UČEBNÍ TEXT
ALGEBRA 2020/21

DAVID STANOVSKÝ

stanovsk@karlin.mff.cuni.cz

Date: 16. února 2021.

Základní algebraické objekty

1. Obory integrity

1.1. **Definice a příklady.**

Ze základní a střední školy znáte základní koncepty týkající se dělitelnosti přirozených čísel: algoritmus dělení se zbytkem, prvočísla, prvočíselné rozklady, pojem největšího společného dělitele (NSD) a nejmenšího společného násobku. Nicméně přirozená čísla nejsou jediným oborem, pro který má smysl tyto pojmy studovat: například algoritmus dělení se zbytkem možná znáte i pro polynomy a ostatní uvedené koncepty lze snadno zobecnit. Dělitelnost lze studovat v mnohem širším kontextu tak zvaných *oborů integrity*, mezi něž patří obor celých čísel, obory polynomů, ale také různá rozšíření celých čísel, např. *Gaussova celá čísla* (komplexní čísla s celočíselnými koeficienty).

V různých oborech pak platí různě silná tvrzení. Například existence a jednoznačnost prvočíselných rozkladů platí nejen pro celá čísla, ale také pro celočíselné polynomy či pro Gaussova celá čísla; přesto existují číselné obory, kde jednoznačnost neplatí. Eukleidův algoritmus na výpočet NSD založený na dělení se zbytkem funguje v oboru celých čísel, pro polynomy jedné proměnné nad tělesy, ale například pro polynomy více proměnných nastává s dělením problém. V kapitole II si uděláme v uvedených vlastnostech pořádek. Nyní, v kapitole úvodní, si přesně definujeme matematické struktury, se kterými budeme dále pracovat.

Abychom mohli studovat všechny zmíněné obory naráz, zavádí se obecná struktura nazývaná *komutativní okruh*, jejíž axiomy vystihují základní aritmetické vlastnosti sčítání a násobení.

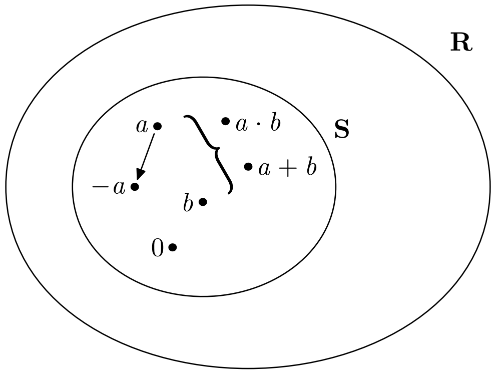
Definice. *Okruhem* $\mathbf{R}$ rozumíme pětici $(R, +, -, \cdot, 0)$, kde R je neprázdná množina, na které jsou definovány binární operace $+, \cdot$ (tj. zobrazení $R \times R \to R$), unární operace $-$ (tj. zobrazení $R \to R$) a prvek $0 \in R$ splňující pro každé $a, b, c \in R$ následující podmínky:

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

Okruh nazveme *komutativní*, pokud je komutativní také operace násobení, tj. $a \cdot b = b \cdot a$ pro všechna $a, b \in R$. *Okruhem s jednotkou* pak rozumíme okruh, ve kterém existuje prvek $1 \in R$ splňující $a \cdot 1 = a$ pro každé $a \in R$.

- Platí-li navíc podmínka $0 \neq 1$ a
$$\text{pokud } a, b \neq 0, \text{ pak } a \cdot b \neq 0,$$
nazýváme $\mathbf{R}$ *obor integrity.*
- Platí-li navíc podmínka $0 \neq 1$ a
$$\text{pro každé } a \neq 0 \text{ existuje } b \text{ splňující } a \cdot b = 1,$$
nazýváme $\mathbf{R}$ *těleso.* Z Tvrzení 1.2(5) plyne, že takové b je 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.



Obrázek 1. Podokruh **S** okruhu **R**.

Příklad. Základní číselné obory $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ jsou komutativní okruhy vzhledem k standardním operacím sčítání, odčítání a násobení. Okruh $\mathbb{Z}$ je oborem integrity (ale nikoliv tělesem), okruhy $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ jsou tělesa. Další zajímavé příklady oborů integrity a těles se nacházejí mezi obory $\mathbb{Z}$ a $\mathbb{C}$, resp. mezi tělesy $\mathbb{Q}$ a $\mathbb{C}$, viz níže.

