označme $u = \mathrm{NSD}(ac, bc)$.

Nejprve dokážeme, že $u \mid c \cdot \mathrm{NSD}(a, b)$. Protože $u \mid ac$, existuje x s vlastností $ac = ux$. Protože $u \mid bc$, existuje y s vlastností $bc = uy$. Protože c je společný dělitel ac, bc, platí $c \mid u$, a tedy existuje z s vlastností $u = cz$. Dostáváme $ac = czx$ a $bc = czy$ a krácením získáme vztahy $a = zx$ a $b = zy$. Tedy z je společný dělitel a, b, tedy z dělí $\mathrm{NSD}(a, b)$, a tudíž $u = cz \mid c \cdot \mathrm{NSD}(a, b)$.

Naopak, protože $\mathrm{NSD}(a, b)$ dělí a i b, tak $c \cdot \mathrm{NSD}(a, b)$ dělí ac i bc, a tudíž musí dělit i jejich největšího společného dělitele. □

Lemma 8.7. *Předpokládejme, že v oboru* $\mathbf{R}$ *existují NSD všech dvojic prvků a buď* $p \in R$ *ireducibilní prvek. Platí-li* $p \mid a \cdot b$, *pak* $p \mid a$ *nebo* $p \mid b$.

Důkaz. Předpokládejme, že $p \nmid a$. Pak $\mathrm{NSD}(a, p) = 1$, protože je p ireducibilní, a tedy podle Lemmatu 8.6

$$\mathrm{NSD}(pb, ab) = b \cdot \mathrm{NSD}(p, a) = b.$$

Ovšem p je společným dělitelem pb a ab, tedy $p \mid \mathrm{NSD}(pb, ab) = b$. □

Nyní se můžeme pustit do důkazu existence a jednoznačnosti rozkladů. Srovnejte tento obecný důkaz s důkazem základní věty aritmetiky pro konkrétní obor $\mathbb{Z}$.

Důkaz Věty 8.5. ($\Rightarrow$) Viz Tvrzení 8.3 a 8.4.

($\Leftarrow$) Nejprve ukážeme, že každý prvek má ireducibilní rozklad, a poté, že jsou tyto rozklady jednoznačné.

Pro spor předpokládejme, že nějaký prvek a nemá ireducibilní rozklad, $0 \neq a \nmid 1$. Rekurzí zkonstruujeme posloupnost, která protiřečí bodu (2).

(i) Položme $a_1 = a$. Tedy $a_1 \nmid 1$ a nemá ireducibilní rozklad.

(ii) Předpokládejme, že $a_i \nmid 1$ a nemá ireducibilní rozklad. Speciálně, prvek a_i není sám ireducibilní, a tedy $a_i = b \cdot c$ pro nějaká $b, c \nmid 1$. Kdyby b i c měly ireducibilní rozklad, pak by ho měl i a_i, takže aspoň jedno z nich ireducibilní rozklad nemá, označme jej a_{i+1}. Tedy a_{i+1} je vlastní dělitel a_i a nemá ireducibilní rozklad.

Tato posloupnost $a_1, a_2, \ldots$ protiřečí předpokladu (2).

Jednoznačnost rozkladu také dokážeme sporem. Mezi všemi prvky s dvěma různými rozklady na ireducibilní činitele zvolme takové a, jehož rozklad je nejkratší, ve smyslu součtu exponentů u všech ireducibilních prvků v tomto rozkladu. Označme tento nejkratší rozklad $a \parallel p_1^{k_1} \cdot \ldots \cdot p_n^{k_n}$ a uvažujme nějaký jiný rozklad $a \parallel q_1^{l_1} \cdot \ldots \cdot q_m^{l_m}$. Protože je p_1 ireducibilní, podle Lemmatu 8.7 musí dělit některé q_i. Protože jsou všechna q_j ireducibilní a $p_1 \nmid 1$, máme $p_1 \parallel q_i$. Pak ale $b = p_1^{k_1-1} p_2^{k_2} \cdot \ldots \cdot p_n^{k_n} \parallel q_1^{l_1} \cdot \ldots \cdot q_{i-1}^{l_{i-1}} \cdot q_i^{l_i-1} q_{i+1}^{l_{i+1}} \cdot \ldots \cdot q_m^{l_m}$ je prvek s kratším nejednoznačným rozkladem, což je spor. □

Věta 8.5 je zajímavá mimo jiné proto, že charakterizuje gaussovské obory dvěma zcela rozdílnými způsoby. Definice pomocí existence a jednoznačnosti rozkladů je čistě aritmetická, formulovaná jako vlastnost operace násobení v oboru $\mathbf{R}$. Naopak druhou stranu charakterizace lze formulovat čistě v jazyku uspořádaných množin: říká, že uspořádaná množina $(\bar{R}, |)$ je svazově uspořádaná (tj. existují NSD a NSN) a zároveň v ní neexistuje nekonečný ostře klesající řetězec.

8.3. **Racionální kořeny polynomů a Eisensteinovo kritérium.**

Následující princip je užitečný pro hledání racionálních kořenů daného celočíselného polynomu. Princip ukážeme v plné obecnosti, pro libovolné gaussovské obory.

Tvrzení 8.8. *Buď $\mathbf{R}$ gaussovský obor a $\mathbf{Q}$ jeho podílové těleso. Má-li polynom $f = \sum_{i=0}^{n} a_i x^i \in R[x]$ kořen $\frac{r}{s} \in Q$ (předpokládáme r, s nesoudělná), pak $r \mid a_0$ a $s \mid a_n$.*

Důkaz. Dosaďme prvek $\frac{r}{s}$ do f. Protože $\sum_{i=0}^{n} a_i (\frac{r}{s})^i = 0$, přenásobením prvkem s^n dostáváme

$$a_0 s^n + a_1 r s^{n-1} + a_2 r^2 s^{n-2} + \ldots + a_{n-1} r^{n-1} s + a_n r^n = 0.$$

Protože r dělí všechny členy $a_1 r s^{n-1}, \ldots, a_n r^n$, musí dělit i první člen $a_0 s^n$. Protože jsou r, s nesoudělné musí $r \mid a_0$ — zde využíváme Tvrzení 8.2 aplikované na všechny ireducibilní prvku v rozkladu r. Analogicky, protože s dělí všechny členy $a_0 s^n, \ldots, a_{n-1} r^{n-1} s$, musí dělit i poslední člen $a_n r^n$, tedy $s \mid a_n$. □

Příklad. Najdeme všechny racionální kořeny polynomu $2x^5 - 3x^4 + 2x - 3$. Podle Tvrzení 8.8 jsou jedinými kandidáty čísla ± 1, ± 3, $\pm\frac{1}{2}$ a $\pm\frac{3}{2}$. Dosazením zjistíme, že vyhovuje pouze číslo $-\frac{3}{2}$.

Pro negaussovské obory tento argument neprojde. Uvažujme například obor $\mathbb{Z}[\sqrt{5}]$, $n \geq 2$, $r = 1 + \sqrt{5}$, $s = 2$, $a_0 = 1$: pak $\mathrm{NSD}(1 + \sqrt{5}, 2) = 1$, $r \mid 1 \cdot 2^n = (1 + \sqrt{5})(-1 + \sqrt{5})2^{n-2}$, ale $r \nmid 1$.

Podobným způsobem lze odvodit jednoduché, ale užitečné kritérium ireducibility polynomu. Polynom $f \in R[x]$ nazýváme *primitivní*, pokud není dělitelný žádným konstantním polynomem. Je-li $\mathbf{R}$ gaussovský obor a $f = \sum a_i x^i \in R[x]$, pak f je primitivní právě tehdy, když $\mathrm{NSD}(a_0, \ldots, a_n) = 1$.

Tvrzení 8.9 (Eisensteinovo kritérium)**.** *Buď $\mathbf{R}$ gaussovský obor a $f = \sum_{i=0}^{n} a_i x^i$ primitivní polynom z $\mathbf{R}[x]$. Pokud existuje ireducibilní prvek $p \in R$ splňující $p \mid a_0$, $p \mid a_1$, …, $p \mid a_{n-1}$ a $p^2 \nmid a_0$, pak je polynom f ireducibilní v $\mathbf{R}[x]$.*

Důkaz. Uvažujme rozklad $f = gh$, kde $g = \sum_{i=0}^{k} b_i x^i$ a $h = \sum_{i=0}^{l} c_i x^i$ jsou polynomy z $\mathbf{R}[x]$ stupně alespoň 1. Protože $p \mid a_0 = b_0 c_0$, podle Tvrzení 8.2 platí $p \mid b_0$ nebo $p \mid c_0$, ale určitě ne oboje zároveň, protože $p^2 \nmid a_0$. Nechť je to bez újmy na obecnosti b_0. Protože $p \mid a_1 = b_0 c_1 + b_1 c_0$ a $p \nmid c_0$, podle Tvrzení 8.2 musí $p \mid b_1$. Protože $p \mid a_2 = b_0 c_2 + b_1 c_1 + b_2 c_0$ a $p \nmid c_0$, musí $p \mid b_2$. Tímto způsobem zjistíme, že p dělí všechny koeficienty b_i, tedy $p \mid gh = f$, což je spor s primitivitou. □

Příkladem použití Eisensteinova kritéria je ireducibilita polynomů $x^n \pm a$ v $\mathbb{Z}[x]$, kde a není dělitelné čtvercem prvočísla.

[DOPLNIT: sofistikovanější použití pomocí substituce]

8.4. **Polynomy nad gaussovskými obory.**

Polynom nazýváme *primitivní*, pokud není dělitelný žádným neinvertibilním konstantním polynomem. Ekvivalentně (pro gaussovské obory), je-li NSD jeho koeficientů 1. Uvažujme nějaký obor integrity $\mathbf{R}$ a jeho podílové těleso $\mathbf{Q}$. Dělitelnost v oborech $\mathbf{R}[x]$ a $\mathbf{Q}[x]$ se pro primitivní polynomy chová velmi podobně.

Tvrzení 8.10. *Buď $\mathbf{R}$ gaussovský obor, $\mathbf{Q}$ jeho podílové těleso a f, g primitivní polynomy z $\mathbf{R}[x]$. Pak*

(1) *$f \mid g$ v $\mathbf{R}[x]$ právě tehdy, když $f \mid g$ v $\mathbf{Q}[x]$;*
(2) *f je ireducibilní v $\mathbf{R}[x]$ právě tehdy, když f je ireducibilní v $\mathbf{Q}[x]$;*
(3) *$\mathrm{NSD}_{\mathbf{R}[x]}(f, g)$ existuje a je roven primitivnímu polynomu $h \in R[x]$ splňujícímu $h = \mathrm{NSD}_{\mathbf{Q}[x]}(f, g)$.*

Takový polynom h v části (3) jistě existuje: stačí vzít libovolný $\mathrm{NSD}(f, g)$ v $\mathbf{Q}[x]$ a přenásobit ho prvkem $q = \frac{a}{b} \in Q$, kde a je NSN jmenovatelů všech koeficientů, a b je NSD všech čitatelů koeficientů.

Je zřejmé, že pokud $f \mid g$ a g je primitivní, pak je i f primitivní. Klíčovým krokem k důkazu uvedeného tvrzení je fakt, že platí také opačné tvrzení: součin fg je primitivní právě tehdy, když jsou oba polynomy f, g primitivní.