Příklad. Důležitými příklady jsou konečné komutativní okruhy

$$\mathbb{Z}_n = (\{0, \dots, n-1\}, +_{\text{ mod } n}, -_{\text{ mod } n}, \cdot_{\text{ mod } n}, 0)$$

s operacemi modulo n. Všimněte si, že

$$\mathbb{Z}_n \text{ je těleso } \Leftrightarrow \mathbb{Z}_n \text{ je obor integrity } \Leftrightarrow n \text{ je prvočíslo.}$$

První implikace plyne z Tvrzení 1.3, každé těleso je oborem integrity. Druhá implikace: pokud by $n = k \cdot l$ bylo složené číslo, $k, l > 1$, pak by v $\mathbb{Z}_n$ platilo $k \cdot l = n \bmod n = 0$, čili by to nebyl obor integrity. A je-li n prvočíslo, pak je $a^{n-2} \bmod n$ inverzním prvkem pro $a \neq 0$, což plyne ihned z malé Fermatovy věty 4.11 (alternativně, inverzní prvek je koeficientem u z Bézoutovy rovnosti $1 = ua + vp$, viz Tvrzení 4.4).

Příklad. Konečných komutativních okruhů je spousta, ale konečných těles mnoho není. Věta 26.3 říká, že všechna konečná tělesa mají velikost mocniny prvočísla a pro každou mocninu prvočísla p^k existuje právě jedno konečné těleso s p^k prvky (právě jedno až na izomorfismus). Konstrukce bude popsána v sekci 9.2.

Příklad. Stěžejní roli hrají v algebře obory polynomů $\mathbf{R}[x]$, tj. formálních výrazů tvaru $a_0 + a_1x + \ldots + a_nx^n$, kde koeficienty bereme z jistého komutativního okruhu $\mathbf{R}$. Formální definici se dozvíte v sekci 2.

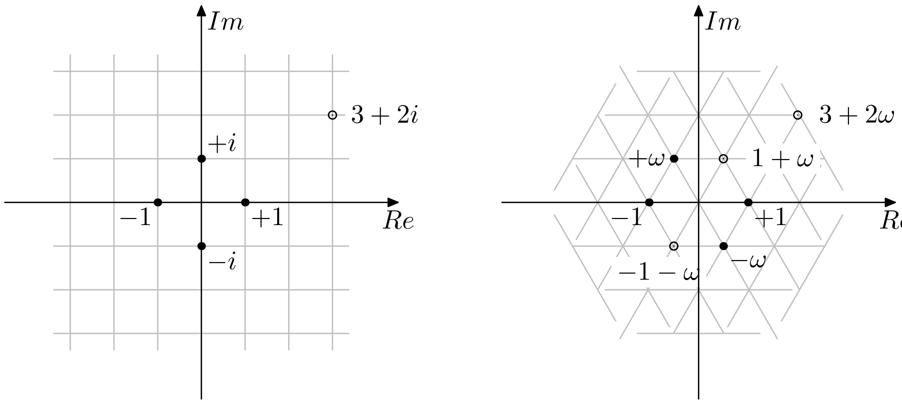
V přednášce se budeme zabývat pouze komutativními okruhy. Přesto, některá tvrzení předpoklad komutativity nevyžadují a proto je dobré mít na paměti nějaký příklad.

Příklad. Základním příkladem nekomutativních okruhů jsou okruhy matic $\mathbf{M}_n(\mathbf{T})$ sestávající z čtvercových matic $n \times n$ nad tělesem $\mathbf{T}$ se standardními operacemi maticového sčítání a násobení.

K popisu oborů a těles mezi $\mathbb{Z}$ a $\mathbb{C}$ se hodí pojem podokruhu a podtělesa.

Definice. Buď **R** komutativní okruh a $S \subseteq R$ podmnožina taková, že $0 \in S$ a kdykoliv $a, b \in S$, pak také $-a \in S$, $a + b \in S$ a $a \cdot b \in S$ (říkáme, že množina S je *uzavřená* na uvedené operace). Vezmeme-li na množině S restrikce operací okruhu **R**, dostaneme také komutativní okruh, který označíme **S** (jsou-li všechny axiomy splněny na větší množině R, pak jistě i na její podmnožině S). Výsledky této konstrukce nazýváme *podokruhy* komutativního okruhu **R**, značíme $\mathbf{S} \leq \mathbf{R}$.

Je-li **R** obor integrity a platí-li navíc $1 \in S$, hovoříme o podoboru. Je-li **R** těleso a platí-li navíc podmínka $0 \neq a \in S \Rightarrow a^{-1} \in S$, hovoříme o *podtělesu.*



Obrázek 2. Gaussova a Eisensteinova celá čísla.

Příklad. Obor $\mathbb{Z}$ je podoborem tělesa $\mathbb{Q}$, které je podtělesem tělesa $\mathbb{R}$, které je podtělesem tělesa $\mathbb{C}$.

V algebraické teorii čísel se studují podobory a podtělesa tělesa $\mathbb{C}$, která vznikají „přidáním konečně mnoha prvků“ k oboru $\mathbb{Z}$, resp. tělesu $\mathbb{Q}$ (podrobněji viz sekce 3).