Lemma 8.11 (Gaussovo lemma)**.** *Buď $\mathbf{R}$ gaussovský obor a f, g primitivní polynomy z $\mathbf{R}[x]$. Pak fg je primitivní polynom.*

Důkaz. Označme $f = \sum_{i=0}^{n} a_i x^i$ a $g = \sum_{i=0}^{m} b_i x^i$ a předpokládejme, že fg není primitivní polynom. Tedy existuje ireducibilní prvek $u \in R$, který dělí součin fg, tj. všechny koeficienty tohoto součinu. Zvolme nejmenší j takové, že $u \nmid a_j$, a nejmenší

k takové, že $u \nmid b_k$ (protože jsou polynomy f, g primitivní, u nemůže dělit všechny jejich koeficienty). Podívejme se na $(j+k)$-tý koeficient polynomu fg:

$$c_{j+k} = a_0 b_{j+k} + \ldots + a_{j-1} b_{k+1} + a_j b_k + a_{j+1} b_{k-1} + \ldots + a_{j+k} b_0.$$

Protože $u \mid a_i$ pro všechna $i < j$, máme

$$u \mid a_0 b_{j+k} + \ldots + a_{j-1} b_{k+1}.$$

Protože $u \mid b_i$ pro všechna $i < k$, máme

$$u \mid a_{j+1} b_{k-1} + \ldots + a_{j+k} b_0.$$

Tedy u dělí všechny členy kromě $a_j b_k$. Ten naopak u dělitelný není, protože u je ireducibilní a nedělí ani a_j, ani b_k. Dostáváme, že $u \nmid c_{j+k}$, spor. □

Důkaz Tvrzení 8.10. (1) Pokud že $f \mid g$ v $\mathbf{R}[x]$, tj. že existuje $h \in R[x] \subseteq Q[x]$ splňující $g = fh$, pak tato rovnost platí i v $\mathbb{Q}[x]$. Opačná implikace je těžší. Předpokládejme, že $f \mid g$ v $\mathbf{Q}[x]$, tj. že existuje $h \in Q[x]$ splňující $g = fh$. Zvolme $q \in Q$ tak, aby qh byl primitivní polynom z $\mathbf{R}[x]$. Pak $qg = f \cdot qh$, na pravé straně je součin primitivních polynomů z $\mathbf{R}[x]$, takže podle Gaussova lemmatu je qg také primitivní polynom z $\mathbf{R}[x]$. Označme $q = \frac{a}{b} \in Q$. Platí $ag = b(qg)$, přitom oba polynomy g, qg jsou primitivní, takže z $a \mid b(qg)$ plyne $a \mid b$, a z $b \mid aq$ plyne $b \mid a$ (využíváme Tvrzení 8.2). Tedy $a \parallel b$ a $1 \parallel q \in R$ a dostáváme $h \in R[x]$.

(2) Dokážeme následující ekvivalentní tvrzení: f má vlastního dělitele v $\mathbf{R}[x]$ právě tehdy, když má vlastního dělitele v $\mathbf{Q}[x]$. ($\Rightarrow$) Protože je f primitivní, jakýkoliv vlastní dělitel je primitivní a má stupeň aspoň 1. Tedy jde zároveň o vlastního dělitele v $\mathbf{Q}[x]$. ($\Leftarrow$) Nechť g je vlastní dělitel f v $\mathbf{Q}[x]$. Pak existuje $q \in Q$ takové, že qg je primitivní polynom z $\mathbf{R}[x]$. Přitom $qg \mid f$ v $\mathbf{Q}[x]$, tedy podle (1) je qg vlastní dělitel f v $\mathbf{R}[x]$.

(3) Polynom h dělí f, g v $\mathbf{Q}[x]$ a je primitivní, tedy podle (1) dělí f, g i v $\mathbf{R}[x]$, takže je to společný dělitel. Kdykoliv máme jiný společný dělitel $d \mid f, g$ v $\mathbf{R}[x]$, pak je jistě primitivní, podle (1) $d \mid f, g$ v $\mathbf{Q}[x]$, tedy $d \mid h$ v $\mathbf{Q}[x]$, a opět podle (1) $d \mid h$ i v $\mathbf{R}[x]$. □

Z Tvrzení 8.10 lze snadno odvodit podobná tvrzení pro obecné polynomy, ne nutně primitivní. Buď $f = \sum_{i=0}^{n} a_i x^i$ polynom z $\mathbf{R}[x]$. Definujeme

$$\mathrm{c}(f) = \mathrm{NSD}(a_0, \ldots, a_n) \qquad \text{a} \qquad \mathrm{pp}(f) = f / \mathrm{c}(f).$$

Polynom $\mathrm{pp}(f)$ je očividně primitivní a nazývá se *primitivní částí* polynomu f.

Věta 8.12. *Buď $\mathbf{R}$ gaussovský obor, $\mathbf{Q}$ jeho podílové těleso a f, g polynomy z $\mathbf{R}[x]$. Pak*

(1) *f je ireducibilní v $\mathbf{R}[x]$ právě tehdy, když*
   - *$\deg f = 0$ a f je ireducibilní v $\mathbf{R}$; nebo*
   - *$\deg f > 0$, f je primitivní a ireducibilní v $\mathbf{Q}[x]$.*

(2) *Pak $\mathrm{NSD}_{\mathbf{R}[x]}(f, g)$ existuje a je roven součinu $c \cdot h$, kde $c = \mathrm{NSD}_{\mathbf{R}}(\mathrm{c}(f), \mathrm{c}(g))$ a h je primitivní polynom z $\mathbf{R}[x]$ splňující $h = \mathrm{NSD}_{\mathbf{Q}[x]}(\mathrm{pp}(f), \mathrm{pp}(g))$.*

Důkaz. (1) Pokud není f primitivní, pak se rozkládá na součin neinvertibilního konstantního polynomu a primitivního polynomu. Jinak je buď konstantní (první položka), nebo primitivní (druhá položka).

(2) Označme pravou stranu r. Protože $\mathrm{NSD}_{\mathbf{R}}(\mathrm{c}(f), \mathrm{c}(g))$ dělí $\mathrm{c}(f)$ i $\mathrm{c}(g)$, a zároveň $\mathrm{NSD}_{\mathbf{R}[x]}(\mathrm{pp}(f), \mathrm{pp}(g))$ dělí $\mathrm{pp}(f)$ i $\mathrm{pp}(g)$, tak jejich součin r dělí oba polynomy

f, g, čili r je společný dělitel. Dokážeme, že je to největší společný dělitel: pokud nějaký h dělí f i g, pak $\mathrm{c}(h)$ dělí $\mathrm{c}(f)$ i $\mathrm{c}(g)$, tedy $\mathrm{c}(h) \mid \mathrm{NSD}_{\mathbf{R}}(\mathrm{c}(f), \mathrm{c}(g))$; analogicky $\mathrm{pp}(h) \mid \mathrm{NSD}_{\mathbf{R}[x]}(\mathrm{pp}(f), \mathrm{pp}(g))$ a dostáváme $h \mid r$. □

Příklady.

- Polynom $2x - 2$ je ireducibilní v $\mathbb{Q}[x]$, ale není ireducibilní v $\mathbb{Z}[x]$, protože není primitivní: rozkládá se jako $2 \cdot (x - 1)$.
- Polynom 2 není ireducibilní v $\mathbb{Q}[x]$, protože je invertibilní, ale je ireducibilní v $\mathbb{Z}[x]$.

Příklad. Uvažujme obor $\mathbb{Z}[x]$ a polynomy

$$f = 4x^2 + 8x + 4, \qquad g = -6x^2 + 6.$$

Pak $c = \mathrm{NSD}_{\mathbb{Z}}(4, -6) = 2$, $h = \mathrm{NSD}_{\mathbb{Q}[x]}(x^2 + 2x + 1, x^2 - 1) = x + 1$, a tedy $\mathrm{NSD}_{\mathbb{Z}[x]}(4x^2 + 8x + 4, -6x^2 + 6) = 2(x + 1)$.

Věta 8.12 nejen zaručuje existenci NSD v $\mathbf{R}[x]$, ale také dává návod, jak je spočítat. Např. vypočet NSD v $\mathbb{Z}[x]$ se redukuje na dva výpočty NSD, jeden v $\mathbb{Z}$ a druhý v $\mathbb{Q}[x]$. Oba lze provést pomocí Eukleidova algoritmu.

Věta 8.13 (Gaussova)**.** *Je-li* $\mathbf{R}$ *gaussovský obor, pak je* $\mathbf{R}[x]$ *také gaussovský obor.*

Důkaz. Použijeme charakterizaci z Věty 8.5. NSD v $\mathbf{R}[x]$ existují podle Věty 8.12. A je-li $f_1, f_2, f_3, \ldots$ posloupnost vlastních dělitelů, pak $\deg f_1 \geq \deg f_2 \geq \deg f_3 \geq \cdots \geq 0$, a tedy existuje n takové, že $\deg f_n = \deg f_{n+1} = \ldots$ Označíme-li a_i vedoucí koeficient polynomu f_i, pak $a_n, a_{n+1}, \ldots$ je posloupnost vlastních dělitelů v $\mathbf{R}$, spor. □

Z Gaussovy věty ihned plyne, že také obory více proměnných nad gaussovským oborem jsou gaussovské: použije se indukce podle počtu proměnných a vztah $\mathbf{R}[x_1, \ldots, x_n] = (\mathbf{R}[x_1, \ldots, x_{n-1}])[x_n]$.

9. Eukleidovské obory

Cíl. *Budeme se zabývat obory, ve kterých, zjednodušeně řečeno, lze dělit se zbytkem. Dělitelnost se pak chová hezky: NSD je možné počítat pomocí Eukleidova algoritmu, platí zde Bézoutova rovnost, a tudíž jde o gaussovské obory.*

9.1. Eukleidův algoritmus.

Definice. *Eukleidovskou normou* na oboru $\mathbf{R}$ rozumíme zobrazení

$$\nu : R \to \mathbb{N} \cup \{0\}$$

splňující

(0) $\nu(0) = 0$;
(1) pokud $a \mid b \neq 0$, pak $\nu(a) \leq \nu(b)$;
(2) pro všechna $a, b \in R$, $b \neq 0$, existují $q, r \in R$ taková, že

$$a = bq + r \quad \text{a} \quad \nu(r) < \nu(b).$$

Obor $\mathbf{R}$ se nazývá *eukleidovský*, pokud na něm existuje eukleidovská norma.

Eukleidovská norma nám umožňuje „měřit“ prvky daného oboru s ohledem na jejich dělitelnost. Podmínka (2) říká, že pro každou dvojici $a, b \neq 0$ existuje „podíl“ q a „zbytek“ r (bez nároku na jejich jednoznačnost!), přičemž zbytek je „menší“ než prvek, kterým dělíme.

Příklad. Řada gaussovských oborů je také eukleidovských:

- Tělesa jsou eukleidovské obory. Eukleidovskou normou je např. zobrazení $\nu(0) = 0$ a $\nu(a) = 1$ pro všechna $a \neq 0$.
- Obor $\mathbb{Z}$ je eukleidovský. Normou je absolutní hodnota, tj. $\nu(a) = |a|$.
- Obor $\mathbb{Z}[i]$ (Gaussova celá čísla) je eukleidovský s normou $\nu(z) = |z|^2$, viz Tvrzení 6.1 a 6.2.
- Obor $\mathbb{Z}[\omega]$ (Eisensteinova celá čísla), kde $\omega = e^{2\pi i/3}$ je komplexní třetí odmocnina z jedné, je eukleidovský. Normou je $\nu(z) = |z|^2$.
- Některé obory $\mathbb{Z}[\sqrt{s}]$ jsou eukleidovské, např. pro $s = -1, \pm 2, 3$, některé ne, např. pro $s = -3, 5$. V uvedených případech je normou
$$\nu(a + b\sqrt{s}) = |a^2 - sb^2|.$$
- Obor $\mathbf{T}[x]$ je eukleidovský pro libovolné těleso $\mathbf{T}$. Normou je
$$\nu(f) = 1 + \deg f.$$
(Proč ne pouze $\deg f$? Protože 0 musí být jediný prvek s normou 0.) Vlastnost (1) je zřejmá a vlastnost (2) plyne z Tvrzení 5.2.

Ne každý gaussovský obor je eukleidovský. Příkladem je obor $\mathbb{Z}[x]$ nebo obory polynomů více proměnných nad tělesem. Rozebereme případ oboru $\mathbb{Z}[x]$. Všimněte si, že zobrazení $\nu(f) = 1 + \deg f$ není eukleidovskou normou: např. pro polynomy $3x$ a $2x$ neexistují $q, r \in \mathbb{Z}[x]$ splňující $3x = q \cdot 2x + r$ a $\deg r = 0$ — po dosazení nuly vidíme, že $r = 0$, a tedy musí platit $3x = 2qx$, ale takový polynom v $\mathbb{Z}[x]$ neexistuje. Pozor, z uveden0ho neplyne, že obor $\mathbb{Z}[x]$ není eukleidovský! Pouze jsme dokázali, že toto konkrétní ν není eukleidovskou normou. Přímý důkaz, že žádné zobrazení $\mathbb{Z}[x] \to \mathbb{N} \cup \{0\}$ nesplňuje podmínky eukleidovské normy, by byl komplikovaný. Jednodušší je využít faktu, že eukleidovské obory splňují Bézoutovu rovnost, viz níže.

Dělitelnost se v eukleidovských oborech chová hezky: NSD je možné počítat pomocí Eukleidova algoritmu, platí Bézoutova rovnost a pomocí Věty 8.5 dokážeme také existenci a jednoznačnost ireducibilních rozkladů.