Příklad (Gaussova čísla)**.**

- Množina $\mathbb{Z}[i] = \{a + bi : \ a, b \in \mathbb{Z}\}$ tvoří podobor tělesa $\mathbb{C}$, zvaný *Gaussova celá čísla.* Uzavřenost na odčítání plyne ze vztahu $-(a + bi) = -a - bi$, na sčítání ze vztahu $(a + bi) + (c + di) = (a + c) + (b + d)i$ a na násobení ze vztahu $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$.
- Množina $\mathbb{Q}(i) = \{a + bi : \ a, b \in \mathbb{Q}\}$ tvoří podtěleso tělesa $\mathbb{C}$, zvané *Gaussova racionální čísla.* Je třeba navíc ověřit, že $(a+bi)^{-1} \in \mathbb{Q}(i)$, což plyne ze vztahu $\frac{1}{a+bi} = \frac{1}{a+bi} \cdot \frac{a-bi}{a-bi} = \frac{a}{a^2+b^2} + \frac{b}{a^2+b^2} i$.

Příklad (Eisensteinova čísla)**.**
Množina $\mathbb{Z}[\zeta_3] = \{a+b\zeta_3 : \ a, b \in \mathbb{Z}\}$, kde $\zeta_3 = e^{2\pi i/3}$ je komplexní třetí odmocnina z jedné, tvoří podokruh oboru $\mathbb{C}$, zvaný *Eisensteinova celá čísla.* Všimněte si, že $\zeta_3^2 = -1 - \zeta_3$, čili uzavřenost na násobení plyne ze vztahu $(a+b\zeta_3)(c+d\zeta_3) = ac+(ad+bc)\zeta_3+bd\zeta_3^2 = (ac-bd)+(ad+bc-bd)\zeta_3$.

1.2. Základní vlastnosti.

V moderní matematice je zvykem definovat abstraktní struktury pomocí co nejmenší sady axiomů. V této části si odvodíme několik jednoduchých vlastností, které plynou z axiomů komutativních okruhů a v dalším textu je budeme zcela automaticky používat.

Tvrzení 1.1. *Buď $*$ asociativní operace na množině X a $a_1, \dots, a_n \in X$. Hodnota výrazu $a_1 * \ldots * a_n$ nezávisí na uzávorkování.*

Důkaz. Tvrzení dokážeme indukcí. Pro $n = 1, 2$ není co dokazovat, případ $n = 3$ je sám asociativní zákon. Dále budeme postupovat indukcí. Předpokládejme, že na uzávorkování nezávisí výrazy s méně než n členy a uvažujme dva různě uzávorkované výrazy, $(a_1 * \ldots * a_i) * (a_{i+1} * \ldots * a_n)$ a $(a_1 * \ldots * a_j) * (a_{j+1} * \ldots * a_n)$, kde $i < j$. Pak

$$\begin{aligned}(a_1 * \ldots * a_i) * (a_{i+1} * \ldots * a_n) &= (a_1 * \ldots * a_i) * ((a_{i+1} * \ldots * a_j) * (a_{j+1} * \ldots * a_n)) = \\ &= ((a_1 * \ldots * a_i) * (a_{i+1} * \ldots * a_j)) * (a_{j+1} * \ldots * a_n) = \\ &= (a_1 * \ldots * a_j) * (a_{j+1} * \ldots * a_n),\end{aligned}$$

přičemž v první a třetí rovnosti jsme použili indukční předpoklad a v druhé asociativitu pro tři prvky. □

Tvrzení 1.2 (základní vlastnosti okruhů)**.** *Buď* **R** *okruh, $a, b, c \in R$. 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 $a = b$. □

Z těchto vlastností plynou dva zajímavé důsledky.

Tvrzení 1.3. *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. V posledním kroku jsme použili Tvrzení 1.2(2)! □

V konečném světě platí i opačné tvrzení: konečné obory integrity jsou nutně tělesa. Tento fakt je založen na následujícím pozorování.

Lemma 1.4. *Buď* $f : X \to Y$ *zobrazení mezi stejně velkými konečnými množinami. Je-li* f *prosté, pak je bijektivní.*

Důkaz. Nechť $n = |X| = |Y|$. Hodnoty, které zobrazení f přiřadí prvkům množiny X, jsou po dvou různé, takže obor hodnot zobrazení f má právě n prvků. Čili to musí být celé Y. □

Tvrzení 1.5. *Každý konečný obor integrity je tělesem.*

Důkaz. Označme tento obor $\mathbf{R}$. Pro nenulové $a \in R$ uvažujme zobrazení

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

Podle Tvrzení 1.2(5) je toto zobrazení prosté, a protože jde o zobrazení na konečné množině, podle Lemmatu 1.4 je to bijekce. Inverzním prvkem k prvku a je tedy $f_a^{-1}(1)$. □

Buď $\mathbf{R}$ 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. Prvek $\underbrace{1 + 1 + \ldots + 1}_{n}$ budeme zotožnovat s číslem n a prvek $\underbrace{(-1) + (-1) + \ldots + (-1)}_{n}$ s číslem $-n$.

Tvrzení 1.6. *Obor integrity má charakteristiku 0 nebo prvočíslo.*

Důkaz. Kdyby byla charakteristika $n = k \cdot l$, $k, l \neq 1$, pak by v tomto oboru platilo

$$0 = \underbrace{1 + 1 + \ldots + 1}_{n} = (\underbrace{1 + 1 + \ldots + 1}_{k}) \cdot (\underbrace{1 + 1 + \ldots + 1}_{l}).$$

Aspoň jeden z těchto činitelů by musel být roven 0, což je spor s minimalitou n. □

1.3. **Izomorfismus.**

Slovem *homomorfismus* se v matematice označují zobrazení, která zachovávají základní strukturu matematických objektů. Například v lineární algebře to jsou zobrazení zachovávající sčítání a skalární násobení. V teorii grafů to jsou zobrazení zachovávající hrany. Podobně, pro okruhy to budou zobrazení zachovávající základní okruhové operace.

Definice. Buď $\mathbf{R}, \mathbf{S}$ dva okruhy. Zobrazení $\varphi : R \to S$ se nazývá *homomorfismem* těchto okruhů, pokud pro každé $a, b \in R$ platí

$$\varphi(a+b) = \varphi(a) + \varphi(b) \quad \text{a} \quad \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b).$$

Fakt, že je zobrazení mezi okruhy homomorfismem, budeme zapisovat $\varphi : \mathbf{R} \to \mathbf{S}$.

Vlastnostem obecných homomorfismů se budeme věnovat později (sekce 20.1), nyní se soustředíme na jeden speciální případ: bijektivní homomorfismy se nazývají *izomorfismy.* Na izomorfismus je možné pohlížet jako na „kopírování“: máme-li okruh $\mathbf{R}$ a bijektivní zobrazení $\varphi : R \to S$, můžeme na množinu S „překopírovat“ obě operace $* \in \{+, \cdot\}$ předpisem

$$a * b = \varphi(\varphi^{-1}(a) \cdot \varphi^{-1}(b)).$$

Vidíme, že zobrazení φ^{-1} bude izomorfismem mezi novým okruhem $\mathbf{S}$ a starým okruhem $\mathbf{R}$. Jeden okruh je kopií druhého, došlo pouze k „přejmenování prvků“ kopírovacím zobrazením φ. Na každý izomorfismus lze pohlížet tímto způsobem.

Dva okruhy $\mathbf{R}, \mathbf{S}$ nazveme *izomorfní*, pokud existuje izomorfismus $\mathbf{R} \to \mathbf{S}$, tento fakt značíme $\mathbf{R} \simeq \mathbf{S}$. Neformálně, jeden okruh je „kopií“ druhého, jsou stejné „až na přejmenování prvků“ zobrazením φ.

Příklad. Je snadné nahlédnout, že množina matic

$$S = \left\{ \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) : \ a, b \in \mathbb{R} \right\}$$

tvoří podokruh okruhu $\mathbf{M}_2(\mathbb{R})$. Zobrazení

$$\varphi : \mathbb{C} \to \mathbf{S}, \qquad a + bi \mapsto \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$$

je izomorfismem těchto okruhů, neboť je bijektivní a pro všechna $a, b, c, d \in \mathbb{R}$ platí

$$\begin{aligned} \varphi((a+bi) + (c+di)) &= \varphi((a+c) + (b+d)i) = \left(\begin{smallmatrix} a+c & b+d \\ -b-d & a+c \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) + \left(\begin{smallmatrix} c & d \\ -d & c \end{smallmatrix}\right) = \varphi(a+bi) + \varphi(c+di) \end{aligned}$$

a podobně

$$\begin{aligned} \varphi((a+bi) \cdot (c+di)) &= \varphi((ac-bd) + (ad+bc)i) = \left(\begin{smallmatrix} ac-bd & ad+bc \\ -ad-bc & ac-bd \end{smallmatrix}\right) \\ &= \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} c & d \\ -d & c \end{smallmatrix}\right) = \varphi(a+bi) + \varphi(c+di). \end{aligned}$$

Tedy těleso $\mathbb{C}$ je izomorfní s maticovým okruhem $\mathbf{S}$, oba okruhy jsou „stejné“ při ztotožnění čísla $a + bi$ a odpovídající matice.