Eukleidův algoritmus. Buď $\mathbf{R}$ eukleidovský obor.

- `VSTUP:` $a, b \in R$, $\nu(a) \geq \nu(b)$.
- `VÝSTUP:` $\mathrm{NSD}(a, b)$ a $u, v \in R$ splňující $\mathrm{NSD}(a, b) = u \cdot a + v \cdot b$.
- $a_0 = a, \quad u_0 = 1, \quad v_0 = 0.$
  $a_1 = b, \quad u_1 = 0, \quad v_1 = 1.$
  $a_{i+1} = r, \quad u_{i+1} = u_{i-1} - u_i q, \quad v_{i+1} = v_{i-1} - v_i q$, kde q, r zvolíme tak, aby
$$a_{i-1} = a_i q + r \quad \text{a} \quad \nu(r) < \nu(a_i).$$
  Pokud $a_{i+1} = 0$, odpověz a_i, u_i, v_i.

Věta 9.1. *Eukleidův algoritmus najde v eukleidovském oboru* $\mathbf{R}$ *pro jakýkoliv vstup* $a, b \in R$ *hodnotu* $\mathrm{NSD}(a, b)$ *a nějaká* $u, v \in R$ *splňující*
$$\mathrm{NSD}(a, b) = u \cdot a + v \cdot b.$$

Důkaz. Vzhledem k tomu, že $\nu(a_0) \geq \nu(a_1) > \nu(a_2) > \nu(a_3) > \ldots \geq 0$, algoritmus se musí po konečně mnoha krocích zastavit; označme K číslo kroku, ve kterém se tak stane. Je třeba dokázat, že

$$\mathrm{NSD}(a,b) = a_K = u_K \cdot a + v_K \cdot b.$$

Vzhledem k tomu, že $\mathrm{NSD}(a_K, 0) = a_K$, stačí dokázat, že NSD dvou po sobě jdoucích prvků posloupnosti $a_0, a_1, \ldots, a_K$ se nemění, tj. že

(1) pro každé $i = 1, \ldots, K$ platí $\mathrm{NSD}(a_{i-1}, a_i) = \mathrm{NSD}(a_i, a_{i+1})$;
(2) pro každé $i = 0, \ldots, K$ platí $a_i = u_i \cdot a + v_i \cdot b$.

Obě tvrzení plynou z vyjádření

$$a_{i-1} = a_i q + a_{i+1}.$$

Pro důkaz (1) si stačí uvědomit, že dvojice a_{i-1}, a_i má stejné společné dělitele jako dvojice a_i, a_{i+1} (jde o analogii Lemmatu 3.2). Indukcí ověříme (2). Pro $i = 0, 1$ výrok zřejmě platí. Dále, předpokládáme-li $a_{i-1} = u_{i-1}a + v_{i-1}b$ a $a_i = u_i a + v_i b$, pak

$$\begin{aligned} a_{i+1} &= a_{i-1} - a_i q = (u_{i-1}a + v_{i-1}b) - (u_i a + v_i b) \cdot q \\ &= (u_{i-1} - u_i q) \cdot a + (v_{i-1} - v_i q) \cdot b = u_{i+1}a + v_{i+1}b. \end{aligned}$$

□

Úloha. Spočtěte $\mathrm{NSD}(3 + 11i, -2 + 9i)$ v oboru $\mathbb{Z}[i]$.

Řešení. Jeden způsob řešení je pomocí rozkladů na ireducibilní prvky. Máme $\nu(3 + 11i) = 130 = 2 \cdot 5 \cdot 13$ a $\nu(-2 + 9i) = 85 = 5 \cdot 17$, čili případný psolečný dělitel by měl normu 5. Snadno vyzkoušíme, že $2 - i \mid 3 + 11i$, zatímco $2 + i \mid -2 + 9i$, a protože $2 + i \nmid 2 - i$, uvedená čísla jsou nesoudělná.

Druhý způsob je provést Eukleidův algoritmus. Postupně dostáváme čísla $a_0 = 3 + 11i$, $a_1 = -2 + 9i$, $a_2 = 4$, $a_3 = -2 + i$, $a_4 = 1$, $a_5 = 0$, největší společný dělitel je tedy 1. □

Platnost Bézoutovy rovnosti je možné využít k důkazu, že daný obor není eukleidovský. Místo zkoumání všech kandidátů na eukleidovskou normu stačí najít jeden příklad, kdy neplatí Bézoutova rovnost.

Příklad. Obor $\mathbb{Z}[x]$ není eukleidovský, protože v něm neplatí Bézoutova rovnost. Např. pro polynomy $f = x$ a $g = 2$ platí $\mathrm{NSD}(x + 1, 2) = 1$, ale přitom neexistuje $u, v \in \mathbb{Z}[x]$ takové, že $1 = u \cdot x + v \cdot 2$ — absolutní člen polynomu na pravé straně je nutně sudý.

Příklad. Obor $\mathbb{Q}[x, y]$ není eukleidovský, protože v něm neplatí Bézoutova rovnost. Např. pro polynomy $f = x$ a $g = y$ platí $\mathrm{NSD}(x, y) = 1$, ale přitom neexistuje $u, v \in \mathbb{Z}[x]$ takové, že $1 = u \cdot x + v \cdot y$ — absolutní člen polynomu na pravé straně je nutně nula.

9.2. **Rozklady na ireducibilní činitele.**

Nyní dokážeme, že v eukleidovských oborech existují jednoznačné rozklady na ireducibilní prvky.

Lemma 9.2. *Buď* $\mathbf{R}$ *eukleidovský obor a* $a, b \in R$, $a, b \neq 0$.

(1) *Pokud* $a \parallel b$, *pak* $\nu(a) = \nu(b)$.
(2) *Pokud* $a \mid b$ *a* $a \nmid b$, *pak* $\nu(a) < \nu(b)$.

Samotná implikace $\nu(a) = \nu(b) \Rightarrow a \parallel b$ neplatí: např. v oborech polynomů jsou jistě neasociované polynomy stejného stupně!

Důkaz. (1) Je-li $a \parallel b$, tedy $a \mid b$ a $b \mid a$, pak $\nu(a) \leq \nu(b) \leq \nu(a)$, tedy $\nu(a) = \nu(b)$. (2) Napišme

- $b = au$ pro nějaké $u \in R$,
- $a = bq + r$ pro nějaká $q, r \in R$, $\nu(r) < \nu(b)$.

Vzhledem k tomu, že $b \nmid a$, platí $r \neq 0$. Dosazením získáme vyjádření $r = a - bq = a - auq = a(1 - uq)$, z kterého plyne, že $a \mid r$. Protože $r \neq 0$, dostáváme $\nu(a) \leq \nu(r) < \nu(b)$. □

Důsledek 9.3. *Eukleidovské obory jsou gaussovské.*

Důkaz. Podle Věty 8.5 stačí dokázat, že v eukleidovských oborech existují NSD a neexistují nekonečné posloupnosti vlastních dělitelů. První fakt jsme dokázali ve Větě 9.1 a druhý plyne bezprostředně z bodu (2) předešlého lemmatu. □

Důsledkem této věty je fakt, že např. obor $\mathbb{Z}[i]$ je gaussovský, a také obory $\mathbf{T}[x]$, $\mathbf{T}$ těleso, jsou gaussovské. V Sekci **??** tento fakt využijeme k důkazu Gaussovy věty.

9.3. Aplikace: řešení diofantických rovnic.

Zatím jsme pomíjeli motivaci ke studiu rozšíření celých čísel. Teď, když máme dokázány všechny základní vlastnosti, si můžeme ukázat, jak je využít k řešení jistého typu *diofantických rovnic*. Asi nejznámější takovou rovnicí je Velká Fermatova věta, tedy tvrzení, že neexistují nenulová celá čísla x, y, z splňující $x^n + y^n = z^n$ pro nějaké $n \geq 3$. Už Leonhard Euler použil v roce 1753 počítání v Eisensteinových číslech k řešení této rovnice pro $n = 3$, řadu dalších exponentů vyřešil Ernst Kummer v polovině 19. století pomocí rozkladu

$$x^n - z^n = (x - z)(x - \omega z)(x - \omega^2 z) \cdots (x - \omega^{n-1} z)$$

v oboru $\mathbb{Z}[\omega]$, kde $\omega = e^{2\pi i/n}$. K důkazu Velké Fermatovy věty nakonec vedla úplně jiná metoda, ale to už je jiná historka. Jako ilustraci si ukážeme řešení jedné speciální diofantické rovnice.

Úloha. Řešte v oboru celých čísel rovnici

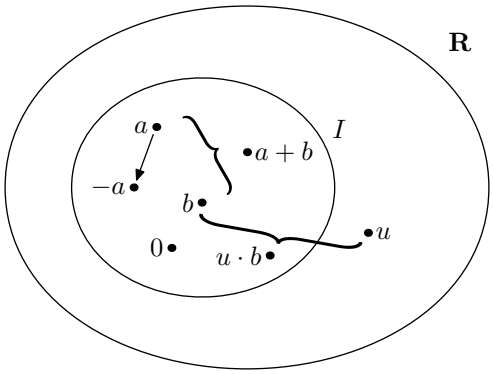
$$x^2 + 1 = y^3.$$

Řešení. Uvažujme řešení $x, y \in \mathbb{Z}$ a počítejme v oboru $\mathbb{Z}[i]$. Nejprve rozložíme $x^2 + 1 = (x + i)(x - i)$ a dokážeme, že jsou čísla $x + i, x - i$ nesoudělná. Podle pravidla $\mathrm{NSD}(a, b) = \mathrm{NSD}(a, a - b)$ dostáváme

$$\mathrm{NSD}(x + i, x - i) = \mathrm{NSD}(x + i, 2i) = \mathrm{NSD}(x - i, 2i),$$

a protože číslo $2i$ má ireducibilní rozklad $(1 + i)^2$, musí být výsledek jedno z čísel $1, 1 + i, (1 + i)^2$. Pokud je x sudé, pak je $\nu(x + i)$ liché, a tedy $\mathrm{NSD}(x + i, x - i) = 1$. Pokud je x liché, pak je $\nu(x + i) = \nu(x - i) \equiv 2 \pmod 4$ (dosaďte $x = 2k + 1$), a tedy $(1 + i)^2$ nedělí $x + i$ ani $x - i$ (tj. v ireducibilním rozkladu těchto čísel je $1 + i$ nejvýše jednou). Protože je součin $(x + i)(x - i)$ třetí mocninou, počet prvočísel $1 + i$ v jeho ireducibilním rozkladu musí být dělitelný třemi; čili jediná možnost je, že tam není žádné. Tedy $\mathrm{NSD}(x + i, x - i) = 1$.

Dokázali jsme, že $x + i$ a $x - i$ jsou nesoudělné v $\mathbb{Z}[i]$. Protože jejich součin je třetí mocninou čísla y, každé z nich musí být třetí mocninou nějakého prvku $\mathbb{Z}[i]$. Uvažujme takové $a + bi$: z rovnosti $(a + bi)^3 = (a^3 - 3ab^2) + (3a^2 b - b^3)i = x + i$



Obrázek 10. Ideál I v oboru $\mathbf{R}$.

plyne $b(3a^2 - b^2) = 1$, což má jediné celočíselné řešení: $b = -1$, $a = 0$. To dává jediné celočíselné řešení původní rovnice $x = 0$, $y = 1$. □

10. Ideály a dělitelnost

Definice. Buď $\mathbf{R}$ komutativní okruh. *Ideálem* v $\mathbf{R}$ nazýváme každou podmnožinu $I \subseteq R$ takovou, že

- $0 \in I$;
- pokud $a, b \in I$, pak $-a \in I$ a $a + b \in I$;
- pokud $a \in I$ a $r \in R$, pak $r \cdot a \in I$.

Všimněte si, že množiny $\{0\}$ a R jsou ideály v libovolném okruhu $\mathbf{R}$, říká se jim *nevlastní*.

Příklad. Uvažujme obor $\mathbb{Z}$. Množiny $n\mathbb{Z} = \{nz : \ z \in \mathbb{Z}\} = \{u \in \mathbb{Z} : \ n \mid u\}$ jsou ideály v $\mathbb{Z}$. (Z Věty 10.2 plyne, že žádné jiné ideály v oboru $\mathbb{Z}$ nejsou.)

Konstrukci ideálů z předchozího příkladu lze zobecnit.