Příklad. Buď $\mathbf{R}$ libovolný okruh charakteristiky $n > 0$. Je snadné nahlédnout, že prvky tvaru $1 + \dots + 1$ tvoří podokruh $\mathbf{P}$, tzv. *prvookruh.* Je snadné nahlédnout, že zobrazení

$$\varphi : \mathbb{Z}_n \to \mathbf{P}, \qquad 0 \mapsto 0, \quad k \mapsto \underbrace{1 + \dots + 1}_{k},$$

je izomorfismem těchto okruhů. Pro charakteristiku 0 platí analogické tvrzení, prvookruh je izomorfní okruhu $\mathbb{Z}$, uvažovat musíme i prvky tvaru $-(1 + \dots + 1)$.

1.4. **Podílová tělesa.**

Obor celých čísel lze přirozeně rozšířit do tělesa racionálních čísel způsobem, který je znám jako zlomky. Tuto myšlenku lze zobecnit na libovolný obor integrity, výsledkem bude tzv. *podílové těleso* tohoto oboru. Podílová tělesa se hodí v situacích, kdy dělení zjednoduší zadanou úlohu. Například, abychom k výpočtu NSD dvou polynomů mohli použít Eukleidův algoritmus.

Buď $\mathbf{R}$ obor integrity a $M \subseteq R$ tzv. *multiplikativní množina*, tj. podmnožina neobsahující 0, obsahující 1 a splňující podmínku, že kdykoliv $a, b \in M$, pak $a \cdot b \in M$. V konstrukci budeme uvažovat zlomky, jejichž čitatel je z R a jmenovatel z M. Pro konstrukci podílového tělesa se používá $M = R \setminus \{0\}$, ale obecně můžeme uvažovat i menší multiplikativní množiny. Např. v oboru $\mathbb{Z}$ lze pro libovolné prvočíslo p uvažovat množinu $M_p = \{a : \ p \nmid a\}$, resp. obecně lze tuto množinu uvažovat pro libovolný *prvočinitel* p v daném oboru (viz sekce 6).

Konstrukce. Buď $\mathbf{R}$ obor integrity a $M \subseteq R$ multiplikativní množina. Definujeme relaci $\sim$ na množině $R \times M$ předpisem

$$(a, b) \sim (c, d) \quad \iff \quad ad = bc.$$

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 krácením prvkem $d \neq 0$ dostaneme $af = be$ (ke krácení potřebujeme předpoklad, že $\mathbf{R}$ je obor integrity!).

Blok $[(a, b)]_\sim$ této ekvivalence budeme nazývat *zlomek* a značit jej $\frac{a}{b}$. Uvažujme množinu Q všech zlomků a definujme na ní operace a konstanty

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \qquad -\frac{a}{b} = \frac{-a}{b}, \qquad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \qquad 0 = \frac{0}{1}, \qquad 1 = \frac{1}{1}.$$

Jsou tyto operace dobře definované? Předně, aby jmenovatel součtu a součinu zůstal v M, potřebujeme fakt, že M je multiplikativní. 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. Pro sčítání potřebujeme 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 $cd' = c'd$. Odčítání a násobení se ověří podobně.

Výsledná struktura $\mathbf{Q} = (Q, +, -, \cdot, 0)$ se nazývá *lokalizace* oboru $\mathbf{R}$ podle M a jak si dokážeme, jde o obor integrity. V případě $M = R \setminus \{0\}$ pak hovoříme o *podílovém tělese* oboru $\mathbf{R}$.

Tvrzení 1.7 (vlastnosti lokalizace). *Buď* $\mathbf{R}$ *obor integrity,* M *multiplikativní podmnožina a* $\mathbf{Q}$ *výsledek právě popsané konstrukce. Pak*

(1) $\mathbf{Q}$ *je obor integrity,*
(2) *množina* $\{\frac{a}{1} : a \in R\}$ *tvoří podobor oboru* $\mathbf{Q}$*, který je izomorfní oboru* $\mathbf{R}$*,*
(3) *je-li* $M = R \setminus \{0\}$*, pak je* $\mathbf{Q}$ *těleso.*

Důkaz. (1) 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 $0 \cdot 1 \neq 1 \cdot 1$.
- Součin nenulových prvků: pokud $0 = \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$, pak $a = 0$ nebo $c = 0$, protože $\mathbf{R}$ je obor integrity.

(2) Vzhledem k tomu, že $\frac{a}{1}+\frac{b}{1}=\frac{a+b}{1}$ a $\frac{a}{1}\cdot\frac{b}{1}=\frac{a\cdot b}{1}$, zlomky tvaru $\frac{a}{1}$ tvoří podobor a zobrazení $a\mapsto\frac{a}{1}$ je izomorfismus z $\mathbf{R}$ na tento podobor.

(3) Všimněte si, že $\frac{a}{b}=0=\frac{0}{1}$ právě tehdy, když $a\cdot 1=b\cdot 0$, čili když $a=0$. Tedy, pokud $M=R\setminus\{0\}$, pak pro každé $\frac{a}{b}\neq 0$ platí $\frac{a}{b}\cdot\frac{b}{a}=\frac{ab}{ba}=\frac{1}{1}$, čili $\mathbf{Q}$ je těleso. □

Příklad. Těleso racionálních čísel $\mathbb{Q}$ je *definováno* jako podílové těleso oboru $\mathbb{Z}$. Volbou $M=M_2=\{a\in\mathbb{Z}:\ a \text{ je liché}\}$ dostaneme obor $\{\frac{a}{b}:\ b \text{ je liché}\}\leq\mathbb{Q}$, který není tělesem (prvek 2 nemá inverz).

Příklad. Je-li $\mathbf{T}$ těleso, jeho podílové těleso je izomorfní s $\mathbf{T}$: každý zlomek je ekvivalentní nějakému zlomku se jmenovatelem 1, neboť $\frac{a}{b}=\frac{ab^{-1}}{1}$ pro každé $a,b\in T$, $b\neq 0$.

Triviální konstrukci poskytuje také volba $M=\{1\}$ v libovolném oboru $\mathbf{R}$, příslušná lokalizace je izomorfní původnímu oboru $\mathbf{R}$.

Příklad. Podílové těleso oboru $\mathbb{Z}[i]$ sestává ze zlomků tvaru $\frac{a+bi}{c+di}$, kde $a,b,c,d\in\mathbb{Z}$, $c+di\neq 0$. Není těžké dokázat, že zobrazení

$$\frac{a+bi}{c+di}\mapsto\frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i$$

je izomorfismem z tohoto podílového tělesa do tělesa $\mathbb{Q}(i)$.

2. Polynomy

2.1. Obory polynomů.

Stěžejním objektem v algebře komutativních okruhů jsou polynomy. V této sekci si ukážeme základní vlastnosti polynomů týkající se dělitelnosti a kořenů. Začneme ovšem definicí polynomu a polynomiálního okruhu.

V celé sekci bude $\mathbf{R}$ značit nějaký komutativní okruh s jednotkou.

Definice. *Polynomem proměnné* x nad okruhem $\mathbf{R}$ rozumíme 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

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

$$(\sum_{i=0}^{m}a_ix^i)\cdot(\sum_{i=0}^{n}b_ix^i)=\sum_{i=0}^{m+n}\big(\sum_{j+k=i}a_jb_k\big)x^i.$$

Jak si nyní ukážeme, dostaneme opět komutativní okruh, který budeme značit $\mathbf{R}[x]$.

Tvrzení 2.1. *Buď* $\mathbf{R}$ *komutativní okruh s jednotkou. Pak*

(1) $\mathbf{R}[x]$ *je komutativní okruh s jednotkou;*
(2) *je-li* $\mathbf{R}$ *oborem integrity, pak* $\mathbf{R}[x]$ *je také obor integrity a platí* $\deg(fg)=\deg f+\deg g$ *pro každé dva polynomy* $f,g\neq 0$.

Důkaz. Označme $f=\sum_{i=0}^{m}a_ix^i$, $g=\sum_{i=0}^{n}b_ix^i$, $h=\sum_{i=0}^{p}c_ix^i$. (1) Ověříme postupně všechny axiomy:

- Axiomy pro sčítání jsou očividné, sčítají se nezávisle koeficienty u jednotlivých mocnin, čili rovnosti pro polynomy ihned plynou z rovností v $\mathbf{R}$.
- Komutativita násobení plyne z toho, že vzorec je symetrický vzhledem k prohození písmen a a b.
- Jednotka: z definice součinu

$$f \cdot 1 = (\sum_{i=0}^{n} a_i x^i) \cdot (1 + 0 + 0 + \ldots) = \sum_{i=0}^{n} (\sum_{j+k=i} a_j b_k) x^i,$$

  kde všechny b_i kromě b_0 jsou nulové, takže výsledkem je opět polynom f.
- Asociativita násobení: z jedné strany, $f \cdot (g \cdot h)$ je rovno

$$\begin{aligned}\left(\sum_{i=0}^{m} a_i x^i\right) \cdot \left(\left(\sum_{i=0}^{n} b_i x^i\right) \cdot \left(\sum_{i=0}^{p} c_i x^i\right)\right) &= \left(\sum_{i=0}^{m} a_i x^i\right) \cdot \left(\sum_{i=0}^{n+p} \left(\sum_{k+l=i} b_k c_l\right) x^i\right) \\ &= \sum_{i=0}^{m+n+p} \left(\sum_{j+k+l=i} a_j b_k c_l\right) x^i,\end{aligned}$$

  a je vidět, že stejně vyjde i výpočet součinu $(f \cdot g) \cdot h$.
- Distributivita se prověří podobně, viz cvičení.