Tvrzení 10.1. *Buď $\mathbf{R}$ komutativní okruh a $a \in R$. Pak*

$$aR = \{ar : \ r \in R\} = \{u \in R : \ a \mid u\}$$

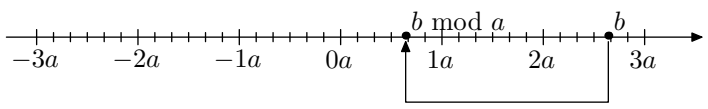
tvoří ideál. Je to nejmenší ideál (nejmenší vzhledem k inkluzi) obsahující prvek a.

Tento ideál se nazývá *hlavní ideál* generovaný prvkem a.

Důkaz. Je zřejmé, že jde skutečně o ideál: $a \mid 0$, tedy $0 \in aR$, součet i rozdíl dvou prvků dělitelných a je dělitelný a, a pokud $a \mid u$, pak $a \mid ru$ pro libovolné $r \in R$. Buď I libovolný ideál obsahující prvek a. Pak I jistě obsahuje i všechny jeho násobky, tedy $aR \subseteq I$, čili aR je nejmenší ideál obsahující prvek a. □

Komutativní okruhy, které neobsahují jiné ideály než hlavní, nazýváme *okruhy hlavních ideálů*; v případě oborů integrity hovoříme o *oborech hlavních ideálů*. Nejdůležitější příklady popisuje následující věta.

Věta 10.2. *V eukleidovských oborech je každý ideál hlavní.*



Obrázek 11. Ilustrace důkazu Věty 10.2 v případě $\mathbf{R} = \mathbb{Z}$.

Důkaz. Buď I ideál v eukleidovském oboru $\mathbf{R}$. Je-li $I = \{0\}$, pak $I = 0R$. V opačném případě označme a takový prvek ideálu I, který má nejmenší nenulovou eukleidovskou normu (libovolný z nich, je-li jich více). Dokážeme, že $I = aR$. Zřejmě $aR \subseteq I$, pro spor tedy předpokládejme, že existuje nějaký prvek $b \in I \setminus aR$. Zvolme q, r splňující $b = aq + r$ a $\nu(r) < \nu(a)$. Samozřejmě $r \neq 0$, protože b není dělitelné a, a tedy $0 < \nu(r) < \nu(a)$. Ovšem

$$r = \underbrace{b}_{\in I} - \underbrace{aq}_{\in I} \in I,$$

což je spor s výběrem a jako prvku I s nejmenší kladnou normou. □

Opačná implikace neplatí, ale vymyslet nějaký protipříklad není snadné: asi nejjednodušším příkladem je obor $\mathbb{Z}[\frac{1+i\sqrt{19}}{2}]$. Důkaz tohoto faktu je poměrně obtížný.

Pro tělesa platí ještě silnější vlastnost. Tento fakt se nám bude hodit později, až budeme konstruovat tělesa jako faktorokruhy (viz Sekce **??**).

Tvrzení 10.3. *Buď* $\mathbf{R}$ *komutativní okruh s jednotkou. Pak* $\mathbf{R}$ *je těleso právě tehdy, když má pouze nevlastní ideály.*

Důkaz. ($\Rightarrow$) Buď I ideál v $\mathbf{R}$ a předpokládejme, že $I \neq \{0\}$. Zvolme libovolné $0 \neq a \in I$. Pak pro každé $b \in R$ platí $b = a \cdot (a^{-1} \cdot b) \in I$, a tedy $I = R$.

($\Leftarrow$) Ke každému $0 \neq a \in R$ hledáme prvek $b \in R$ takový, že $a \cdot b = 1$. Uvažujme hlavní ideál aR. Ten obsahuje prvek a, čili je různý od $\{0\}$, a tudíž podle předpokladu $aR = R$. Speciálně $1 \in aR$, tj. existuje $b \in R$ splňující $1 = a \cdot b$. □

V Sekci 9 jsme ukázali, že obory $\mathbb{Z}[x]$ ani obory polynomů více proměnných nejsou eukleidovské. Ukážeme, že to dokonce nejsou ani obory hlavních ideálů. Oba důkazy jsou založené na následující myšlence. Hlavní ideál aR, který obsahuje dva nesoudělné prvky u, v, je roven celému R: je-li $u, v \in aR$, tj. $a \mid u$ i $a \mid v$, pak musí být $a \parallel 1$, z čehož plyne $aR = R$. Toto pozorování lze snadno použít k hledání ideálů, které nejsou hlavní.

Příklad. Obor $\mathbb{Z}[x]$ není obor hlavních ideálů. Uvažujme množinu

$$I = \{f \in \mathbb{Z}[x] : \ f(0) \text{ je sudé}\} \subset \mathbb{Z}[x].$$

Je vidět, že jde o ideál. Přitom I obsahuje polynomy 2 a x, které jsou nesoudělné, nemůže tedy být hlavní.

Příklad. Obor $\mathbf{R}[x_1, \ldots, x_k]$ (kde $\mathbf{R}$ je libovolný obor integrity a $k > 1$) není obor hlavních ideálů. Uvažujme množinu

$$I = \{f \in R[x_1, \ldots, x_k] : \ f(0, \ldots, 0) = 0\} \subset R[x_1, \ldots, x_k].$$

Je vidět, že jde o ideál. Přitom I obsahuje polynomy x_1 a x_2, které jsou nesoudělné, nemůže tedy být hlavní.

Hlavní ideály hrají v teorii dělitelnosti důležitou roli z následujícího důvodu:

- $a \mid b$ právě tehdy, když $bR \subseteq aR$;
- $a \parallel b$ právě tehdy, když $aR = bR$.

Obory hlavních ideálů jsou důležitou třídou oborů integrity z toho důvodu, že struktura ideálů věrně odráží pojem dělitelnosti až na asociovanost. Jako ukázku práce s hlavními ideály si ukážeme větu, která obory hlavních ideálů zařazuje do hierarchie oborů z hlediska teorie dělitelnosti.

Věta 10.4. *Obory hlavních ideálů jsou gaussovské a platí v nich Bézoutova rovnost.*

Důkaz. Buď $\mathbf{R}$ obor hlavních ideálů. Podle Věty 8.5 stačí dokázat, že v $\mathbf{R}$ (1) existují NSD a (2) neexistují nekonečné posloupnosti vlastních dělitelů. Připomeňme, že pro libovolná u, v platí $u \mid v \Leftrightarrow vR \subseteq uR$.

(1) Zvolme $a, b \in R$ a označme I nejmenší ideál obsahující množinu $aR \cup bR$. Existuje tedy $c \in R$ takové, že $I = cR$. Protože $aR \subseteq cR$, máme $c \mid a$, a analogicky $c \mid b$. Přitom pokud je d společným dělitelem a, b, pak $aR \subseteq dR$ a $bR \subseteq dR$, tedy $I = cR \subseteq dR$ a dostáváme $d \mid c$. Tedy $c = \mathrm{NSD}(a, b)$.

(2) Pro spor předpokládejme, že v $\mathbf{R}$ existuje nekonečná posloupnost vlastních dělitelů $a_1, a_2, \ldots$ (tj. $a_{i+1} \mid a_i$ a $a_i \nmid a_{i+1}$). Pak $a_1R \subset a_2R \subset a_3R \subset \ldots$ a označme $I = \bigcup_{i=1}^{\infty} a_iR$. Tato množina také tvoří ideál (dokáže se podobně jako Tvrzení **??**), takže $I = bR$ pro nějaké $b \in I$. Ovšem protože $b \in I = \bigcup_{i=1}^{\infty} a_iR$, existuje i takové, že $b \in a_iR$. Pak ale $bR = a_iR = a_{i+1}R = \ldots$, spor.

K dokázání Bézoutovy rovnostu stačí nahlédnout, že nejmenší ideál obsahující množinu $aR \cup bR$ je ideál $aR + bR$. Protože tento ideál obsahuje $\mathrm{NSD}(a, b)$, dostáváme $\mathrm{NSD}(a, b) = au + bv$ pro nějaká $u, v \in R$. □

SHRNUTÍ

V celé kapitole jsme dokázali následující hierarchii oborů integrity:

eukleidovský obor $\Longrightarrow$ *obor hlavních ideálů* $\Longrightarrow$ *gaussovský obor*

Některé vlastnosti těchto tříd jsou shrnuty v následující tabulce:

obory	ired. rozklady	ex. NSD	Bézout. rovnost	Eukleidův alg.
eukleidovské	Věta 9.3	Věta 9.1	Věta 9.1	Věta 9.1
hlavních ideálů	Věta 10.4	Věta 10.4	ano	NE
gaussovské	definice	Věta 8.3	NE	NE
obecné	NE	NE	NE	NE

A na závěr pár příkladů, které stojí za zapamatování.

eukleidovské	tělesa, $\mathbb{Z}$, $\mathbf{T}[x]$ ($\mathbf{T}$ těleso), $\mathbb{Z}[i]$, $\mathbb{Z}[\sqrt{2}]$, $\mathbb{Z}[i\sqrt{2}]$
hlavních ideálů, neeukleidovské	$\mathbb{Z}[\frac{1+i\sqrt{19}}{2}]$
gaussovské, ne hlavních ideálů	$\mathbb{Z}[x]$, $\mathbf{R}[x, y, \ldots]$ ($\mathbf{R}$ gaussovský)
negaussovské	$\mathbb{Z}[\sqrt{5}]$, $\mathbb{Z}[i\sqrt{3}]$

Rozšíření těles

11. Počítání modulo a konstrukce těles

Tělesem rozumíme komutativní okruh s jednotkou, jehož každý nenulový prvek je invertibilní. Mezi nejdůležitější příklady patří tělesa racionálních čísel, komplexních čísel a jejich mezitělesa, především ta, jejichž stupeň rozšíření nad $\mathbb{Q}$ je konečný; jim bude věnována Sekce 12. Druhou důležitou rodinou jsou konečná tělesa, jež stručně zmíníme v Sekci **??**.

Každý obor integrity lze rozšířit do tzv. *podílového tělesa* pomocí konstrukce analogické konstrukci zlomků. Důležitým příkladem jsou tělesa racionálních funkcí. Druhou stěžejní konstrukcí jsou tzv. faktorokruhy, viz Sekce 11; typickým přikladem je počítání s polynomy modulo nějaký ireducibilní polynom. Tyto konstrukce jsou zásadní pro kontrukci různých nadtěles se specifickými vlastnostmi (např. nadtělesa, ve kterém má daný polynom kořen).

Buď $\mathbf{R}$ komutativní okruh s jednotkou a $\mathbf{I}$ jeho hlavní ideál, tj. $I = mR$ pro nějaké $m \in R$. Definujeme ekvivalenci na množině R předpisem $a \sim_m b$ právě tehdy, když $m \mid a - b$, což symbolicky značíme $a \equiv b \pmod m$. *Faktorokruhem* komutativního okruhu $\mathbf{R}$ podle mR rozumíme množinu $R/\!\sim = \{[a]_\sim : a \in R$ všech bloků ekvivalence $\sim_m$, spolu s operacemi definovanými předpisy

$$[a] + [b] = [a + b], \quad -[a] = [-a], quad[a] \cdot [b] = [a \cdot b]$$

a konstantami [0] a [1]. Důležitým pozorováním je, že touto konstrukcí dostaneme opět komutativní okruh s jednotkou (DOPLNIT DůKAZ). Faktorokruh $\mathbf{R}$ podl mR zapisujeme zkráceně $\mathbf{R}/(m)$.

Následující pozorování je stěžejní pro většinu aplikací. Je-li v okruhu $\mathbf{R}$ definováno dělení se zbytkem (např. celá čísla, polynomy) , prvky $\mathbf{R}/(m)$ můžeme reprezentovat jako všechny možné zbytky po dělení prvkem m, přičemž operace v $\mathbf{R}/(m)$ budou jako operace v původním okruhu modulo m:

$$[a] \pm [b] = [a \pm b] = [a \pm b \bmod m], \qquad [a] \cdot [b] = [a \cdot b] = [a \cdot b \bmod m].$$

Příklad. Obor $\mathbb{Z}$ je eukleidovský, s jednoznačně definovaným podílem a zbytkem. Prvky faktorokruhu $\mathbb{Z}/(n)$ tedy můžeme reprezentovat jako všechny možné zbytky po dělení číslem n, tj. jako čísla $0, \dots, n-1$, přičemž operace provádíme modulo n. Je vidět, že dostaneme okruh $\mathbb{Z}_n$, při ztotožnění čísla a s blokem $[a]$.

Příklad. Obor $\mathbf{T}[x]$, $\mathbf{T}$ těleso, je eukleidovský, s jednoznačně definovaným podílem a zbytkem. Prvky faktorokruhu $\mathbf{T}[x]/(f)$, kde $f \in T[x]$, tedy můžeme reprezentovat jako všechny možné zbytky po dělení polynomem f, tj. jako všechny polynomy stupně menšího než $\deg f$, přičemž operace provádíme modulo f.

Věta 11.1. *Buď $\mathbf{R}$ obor integrity hlavních ideálů. Pak $\mathbf{R}/(m)$ je těleso právě tehdy, když m je ireducibilní prvek.*

Důkaz. ($\Rightarrow$) Kdyby v $\mathbf{R}$ platilo $m = a \cdot b$, kde $a, b \nmid 1$, pak by v $\mathbf{R}/(m)$ platilo $[0] = [a \cdot b] = [a] \cdot [b]$, a tedy by $\mathbf{R}/(m)$ nebyl ani oborem integrity, natož tělesem (viz Tvrzení 4.3).

($\Leftarrow$) Buď $[a] \in R/(m)$, $[a] \neq [0]$, hledáme jeho inverz. Protože v $\mathbf{R}$ platí Bézoutova rovnost (Věta 10.4), existují $u, v \in R$ taková, že $\mathrm{NSD}(a, m) = ua + vm$. Přitom $\mathrm{NSD}(a, m) = 1$, protože m je ireducibilní a $m \nmid a$ (což je ekvivalentní vyjádření $a \not\sim_m 0$). Čili $1 = ua + vm \equiv ua \pmod{m}$, a tedy $[u]$ je inverzním prvkem k $[a]$. □

Příklad. Faktorokruh $\mathbb{Z}/(n) \simeq \mathbb{Z}_n$ je těleso právě tehdy, když n je prvočíslo.

Příklad. Obor $\mathbb{Z}[x]$ není oborem hlavních ideálů, čili Věta 11.1 nedává žádnou garanci, že faktorokruhy podle ireducibilních prkvů dávají tělesa. A skutečně, ačkoliv je polynom $x - 1$ ireducibilní, faktorokruh $\mathbb{Z}[x]/(x - 1)$ není tělesem, např. prvek $[2]$ nemá inverz.

Příklad. Uvažujme ireducibilní polynom $f \in \mathbb{Z}_p[x]$, p prvočíslo, stupně k. Protože je $\mathbb{Z}_p[x]$ oborem hlavních ideálů, faktorokruh $\mathbb{Z}_p[x]/(f)$ je tělesem. Jeho prvky lze reprezentovat jako polynomy stupně $< k$. Tyto mají přesně k koeficientů ze $\mathbb{Z}_p$, a tedy $\mathbb{Z}_p[x]/(f)$ je *konečným tělesem*, které má p^k prvků. Například:

- $\mathbb{F}_p = \mathbb{Z}_p$.
- $\mathbb{F}_4 = \mathbb{Z}_2[x]/(x^2 + x + 1)$, $\mathbb{F}_8 = \mathbb{Z}_2[x]/(x^3 + x + 1)$, $\mathbb{F}_9 = \mathbb{Z}_3[x]/(x^2 + 1)$.
- $\mathbb{F}_{p^k}$ není ani $\mathbb{Z}_{p^k}$, ani $(\mathbb{Z}_p)^k$, protože to vůbec nejsou tělesa!

V Sekci **??** o konečných tělesech dokážeme následující netriviální fakta: pro každé p, k takový ireducibilní polynom existuje, na jeho volbě (až na izomorfismus zkonstruovaných těles) nezáleží a každé konečné těleso lze tímto způsobem zkonstruovat.

12. Algebraická rozšíření

12.1. Motivace: algebraická a transcendentní čísla.

Jako motivaci ke studiu rozšíření těles začneme s klasickým tématem 19. století: která čísla jsou kořenem nějakého celočíselného polynomu a jak takový polynom najít? To nebývá úplně snadné, důkazy transcendence čísel jako e nebo π patří k významnějším výsledkům matematiky 19. století. Ukážeme si geniální Cantorovu myšlenku, která ukazuje, že skoro každé číslo je transcendentní, aniž bychom museli uvést byť jediný příklad. Jde o jeden z argumentů, který formoval moderní teorii množin. V dalších sekcích pak vyložíme systematický algebraický přístup, který, mimo jiné, umožňuje popsat řadu vlastností algebraických čísel.

Definice. Reálné číslo a se nazývá *algebraické*, pokud existuje nenulový celočíselný polynom f takový, že $f(a) = 0$. V opačném případě se a nazývá *transcendentní*.

Příklad. Spousta čísel „ze života“ je algebraických:

- Racionální čísla jsou algebraická, racionální číslo $\frac{a}{b}$ je kořenem polynomu $bx - a$.
- Některá iracionální čísla jsou algebraická, např. $\sqrt{2}$ je kořenem polynomu $x^2 - 2$.
- Leckterá iracionální jsou algebraická, i když to není vidět na první pohled, např. $\sqrt{2} + \sqrt{3}$. Z Věty 12.10 plyne, že součet, rozdíl, součin a podíl algebraických čísel je algebraické číslo.

Příklad. Ač matematici dlouho tušili, že je řada čísel transcendentních, nedařilo se jim tuto vlastnost o žádném čísle dokázat.

- První prokazatelně transcendentní číslo předvedl v roce 1840 francouzský matematik Joseph Liouville: byl jím součet řady $\sum_{i=0}^{\infty} 10^{-i!}$, tj. číslo, které má v desetinném rozvoji jedničku právě na pozicích tvaru $i!$, jinak nuly.
- V roce 1873 dokázal Charles Hermite, že číslo e je transcendentní, a až v roce 1882 našel Ferdinand von Lindemann důkaz transcendence čísla π.
- O to více udivil matematiky v roce 1874 Georg Cantor, když dokázal, že *skoro všechna reálná čísla jsou transcendentní.*

Ač všechny důkazy transcendence konkrétních čísel jako e nebo π jsou poměrně komplikované, Cantorův důkaz je překvapivě jednoduchý.

Spočetnou množinou rozumíme takovou nekonečnou množinu, jejíž prvky lze seřadit do posloupnosti indexované přirozenými čísly (tj. jde o množinu stejně velkou jako $\mathbb{N}$). Všechny ostatní (tj. větší) nekonečné množiny nazýváme *nespočetné.*

Např. množina $\mathbb{Z}$ je spočetná: $0, 1, -1, 2, -2, 3, -3, \ldots$ Dokonce i množina $\mathbb{Q}$ je spočetná: seřaďte kladná racionální čísla do posloupnosti podle součtu čitatele a jmenovatele (ty se stejným součtem seřaďte libovolně) a vložte záporná čísla analogickým trikem.

Tvrzení 12.1. *Množina algebraických reálných čísel je spočetná.*

Důkaz. Definujme *index polynomu* $f = a_0 + a_1 x + \ldots + a_n x^n \neq 0$ jako součet $|a_0| + |a_1| + \ldots + |a_n| + n$. Všimněte si, že existuje jen konečně mnoho polynomů daného indexu (např. index 1: $f = \pm 1$; index 2: $f = \pm 2$, $f = \pm x$; index 3: $f = \pm 3$, $f = \pm 2x$, $f = \pm x \pm 1$, $f = \pm x^2$), všechny celočíselné polynomy tedy lze seřadit do posloupnosti podle vzrůstajícího indexu. Přitom každý nenulový polynom má jen konečně mnoho kořenů, tedy nahrazením polynomu za jeho kořeny získáme posloupnost obsahující všechna algebraická čísla. □

Tvrzení 12.2. *Množina reálných čísel je nespočetná.*

Důkaz. Kdyby byla množina reálných čísel spočetná, byl by jistě spočetný i interval $\langle 0, 1)$, a tudíž bychom mohli seřadit čísla z tohoto intervalu do posloupnosti

$$\begin{aligned} a_1 &= 0, a_{11} a_{12} a_{13} \ldots \\ a_2 &= 0, a_{21} a_{22} a_{23} \ldots \\ a_3 &= 0, a_{31} a_{32} a_{33} \ldots \\ &\ldots \end{aligned}$$

Nyní definujme číslo $b = 0, b_1 b_2 b_3 \ldots$ tak, že $b_1 \neq a_{11}$, $b_2 \neq a_{22}$, atd. Toto číslo nemůže být na seznamu, neboť se od i-tého prvku liší v i-té pozici rozvoje. Což je spor s tím, že tam měla být všechna čísla z intervalu $\langle 0, 1)$. (K tomu, aby byl tento argument korektní, je třeba se vyhnout rozvojům končícím samými devítkami.) □

Tedy reálných čísel je mnohem více než algebraických; z toho důvodu musí existovat nějaká transcendentní čísla. Vhledem k tomu, že spočetné množiny mají míru 0, tvrzení lze dokonce interpretovat tak, že skoro všechna reálná čísla jsou transcendentní (ve smyslu: náhodné reálné číslo je s pravděpodobností 1 transcendentní).

12.2. **Okruhová a tělesová rozšíření.**

Buď $\mathbf{R} \leq \mathbf{S}$ komutativní okruhy a $a_1, \dots, a_n \in S$. Připomeňme, že nejmenší podokruh okruhu $\mathbf{R}$ obsahující danou množinu $X \subseteq R$ se nazývá *podokruh generovaný množinou* X a značí se $\langle X \rangle_{\mathbf{R}}$. Definujeme

$$\mathbf{R}[a_1, \dots, a_n] = \langle R \cup \{a_1, \dots, a_n\} \rangle_{\mathbf{S}}$$

a hovoříme o *okruhovém rozšíření* $\mathbf{R}$ *o prvky* $a_1, \dots, a_n$.

Nyní uvažujme tělesa $\mathbf{T} \leq \mathbf{S}$ a $a_1, \dots, a_n \in S$. Definujeme

$$\mathbf{T}(a_1, \dots, a_n)$$

jako nejmenší podtěleso obsahující množinu $T \cup \{a_1, \dots, a_n\}$ a hovoříme o *tělesovém rozšíření* $\mathbf{R}$ *o prvky* $a_1, \dots, a_n$.

Zřejmě $\mathbf{T}[a_1, \dots, a_n] \leq \mathbf{T}(a_1, \dots, a_n)$. Nemalá část této sekce se bude zabývat problémem, za jakých podmínek jsou obě rozšíření totožná.

Příklad. $\mathbb{Q}[\sqrt{2}] = \mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\}$. DOPSAT

Příklad. $\mathbb{Q}[\pi] \neq \mathbb{Q}(\pi)$. DOPSAT

Tvrzení 12.3. *Buď* $\mathbf{R} \leq \mathbf{S}$ *komutativní okruhy a* $a_1, \dots, a_n \in S$. *Pak*

$$\mathbf{R}[a_1, \dots, a_n] = \{f(a_1, \dots, a_n) : \ f \in R[x_1, \dots, x_n]\}.$$

Důkaz. Označme $M = \{f(a_1, \dots, a_n) : \ f \in R[x_1, \dots, x_n]\}$. Je třeba dokázat, že

(1) množina M obsahuje $R \cup \{a_1, \dots, a_n\}$,
(2) všechny prvky množiny M lze nagenerovat z prvků $R \cup \{a_1, \dots, a_n\}$,
(3) množina M je uzavřená na všechny operace okruhu $\mathbf{S}$.

(1) Prvky R dostaneme skrze konstantní polynomy, prvek a_i pomocí polynomu $x_i \in S[x_1, \dots, x_n]$. (2) Je-li $f = \sum c_{k_1,\dots,k_n} x_1^{k_1} \cdots x_n^{k_n}$ polynom z $R[x_1, \dots, x_n]$, pak $f(a_1, \dots, a_n) = \sum c_{k_1,\dots,k_n} \cdot a_1^{k_1} \cdot \ldots \cdot a_n^{k_n}$ je prvek podokruhu $\mathbf{R}[a_1, \dots, a_n]$, neboť jde o součet součinů prvků $c_{k_1,\dots,k_n} \in R$ a $a_1^{k_1} \cdot \ldots \cdot a_n^{k_n} \in \langle a_1, \dots, a_n \rangle$. (3) Označme $\bar{a} = (a_1, \dots, a_n)$. Pak $0 = 0(\bar{a})$ a je-li $f(\bar{a}), g(\bar{a}) \in M$, pak $-f(\bar{a}) = (-f)(\bar{a}) \in M$, $f(\bar{a}) + g(\bar{a}) = (f + g)(\bar{a}) \in M$ a $f(\bar{a}) \cdot g(\bar{a}) = (f \cdot g)(\bar{a}) \in M$. □

Příklady.