(2) Vedoucím koeficientem polynomu $f \cdot g$ je prvek $a_m b_n$, který je nenulový díky tomu, že $\mathbf{R}$ je oborem integrity. □

Pro libovolné polynomy f, g platí $\deg(f + g) \leq \max(\deg f, \deg g)$, ale rovnost nastat nemusí, například pro $g = -f$. Pokud $\mathbf{R}$ není oborem integrity, stupeň součinu nemusí být součtem stupňů, například v $\mathbb{Z}_4[x]$ platí $(2x + 1) \cdot (2x + 1) = 1$.

Na závěr vysvětlíme, co jsou to *polynomy více proměnných*. Nejjednodušší definice je induktivní: polynomem v proměnných $x_1, \ldots, x_m$ se rozumí polynom v proměnné x_m, jehož koeficienty jsou polynomy v proměnných $x_1, \ldots, x_{m-1}$. Ve zkratce,

$$\mathbf{R}[x_1, \ldots, x_m] = (\mathbf{R}[x_1, \ldots, x_{m-1}])[x_m].$$

Několikanásobnou aplikací Tvrzení 2.1 dostaneme, že polynomy více proměnných nad oborem integrity tvoří obor integrity. Díky distributivitě můžeme libovolný polynom více proměnných přepsat právě jedním způsobem do tzv. *distribuovaného tvaru*

$$\sum_{k_1, \ldots, k_m = 0}^{n} a_{k_1, \ldots, k_m} x_1^{k_1} \cdots x_m^{k_m}$$

s koeficienty $a_{k_1, \ldots, k_m} \in R$. Alternativně bychom mohli definovat okruh $\mathbf{R}[x_1, \ldots, x_m]$ tak, že jeho prvky jsou všechny výrazy uvedeného tvaru, uvést vzorce pro okruhové operace a podobně jako v Tvrzení 2.1 dokázat, že jde o komutativní okruh, resp. obor integrity. Formálně jde o různé konstrukce, ale intuice velí, že výsledkem je ten samý obor; formálně bychom řekli, že výsledné okruhy jsou izomorfní. Stejně tak, až na izomorfismus, nezáleží na pořadí proměnných v induktivní definici.

Příklad. Uvažujme polynom $xy^2 + 2y^2 - 3xy + 2x^2 + 1 \in \mathbb{Z}[x, y]$. Jeho zápis vzhledem k proměnné y bude

$$(x + 2)y^2 - (3x)y + (2x^2 + 1) \in (\mathbb{Z}[x])[y],$$

a vzhledem k proměnné x to bude

$$2x^2 + (y^2 - 3y)x + (2y^2 + 1) \in (\mathbb{Z}[y])[x].$$

2.2. **Hodnota polynomu a polynomiální zobrazení.**

Definice. Buď $\mathbf{R} \leq \mathbf{S}$ obory integrity a uvažujme polynom

$$f = a_0 + a_1 x + \ldots + a_n x^n \in R[x]$$

a prvek $u \in S$. *Hodnota polynomu* f po dosazení u je definována přepisem

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

přičemž v uvedeném vzorci provádíme všechny operace (mocnění, násobení i sčítání) v oboru $\mathbf{S}$. Zobrazení

$$S \to S, \qquad u \to f(u)$$

se nazývá *polynomiální zobrazení* dané polynomem f.

Například pro $\mathbf{R} = \mathbb{Z}$, $\mathbf{S} = \mathbb{C}$, $f = x^2 + x + 1$ a $u = i$ máme $f(i) = i$. Příslušné polynomiální zobrazení je dáno předpisem $u \mapsto u^2 + u + 1$ pro $u \in \mathbb{C}$.

Je třeba striktně rozlišovat mezi polynomem jako *výrazem* (tj. jeho zápisem ve formě „vzorečku") a odvozeným *polynomiálním zobrazením*, daným hodnotami po dosazení. Pozor, různé polynomy mohou dávat stejná polynomiální zobrazení! Například pro polynom $f = x^p \in \mathbb{Z}_p[x]$ platí díky malé Fermatově větě $f(u) = u$ pro každé $u \in \mathbb{Z}_p$, a tedy určuje stejné polynomiální zobrazení jako polynom $g = x$. (Jinak to pro konečné obory být ani nemůže, protože existuje nekonečně mnoho polynomů, ale pouze konečně mnoho zobrazení na konečné množině.)

2.3. **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 $f = gh$. Je-li $\mathbf{R}$ obor integrity a $g \mid f \neq 0$, pak $\deg g \leq \deg f$ díky Tvrzení 2.1. Pokud g nedělí f, má smysl se ptát po zbytku po dělení.