- Protože $\sqrt{2}^2 = 2 \in \mathbb{Q}$, a tedy $\sqrt{2}^k \in \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\}$ pro všechna k, hodnota polynomu $f \in \mathbb{Q}[x]$ v bodě $\sqrt{2}$ je rovna nějakému číslu tvaru $a + b\sqrt{2}$, $a, b \in \mathbb{Q}$. Tedy

$$\mathbb{Q}[\sqrt{2}] = \{f(\sqrt{2}) : \ f \in \mathbb{Q}[x]\} = \{f(\sqrt{2}) : \ f \in \mathbb{Q}[x], \deg f \leq 1\} = \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\}.$$

- Podobně, protože $(\sqrt[3]{2})^3 \in \mathbb{Q}$, platí

$$\mathbb{Q}[\sqrt[3]{2}] = \{f(\sqrt[3]{2}) : \ f \in \mathbb{Q}[x]\} = \{f(\sqrt[3]{2}) : \ f \in \mathbb{Q}[x], \deg f \leq 2\} = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} : \ a, b, c \in \mathbb{Q}\}.$$

- Podobně také

$$\begin{aligned}\mathbb{Q}[\sqrt{2}, \sqrt{3}] = \{f(\sqrt{2}, \sqrt{3}) : \ f \in \mathbb{Q}[x, y]\} &= \{f(\sqrt{2}, \sqrt{3}) : \ f = a + bx + cy + dxy \in \mathbb{Q}[x, y]\} \\ &= \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : \ a, b, c, d \in \mathbb{Q}\}.\end{aligned}$$

Příklad. $\mathbb{Q}(\pi)$. DOPSAT

Tvrzení 12.4. *Buď* $\mathbf{T} \leq \mathbf{S}$ *tělesa a* $a_1, \dots, a_n \in S$. *Pak*

$$\mathbf{T}(a_1, \dots, a_n) = \{\frac{f(a_1, \dots, a_n)}{g(a_1, \dots, a_n)} : \ f, g \in T[x_1, \dots, x_n], g(a_1, \dots, a_n) \neq 0\}.$$

Důkaz se provede analogicky.

Příklad. $\mathbb{Q}(\sqrt{2}) = \mathbb{Q}[\sqrt{2}]$, protože $\frac{1}{g(\sqrt{2})} \in \mathbb{Q}[\sqrt{2}]$. DOPSAT

Příklad. $\mathbb{Q}(\pi)$. DOPSAT

12.3. **Stupeň rozšíření a algebraické prvky.**

Klíčem k pochopení této sekce je myšlenka, že nadtěleso $\mathbf{S} \geq \mathbf{T}$ lze považovat za vektorový prostor nad tělesem $\mathbf{T}$: sčítání a odčítání přebereme beze změny a místo násobení jako operace $S \times S \to S$ uvažujeme pouze restrikci $T \times S \to S$, tj. násobíme prvky většího tělesa $\mathbf{S}$ (vektory) pouze prvky menšího tělesa $\mathbf{T}$ (skaláry). Tento vektorový prostor budeme značit $\mathbf{S_T} = (S, +, -, 0, a\cdot : a \in T)$, jeho dimenze se nazývá *stupeň rozšíření* $\mathbf{T} \leq \mathbf{S}$ a značí se $[\mathbf{S} : \mathbf{T}] = \dim \mathbf{S_T}$.

Příklady.

- $[\mathbb{C} : \mathbb{R}] = 2$. Prvek $a+bi \in \mathbb{C}$ lze považovat za dvojdimenzionální vektor nad $\mathbb{R}$, sčítání i násobení reálným číslem probíhá po složkách. Báze prostoru $\mathbb{C}_\mathbb{R}$ je např. $1, i$.
- $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$. Prvek $a + b\sqrt[3]{2} + c\sqrt[3]{4} \in \mathbb{Q}(\sqrt[3]{2}) = \mathbb{Q}[\sqrt[3]{2}]$ lze považovat za třídimenzionální vektor nad $\mathbb{Q}$ ze stejného důvodu. Báze prostoru $\mathbb{Q}(\sqrt[3]{2})_\mathbb{Q}$ je např. 1, $\sqrt[3]{2}$, $\sqrt[3]{4}$.
- $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$, báze prostoru $\mathbb{Q}(\sqrt{2}, \sqrt{3})_\mathbb{Q}$ je např. $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$.
- Rozšíření mohou mít i nekonečný stupeň: např. stupeň $[\mathbb{Q}(\pi) : \mathbb{Q}]$ je spočetný, zatímco stupeň $[\mathbb{R} : \mathbb{Q}]$ je nespočetný (kontinuum).

Poznámka. Počet prvků libovolného konečného tělesa je mocnina prvočísla, neboť je vektorovým prostorem nad nějakým tělesem $\mathbb{Z}_p$. Uvažujme nějaké konečné těleso $\mathbf{T}$, buď p jeho charakteristika. Není těžké ověřit, že množina

$$\{\underbrace{1 + \ldots + 1}_{k} : \ k = 0, \ldots, p-1\}$$

je uzavřená na všechny tělesové operace, čili tvoří p-prvkové podtěleso $\mathbf{S}$ (při ztotožnění čísla k a k-násobného součtu $1 + \ldots + 1$ vidíme, že jde o těleso $\mathbb{Z}_p$). Tedy $\mathbf{S} \leq \mathbf{T}$ je tělesovým rozšířením a vektorový prostor $\mathbf{T_S}$ bude mít p^k prvků, pro nějaké přirozené číslo k.

Je-li stupeň $[\mathbf{S} : \mathbf{T}]$ konečný, říkáme, že jde o rozšíření *konečného stupně*.

Definice. Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles a $a \in S$. Řekneme, že prvek a je *algebraický* nad $\mathbf{T}$, pokud existuje nenulový polynom z $\mathbf{T}[x]$, jehož je a kořenem. V opačném případě se prvek a nazývá *transcendentní* nad $\mathbf{T}$. Je-li každý prvek tělesa $\mathbf{S}$ algebraický nad $\mathbf{T}$, hovoříme o *algebraickém rozšíření*.

Tvrzení 12.5. *Rozšíření konečného stupně jsou algebraická.*

Důkaz. Označme $n = [\mathbf{S} : \mathbf{T}]$ a uvažujme libovolný prvek $a \in S$; dokážeme, že prvek a je algebraický nad $\mathbf{T}$. Prvky $1, a, a^2, \ldots, a^{n-1}, a^n$ jsou lineárně závislé, protože jich je více než je dimenze vektorového prostoru $\mathbf{S_T}$. Tedy existují koeficienty $b_i \in T$, aspoň jeden z nich nenulový, kterými lze lineárně nakombinovat nulu, tj. $\sum_{i=0}^{n} b_i a^i = 0$. Prvek a je tedy kořenem nenulového polynomu $\sum_{i=0}^{n} b_i x^i \in T[x]$. □

Opačná implikace neplatí: mohou existovat i velká algebraická rozšíření, např. taková, kde má každý polynom kořen (např. algebraický uzávěr tělesa $\mathbb{Q}$ nemá konečný stupeň nad $\mathbb{Q}$). Pokud ovšem rozšiřujeme těleso $\mathbf{T}$ o jediný algebraický prvek,

nebo o konečné množství algebraických prvků, stupeň konečný je (viz Tvrzení 12.7, 12.8). Stupeň rozšíření o jeden algebraický prvek je dán stupněm tzv. minimálního polynomu.

Definice. Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles a $a \in S$ algebraický prvek nad $\mathbf{T}$. *Minimálním polynomem* prvku a nad $\mathbf{T}$ rozumíme monický polynom $m_{a,\mathbf{T}} \in T[x]$ splňující

(1) $m_{a,\mathbf{T}}(a) = 0$;
(2) kdykoliv je a kořenem polynomu $f \in T[x]$, pak $m_{a,\mathbf{T}} \mid f$.

Existuje takový polynom pro každý algebraický prvek? Ano, neboť množina

$$I = \{f \in T[x] : \ f(a) = 0\}$$

tvoří ideál v oboru $\mathbf{T}[x]$; protože jde o obor hlavních ideálů (Věta 10.2), $\mathbf{I}$ má (monický) generátor m, tj. $I = mT[x]$, čili m je hledaný polynom $m_{a,\mathbf{T}}$.

Polynom $m_{a,\mathbf{T}}$ je v $\mathbf{T}[x]$ ireducibilní: kdyby se rozkládal na součin $f \cdot g$, pak by prvek a byl kořenem f nebo g (nebo obou), což by bylo ve sporu s minimalitou. Naopak, je-li a kořen monického ireducibilního polynomu $h \in T[x]$, pak $h = m_{a,\mathbf{T}}$: to proto, že $m_{a,\mathbf{T}}$ musí dělit h, jenže ten nemá vlastní dělitele.

Příklad. Je ihned vidět, že

$$m_{1,\mathbb{Q}} = x - 1, \quad m_{i,\mathbb{Q}} = x^2 + 1, \quad m_{\sqrt[3]{2},\mathbb{Q}} = x^3 - 2, \quad m_{\sqrt{2}+\sqrt{3},\mathbb{Q}} = x^4 - 10x^2 + 1,$$

neboť jde o ireducibilní polynomy, které mají daný prvek za kořen.

Tvrzení 12.6. *Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles a $a \in S$ algebraický prvek nad $\mathbf{T}$. Pak*

$$\mathbf{T}(a) = \mathbf{T}[a].$$

Důkaz. Podle Tvrzení 12.3 tvoří

$$T[a] = \{f(a) : \ f \in T[x]\}$$

podokruh tělesa $\mathbf{S}$. Dokážeme, že to je podtěleso, tj. že v něm existují inverzní prvky ke všem nenulovým prvkům. Mějme tedy nějaký prvek $0 \neq f(a) \in T[a]$; hledáme polynom $g \in T[x]$ takový, že $f(a)g(a) = 1$. Protože $f(a) \neq 0$, polynom $m_{a,\mathbf{T}}$ nedělí f. Z ireducibility $m_{a,\mathbf{T}}$ plyne $\mathrm{NSD}(m_{a,\mathbf{T}}, f) = 1$, a tak podle Bézoutovy rovnosti existují polynomy $u, g \in T[x]$ takové, že $1 = um_{a,\mathbf{T}} + gf$. Dosazením prvku a dostáváme

$$1 = u(a)m_{a,\mathbf{T}}(a) + g(a)f(a) = u(a) \cdot 0 + g(a)f(a) = f(a)g(a),$$

čili $g(a)$ je inverzní k $f(a)$. □

Příklad. Číslo $\sqrt{s}$ ($s \in \mathbb{Z}$) je algebraické nad $\mathbb{Q}$, tedy $\mathbb{Q}(\sqrt{s}) = \mathbb{Q}[\sqrt{s}]$. Skutečně,

$$(a + b\sqrt{s})^{-1} = \frac{a}{a^2 - b^2 s} - \frac{b}{a^2 - b^2 s}\sqrt{s} \in \mathbb{Q}[\sqrt{s}].$$

Příklad. Číslo π je transcendentní nad $\mathbb{Q}$. Díky tomu má homomorfismus

$$\mathbb{Q}[x] \to \mathbb{Q}[\pi], \qquad f \mapsto f(\pi)$$

triviální jádro a z Tvrzení 12.3 plyne, že to je izomorfismus. Ovšem $\mathbb{Q}[x]$ není těleso, takže ani $\mathbb{Q}[\pi]$ není těleso a z toho důvodu

$$\mathbb{Q}[\pi] \neq \mathbb{Q}(\pi).$$

Všimněte si, že např. $\frac{1}{\pi} \in \mathbb{Q}(\pi) \setminus \mathbb{Q}[\pi]$. Ve skutečnosti je $\mathbb{Q}(\pi)$ podílové těleso oboru $\mathbb{Q}[\pi]$.

Tvrzení 12.7. *Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles a $a \in S$ algebraický prvek nad $\mathbf{T}$. Pak*

$$[\mathbf{T}(a) : \mathbf{T}] = \deg m_{a,\mathbf{T}}.$$

Důkaz. Označme $n = \deg m_{a,\mathbf{T}}$. Dokážeme, že prvky $1, a, a^2, \ldots, a^{n-1}$ tvoří bázi vektorového prostoru $\mathbf{T}(a)_{\mathbf{T}}$, a tedy že jeho dimenze je n.

Kdyby byly prvky $1, a, a^2, \ldots, a^{n-1}$ lineárně závislé, pak by platilo $\sum_{i=0}^{n-1} b_i a^i = 0$ pro nějaká $b_i \in T$, z nichž by aspoň jedno bylo nenulové. Prvek a by tedy byl kořenem (nenulového) polynomu $\sum_{i=0}^{n-1} b_i x^i \in T[x]$ s menším stupněm než $m_{a,\mathbf{T}}$, což by byl spor s minimalitou $m_{a,\mathbf{T}}$.

Nyní dokážeme, že prvky $1, a, \ldots, a^{n-1}$ generují vektorový prostor $\mathbf{T}(a)_{\mathbf{T}}$. Uvažujme prvek $f(a)$ tělesa $\mathbf{T}(a) = \mathbf{T}[a]$, vyjádříme jej jako lineární kombinaci. Buď $q, r \in T[x]$ takové, že $f = q \cdot m_{a,\mathbf{T}} + r$ a $\deg r < \deg m_{a,\mathbf{T}} = n$. Pak

$$f(a) = q(a) \cdot m_{a,\mathbf{T}}(a) + r(a) = q(a) \cdot 0 + r(a) = r(a),$$

a protože je stupeň r menší než n, máme $f(a) = r(a) = \sum_{i=0}^{n-1} b_i a^i$, kde $b_i \in T$ jsou koeficienty polynomu r. □

Příklad. Uvedené tvrzení dává návod, jak určit stupeň rozšíření o jeden prvek.

- $[\mathbb{C} : \mathbb{R}] = [\mathbb{R}(i) : \mathbb{R}] = \deg(x^2 + 1) = 2$.
- $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = \deg(x^3 - 2) = 3$.
  Obecněji, $[\mathbb{Q}(\sqrt[n]{p}) : \mathbb{Q}] = \deg(x^n - p) = n$ pro libovolné n a prvočíslo p, protože uvedený polynom je podle Eisensteinova kritéria ireducibilní. (Pokud p není prvočíslo, situace je složitější.)
- $[\mathbb{Q}(e^{2\pi i/3}) : \mathbb{Q}] = 2$. Číslo $e^{2\pi i/3}$ je kořen polynomu $x^3 - 1$, ten však není ireducibilní, rozkládá se jako $(x-1)(x^2+x+1)$.
  Obecněji, $[\mathbb{Q}(e^{2\pi i/p}) : \mathbb{Q}] = p - 1$ pro libovolné prvočíslo p.

K výpočtu stupně komplikovanějších rozšíření se může hodit následující tvrzení.

Tvrzení 12.8. *Buď $\mathbf{T} \leq \mathbf{S} \leq \mathbf{U}$ rozšíření těles. Pak*

$$[\mathbf{U} : \mathbf{T}] = [\mathbf{U} : \mathbf{S}] \cdot [\mathbf{S} : \mathbf{T}].$$

Abychom zjednodušili zápis, důkaz tvrzení provedeme pouze pro případ, kdy jde o rozšíření konečného stupně. V nekonečném případě lze postupovat analogicky a čtenář zběhlý v práci s prostory nekonečné dimenze si důkaz snadno sám upraví (v dalším textu nebudeme tento případ potřebovat).

Důkaz. Označme $m = [\mathbf{U} : \mathbf{S}]$, $n = [\mathbf{S} : \mathbf{T}]$ a zvolme bázi $a_1, \ldots, a_n$ vektorového prostoru $\mathbf{S}_{\mathbf{T}}$ a bázi $b_1, \ldots, b_m$ vektorového prostoru $\mathbf{U}_{\mathbf{S}}$. Dokážeme, že prvky

$$a_1 b_1, \ldots, a_1 b_m,\ a_2 b_1, \ldots, a_2 b_m,\ \ldots,\ a_n b_1, \ldots, a_n b_m$$

tvoří bázi vektorového prostoru $\mathbf{U}_{\mathbf{T}}$.

Nejprve dokážeme, že tyto prvky generují $\mathbf{U}_{\mathbf{T}}$. Je-li $u \in U$, pak $u = \sum_i s_i b_i$ pro nějaká $s_i \in S$. Každé s_i lze napsat jako $s_i = \sum_j t_{ij} a_j$ pro nějaká $t_{ij} \in T$ a dosazením druhé rovnosti do první dostáváme

$$u = \sum_i \big(\sum_j t_{ij} a_j \big) b_i = \sum_{i,j} t_{ij} \cdot a_j b_i.$$

Tedy u je lineární kombinací uvedených prvků s koeficienty z tělesa $\mathbf{T}$.

Nyní dokážeme lineární nezávislost. Předpokládejme, že $\sum_{i,j} t_{ij} \cdot a_i b_j = 0$ pro nějaká $t_{ij} \in T$. Rozepíšeme

$$0 = \sum_{i,j} t_{ij} a_i b_j = \sum_j \underbrace{\big(\sum_i t_{ij} a_i\big)}_{\in S} b_j.$$

Lineární nezávislost prvků $b_1, \ldots, b_m$ nad tělesem $\mathbf{S}$ nám dává $\sum_i t_{ij} a_i = 0$ pro každé j a z lineární nezávislosti $a_1, \ldots, a_n$ nad tělesem $\mathbf{T}$ dostáváme $t_{ij} = 0$ pro všechna i, j. □

Příklad. Pomocí výpočtu dimenze předvedeme, že

$$\mathbb{Q}(\sqrt{2} + \sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3}).$$

Zřejmě $\mathbb{Q}(\sqrt{2} + \sqrt{3}) \leq \mathbb{Q}(\sqrt{2}, \sqrt{3})$. Pokud tedy dokážeme, že oba prostory mají stejnou dimenzi, musí být totožné. Spočteme minimální polynomy:

- $m_{\sqrt{2}+\sqrt{3},\mathbb{Q}} = x^4 - 10x^2 + 1$;
- $m_{\sqrt{2},\mathbb{Q}} = x^2 - 2$;
- $m_{\sqrt{3},\mathbb{Q}(\sqrt{2})} = x^2 - 3$ — uvědomte si, že skutečně jde o ireducibilní polynom v $\mathbb{Q}(\sqrt{2})[x]$.

Podle Tvrzení 12.7 a 12.8 dostáváme $[\mathbb{Q}(\sqrt{2} + \sqrt{3}) : \mathbb{Q}] = 4$ a $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] \cdot [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \cdot 2 = 4$.

12.4. **Rozšíření konečného stupně.**

[POZNÁMKA: možná by se sem hodila věta o primitivním prvku, pokud najdeme jednoduchý důkaz.]

Věta 12.9. *Rozšíření $\mathbf{T} \leq \mathbf{S}$ má konečný stupeň právě tehdy, když*

$$\mathbf{S} = \mathbf{T}(a_1, \ldots, a_n)$$

pro nějaké prvky $a_1, \ldots, a_n \in S$ algebraické nad $\mathbf{T}$.

Důkaz. ($\Leftarrow$) Uvažujme postupná rozšíření

$$\mathbf{T} \leq \mathbf{T}(a_1) \leq \mathbf{T}(a_1, a_2) \leq \ldots \leq \mathbf{T}(a_1, \ldots, a_n).$$

Podle Tvrzení 12.8 je $[\mathbf{T}(a_1, \ldots, a_n) : \mathbf{T}]$ rovno

$$[\mathbf{T}(a_1) : \mathbf{T}] \cdot [\mathbf{T}(a_1, a_2) : \mathbf{T}(a_1)] \cdot \ldots \cdot [\mathbf{T}(a_1, \ldots, a_n) : \mathbf{T}(a_1, \ldots, a_{n-1})].$$

Všechny stupně v součinu jsou konečné díky Tvrzení 12.7, tedy i $[\mathbf{T}(a_1, \ldots, a_n) : \mathbf{T}]$ je konečný.

($\Rightarrow$) Budeme postupovat indukcí podle $k = [\mathbf{S} : \mathbf{T}]$. Pro $k = 1$ je $\mathbf{S} = \mathbf{T}$ a věta platí. Dále předpokládejme platnost tvrzení pro všechna rozšíření dimenze méně než k. Zvolme prvek $a \in S \setminus T$ a uvažujme rozšíření $\mathbf{T} < \mathbf{T}(a) \leq \mathbf{S}$. Podle Tvrzení 12.8 platí

$$\underbrace{[\mathbf{S} : \mathbf{T}]}_{k} = \underbrace{[\mathbf{S} : \mathbf{T}(a)]}_{<k} \cdot \underbrace{[\mathbf{T}(a) : \mathbf{T}]}_{>1}.$$

Z indukčního předpokladu dostáváme, že

$$\mathbf{S} = (\mathbf{T}(a))(b_1, \ldots, b_n) = \mathbf{T}(a, b_1, \ldots, b_n)$$

pro nějaké prvky $b_1, \ldots, b_n$. Protože jde o rozšíření konečného stupně, všechny prvky $a, b_1, \ldots, b_n$ jsou podle Tvrzení 12.5 algebraické nad $\mathbf{T}$. □

Na závěr uvedeme dvě drobné aplikace této věty. Za prvé, algebraická čísla jsou uzavřená na tělesové operace, tj. tvoří podtěleso.

Věta 12.10. *Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles. Pak prvky $\mathbf{S}$, které jsou algebraické nad $\mathbf{T}$, tvoří podtěleso tělesa $\mathbf{S}$.*

Důkaz. Uvažujme prvky $a, b \in S$ algebraické nad $\mathbf{T}$ a uvažujme rozšíření $\mathbf{T} \leq \mathbf{T}(a,b)$. Podle Věty 12.9 jde o rozšíření konečného stupně, a tudíž o rozšíření algebraické, díky Tvrzení 12.5. Čili všechny prvky $\mathbf{T}(a,b)$ jsou algebraické nad $\mathbf{T}$, speciálně také prvky $a+b$, $a\cdot b$, $-a$ i a^{-1} (pro $a \neq 0$). Tedy algebraické prvky tvoří podtěleso tělesa $\mathbf{S}$. □

Za druhé, ukážeme si strukturu tzv. *kvadratických rozšíření*, tj. rozšíření stupně 2. Dokážeme, že je-li $\mathbf{T} < \mathbf{S} \leq \mathbb{C}$ a $[\mathbf{S} : \mathbf{T}] = 2$, pak

$$\mathbf{S} = \mathbf{T}(\sqrt{s}) \text{ pro nějaké } s \in T.$$

Podle Věty 12.9 je $\mathbf{S} = \mathbf{T}(a)$ a podle Tvrzení 12.7 je a kořenem nějakého polynomu z $\mathbf{T}[x]$ stupně 2. Známý vzorec na výpočet kořenů kvadratického polynomu říká, že $a = u + v\sqrt{s}$ pro nějaká $u, v, s \in T$, a tak $\mathbf{S} = \mathbf{T}(u+v\sqrt{s}) = \mathbf{T}(\sqrt{s})$. (Tvrzení platí obecněji, pro libovolné kvadratické rozšíření těles charakteristiky $\neq 2$. Místo komplexních čísel stačí uvažovat libovolné nadtěleso, kde existují odmocniny, jako např. algebraický uzávěr.)

Rozšíření vyšších stupňů už tak snadno popsat nejdou. Z Cardanových vzorců např. plyne, že $[\mathbb{Q}(\sqrt[3]{-1+\sqrt{2}}) : \mathbb{Q}] = 3$, neboť minimální polynom tohoto čísla je $x^3 + 3x + 2$. Oproti tomu jistě není pravda, že každé rozšíření typu $\mathbb{Q}(\sqrt[3]{a+\sqrt{b}})$ je kubické, např. pro $a = 0$ a $b = 2$.

Shrnutí. Zapamatujte si následující vlastnosti rozšíření $\mathbf{T} \leq \mathbf{S}$:

(1) Je-li $a \in S$ algebraický nad $\mathbf{T}$, pak
$$\mathbf{T}(a) = \mathbf{T}[a] = \{f(a) : \ f \in T[x]\} \quad \text{a} \quad [\mathbf{T}(a) : \mathbf{T}] = \deg m_{a,\mathbf{T}}.$$
(2) Je-li $[\mathbf{S} : \mathbf{T}] < \infty$, pak každý prvek $\mathbf{S}$ je algebraický nad $\mathbf{T}$ a
$$\mathbf{S} = \mathbf{T}(a_1, \ldots, a_n) = \mathbf{T}[a_1, \ldots, a_n]$$
pro jistá $a_1, \ldots, a_n \in T$.
(3) Je-li $\mathbf{T} \leq \mathbf{S} \leq \mathbf{U}$, pak $[\mathbf{U} : \mathbf{T}] = [\mathbf{U} : \mathbf{S}] \cdot [\mathbf{S} : \mathbf{T}]$.

13. Aplikace: konstrukce pravítkem a kružítkem

> **Cíl.** *Geometrickým konstrukcím pravítkem a kružítkem odpovídají tělesa konstruovatelných čísel. Pomocí poznatků z předchozí sekce lze dokázat, že některá čísla konstruovatelná nejsou, což umožňuje dokázat neřešitelnost některých konstrukčních úloh.*

Mezi klasické starořecké úlohy patřily konstrukce pomocí pravítka a kružítka. Postupem času vykrystalizovala čtyři slavná zadání, která se přes velkou snahu nedařilo vyřešit:

- *Rektifikace kružnice*: k dané kružnici sestrojit úsečku, která je stejně dlouhá jako obvod této kružnice.
- *Kvadratura kruhu*: k danému kruhu sestrojit úsečku takovou, že čtverec s touto hranou má stejnou plochu jako daný kruh.

- *Zdvojení krychle*: k dané úsečce u sestrojit úsečku v takovou, že krychle s hranou dlouhou jako v má dvakrát větší objem, než krychle s hranou dlouhou jako u.
- *Trisekce úhlu*: k danému úhlu sestrojit třetinový úhel.

V moderní řeči bychom první tři úlohy přeložili jako „je-li dána jednotková úsečka, zkonstruujte úsečku délky 2π, resp. $\sqrt{\pi}$, resp. $\sqrt[3]{2}$.“ Přes 2000 let trvaly snahy tyto úlohy vyřešit. Až rozvoj algebry v 19. století umožnil dokázat, že to není možné. Pro zdvojení krychle a trisekci úhlu nalezl důkaz Pierre Wantzel roku 1837. Stejná metoda řeší i rektifikaci kružnice a kvadraturu kruhu, k dokončení však bylo třeba počkat dalších téměř 50 let na Lindemannův důkaz transcendence čísla π. Wantzelovu metodu zde předvedeme.

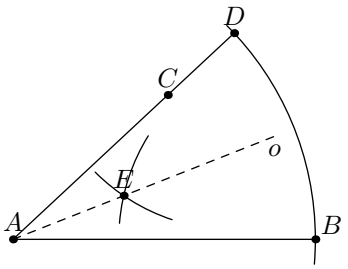
Předně musíme upřesnit, co vlastně rozumíme konstrukcí pomocí pravítka a kružítka. Na začátku je daná jistá konečná množina $\mathcal{M}_0$ bodů v rovině. Z ní můžeme zkonstruovat nový bod jako průsečík přímek nebo kružnic určených již zkonstruovanými body; a tento postup lze několikrát opakovat. Formálně, konstrukce pomocí pravítka a kružítka je posloupnost $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \ldots \subseteq \mathcal{M}_n$ konečných množin bodů v rovině taková, že $\mathcal{M}_{i+1} = \mathcal{M}_i \cup \{X\}$, kde X vznikne jako

(1) průsečík přímky AB a přímky CD;
(2) průsečík přímky AB a kružnice se středem C a poloměrem $|DE|$;
(3) průsečík kružnice se středem A a poloměrem $|BC|$ a kružnice se středem D a poloměrem $|EF|$

pro nějaké body $A, B, C, D, E, F \in \mathcal{M}_i$.

Princip Wantzelovy metody je převedení konstrukcí pravítkem a kružítkem do jazyka algebry. Zvolme v rovině souřadnice a uvažujme nejmenší těleso $\mathbf{T}_i$, které obsahuje x-ové i y-ové souřadnice všech bodů z $\mathcal{M}_i$. Dostáváme tak řetězec rozšíření těles $\mathbf{T}_0 \leq \mathbf{T}_1 \leq \mathbf{T}_2 \leq \ldots \leq \mathbf{T}_n$.

Příklad (Půlení úhlu)**.** Podívejme se, jak se formalizuje úloha k danému úhlu sestrojit poloviční úhel. Mějme dán úhel třemi body A, B, C (kde A je vrchol).



Sestrojíme body

$$D = k(A, |AB|) \cap AC \quad \text{a} \quad E = k(B, |BD|) \cap k(D, |BD|),$$

výsledkem bude úhel daný body A, B, E. Tedy

$$\mathcal{M}_0 = \{A, B, C\}, \quad \mathcal{M}_1 = \mathcal{M}_0 \cup \{D\}, \quad \mathcal{M}_2 = \mathcal{M}_1 \cup \{E\}.$$

Zvolme souřadnice tak, že $A = (0, 0)$, $B = (1, 0)$ a $C = (a, b)$. Není těžké spočítat, že $D = (\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}})$ a $E = (\frac{1}{2} + \frac{a-b\sqrt{3}}{2\sqrt{a^2+b^2}}, \frac{\sqrt{3}}{2} + \frac{b+a\sqrt{3}}{2\sqrt{a^2+b^2}})$, tedy

$$\mathbf{T}_0 = \mathbb{Q}(a, b), \quad \mathbf{T}_1 = \mathbf{T}_0(\sqrt{a^2+b^2}), \quad \mathbf{T}_2 = \mathbf{T}_0(\sqrt{a^2+b^2}, \sqrt{3}).$$

Stěžejním krokem Wantzelovy metody je následující vlastnost.

Tvrzení 13.1. $[\mathbf{T}_n : \mathbf{T}_0]$ *je mocnina čísla 2.*

Důkaz. Podle Tvrzení 12.8 je

$$[\mathbf{T}_n : \mathbf{T}_0] = [\mathbf{T}_n : \mathbf{T}_{n-1}] \cdot \ldots \cdot [\mathbf{T}_2 : \mathbf{T}_1] \cdot [\mathbf{T}_1 : \mathbf{T}_0].$$

Ukážeme, že

$$[\mathbf{T}_{i+1} : \mathbf{T}_i] \in \{1, 2\}.$$

Probereme postupně všechny tři možnosti, jak se konstruuje nový bod.

(1) Jde-li o průsečík dvou přímek, získáme souřadnice nového bodu řešením soustavy dvou lineárních rovnic o dvou neznámých nad tělesem $\mathbf{T}_i$. Řešením soustavy je opět prvek tělesa $\mathbf{T}_i$ (viz algoritmus Gaussovy eliminace), takže máme $\mathbf{T}_{i+1} = \mathbf{T}_i$ a

$$[\mathbf{T}_{i+1} : \mathbf{T}_i] = 1.$$

(2) Jde-li o průsečík přímky a kružnice, získáme souřadnice nového bodu řešením soustavy jedné lineární a jedné kvadratické rovnice o dvou neznámých nad tělesem $\mathbf{T}_i$. Vyjádříme-li z lineární rovnice y a dosadíme jej do kvadratické, dostaneme kvadratickou rovnici pro x, jejímž řešením je číslo tvaru $a + b\sqrt{s}$, $a, b, s \in T_i$; podobný tvar bude mít i y. Tedy $\mathbf{T}_{i+1} = \mathbf{T}_i(\sqrt{s})$, z čehož plyne, že

$$[\mathbf{T}_{i+1} : \mathbf{T}_i] \in \{1, 2\}$$

v závislosti na tom, zda je $\sqrt{s} \in T_i$ nebo ne.

(3) Jde-li o průsečík dvou kružnic, získáme souřadnice nového bodu řešením soustavy dvou kvadratických rovnic o dvou neznámých nad tělesem $\mathbf{T}_i$. Odečtením rovnic od sebe se zbavíme se kvadratických členů (všechny mají koeficient 1) a získáme tak ekvivalentní soustavu sestávající z jedné lineární a jedné kvadratické rovnice. Stejným argumentem jako v (2) dostaneme

$$[\mathbf{T}_{i+1} : \mathbf{T}_i] \in \{1, 2\}.$$

(Proveďte popsané výpočty podrobně s obecnými rovnicemi přímky a kružnice v rovině!) □

Důsledkem Tvrzení 13.1 je neřešitelnost uvedených úloh pravítkem a kružítkem.

Rektifikace kružnice a *kvadratura kruhu.* Zvolme souřadnice tak, že krajní body zadané úsečky (udávající střed a poloměr kružnice) jsou $(0, 0)$ a $(1, 0)$; čili $\mathbf{T}_0 = \mathbb{Q}$. Cílem úlohy je sestrojit úsečku délky 2π, resp. $\sqrt{\pi}$, a bez újmy na obecnosti můžeme předpokládat, že výsledná úsečka má krajní body $(0, 0)$ a $(2\pi, 0)$, resp. $(\sqrt{\pi}, 0)$. V tom případě ale π, resp. $\sqrt{\pi}$, náleží tělesu $\mathbf{T}_n$ a to je spor, neboť rozšíření $\mathbf{T}_0 \leq \mathbf{T}_n$ má být konečného stupně, tedy podle Tvrzení 12.5 algebraické, zatímco π i $\sqrt{\pi}$ jsou transcendentní čísla.

(Obecněji bychom mohli říci, že z jednotkové úsečky nelze sestrojit úsečku žádné transcendentní délky.)

Zdvojení krychle. Podobně, zvolme souřadnice tak, že krajní body zadané úsečky jsou $(0, 0)$ a $(1, 0)$; čili $\mathbf{T}_0 = \mathbb{Q}$. Cílem úlohy je sestrojit úsečku délky $\sqrt[3]{2}$ a bez újmy na obecnosti můžeme předpokládat, že výsledná úsečka má krajní body $(0, 0)$ a $(\sqrt[3]{2}, 0)$. V tom případě ale $\sqrt[3]{2}$ náleží tělesu $\mathbf{T}_n$, z čehož plyne, že 3 dělí $[\mathbf{T}_n : \mathbf{T}_0]$ – uvažujeme-li rozšíření $\mathbb{Q} \leq \mathbb{Q}(\sqrt[3]{2}) \leq \mathbf{T}_n$, Tvrzení 12.8 říká, že

$$[\mathbf{T}_n : \mathbf{T}_0] = [\mathbf{T}_n : \mathbb{Q}(\sqrt[3]{2})] \cdot [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3 \cdot [\mathbf{T}_n : \mathbb{Q}(\sqrt[3]{2})].$$

Spor s Tvrzením 13.1.

(Obecněji bychom mohli říci, že z jednotkové úsečky nelze sestrojit úsečka žádné délky a takové, že polynom $m_{a,\mathbb{Q}}$ má stupeň, který není mocnina dvojky.)

Trisekce úhlu. Stačí najít jedno konkrétní zadání, které není řešitelné pravítkem a kružítkem. Uvažujme tedy úhel $60°$ zadaný body $(0,0)$, $(1,0)$ a $(\frac{1}{2}, \frac{\sqrt{3}}{2})$; čili $\mathbf{T}_0 = \mathbb{Q}(\sqrt{3})$. Dokážeme, že není možné sestrojit bod

$$(\cos 20°, \sin 20°).$$

(Kdybychom zkonstruovali přímku se směrnicí $20°$ pomocí jiného bodu, dostaneme tento jako její průsečík s jednotkovou kružnicí.) Dokážeme-li, že

$$[\mathbb{Q}(\sqrt{3}, \cos 20°) : \mathbb{Q}(\sqrt{3})] = 3,$$

můžeme použít stejný argument jako pro zdvojení krychle. K tomuto cíli stačí podle Tvrzení 12.7 nalézt minimální polynom čísla $\cos 20°$ nad tělesem $\mathbb{Q}(\sqrt{3})$, tj. nějaký ireducibilní polynom, jehož je číslo $\cos 20°$ kořenem. Použijeme-li vzorec

$$\cos 3\alpha = 4(\cos\alpha)^3 - 3\cos\alpha$$

(viz nějaká sbírka goniometrických vzorců), dostáváme $\cos 20°$ jako kořen polynomu $4x^3 - 3x - \frac{1}{2} \in \mathbb{Q}(\sqrt{3})[x]$. Tento polynom je v $\mathbb{Q}(\sqrt{3})[x]$ ireducibilní, neboť nemá v $\mathbb{Q}(\sqrt{3})$ kořen (jak snadno zjistíme dosazením $x = a + b\sqrt{3}$). Tedy

$$m_{\cos 20°, \mathbb{Q}(\sqrt{3})} = x^3 - \frac{3}{4}x - \frac{1}{8}$$

a dostáváme $[\mathbb{Q}(\sqrt{3}, \cos 20°) : \mathbb{Q}(\sqrt{3})] = \deg m_{\cos 20°, \mathbb{Q}(\sqrt{3})} = 3$.

Rejstřík