Tvrzení 2.2 (dělení polynomů se zbytkem)**.** *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*

$$f = gq + r \qquad a \qquad \deg r < \deg g.$$

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. Existenci dokážeme tak, že popíšeme algoritmus, který podíl a zbytek najde. Na začátku vezmeme $q_0 = 0$, $r_0 = f$, a poté definujeme rekurzivně

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

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 v každém kroku zmenšíme stupeň zbytku, tj. $\deg r_{i+1} < \deg r_i$ pro včechna i. Indukcí snadno ověříme, že vztah $f = gq_i + r_i$ platí pro každé 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 $\mathbf{R}[x]$.

Na závěr dokážeme jednoznačnost. 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 $q_1 = q_2$, protože $g \neq 0$ a jsme v oboru integrity. $\square$

2.4. **Kořeny a dělitelnost.**

Definice. Buď $\mathbf{R} \leq \mathbf{S}$ obory integrity, $f \in R[x]$ a $a \in S$. Řekneme, že a je *kořen* polynomu f, pokud $f(a) = 0$.

Například $i \in \mathbb{C}$ je kořenem polynomu $x^2 + 1 \in \mathbb{Z}[x]$ v tělese $\mathbb{C}$. Ukážeme si, jak existence kořene souvisí s děliteli daného polynomu.

Tvrzení 2.3. *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 $q, r \in R[x]$ podíl a zbytek při dělení polynomu f polynomem $x - a$ (dělíme monickým polynomem, čili nepotřebujeme podílové těleso). Tedy $f = (x-a) \cdot q + r$ a $\deg r < \deg(x-a) = 1$, čili r je konstantní polynom. 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 je vidět jedno důležité pozorování: je-li $f \in R[x]$ a $a \in R$, pak

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

Věta 2.4 (počet kořenů polynomu). *Buď* $\mathbf{R}$ *obor integrity,* $0 \neq f \in R[x]$ *a* $\deg f = n$. *Pak má polynom* f *nejvýše* n *kořenů v* $\mathbf{R}$.

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. □

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 2.4 neplatí, není-li $\mathbf{R}$ oborem integrity, ale třeba jen komutativním okruhem. Předpoklad jsme použili v poslední fázi důkazu, když z $f(b) = (b - a) \cdot g(b) = 0$ plynulo $b - a = 0$ nebo $g(b) = 0$. Například polynom $2x \in \mathbb{Z}_4[x]$ má v $\mathbb{Z}_4$ dva kořeny $0, 2$ a polynom $x^2 + x \in \mathbb{Z}_6[x]$ má v $\mathbb{Z}_6$ čtyři kořeny $0, 2, 3, 5$.

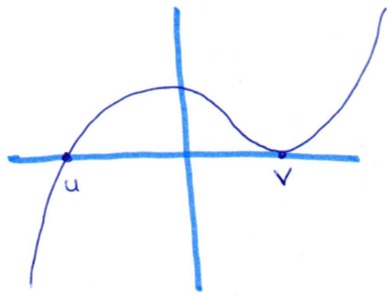
Poznámka. Věta 2.4 neplatí ani pro nekomutativní tělesa. Například v kvaternionech má polynom $x^2 + 1$ šest kořenů, $\pm i$, $\pm j$, $\pm k$.

2.5. **Vícenásobné kořeny a derivace.**
Na násobnost kořene lze pohlížet dvěma způsoby: algebraicky a geometricky. Algebraická definice vychází z Tvrzení 2.3, které zajišťuje, že každý kořen má nějakou násobnost.

Definice. Buď f polynom z $\mathbf{R}[x]$ a $a \in R$. Řekneme, že a je *n-násobným kořenem* polynomu f, pokud

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

Jinými slovy, pokud existuje polynom $g \in R[x]$ takový že $f = (x - a)^n \cdot g$ a $g(a) \neq 0$.



OBRÁZEK 3. Jednonásobný kořen u a vícenásobný kořen v.

Geometrický pohled využívá tečny grafu. Kořen a dané reálné funkce f se nazývá jednoduchý (v našem kontextu: jednonásobný), pokud křivka daná funkcí f protíná osu x „napříč“, tj. pokud tečna k f v bodě a není identická s osou x. Jinými slovy, pokud $f'(a) \neq 0$. Body, ve kterých se křivka osy „dotýká“, se nazývají násobné. Násobnost se pak dá definovat jako nejmenší n takové, že n-tá derivace $f^{(n)}$ je v bodě a nenulová.

Za chvíli si dokážeme (Věta 2.7), že obě definice násobnosti jsou ekvivalentní, a to nejen pro reálné polynomy. Nejprve však musíme definovat derivaci polynomu nad obecným oborem. Definice z diferenciálního počtu (tečna grafu) není použitelná. Jednou možností je použít známý vzorec (viz Lemma 2.5), ale elegantnější je zavést derivaci axiomaticky (navíc, tento typ definice dává smysl v libovolném komutativním okruhu funkcí obsahujícím identitu, byť my se omezíme na polynomy).

Definice. *Derivace* v $\mathbf{R}[x]$ je 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.

Lemma 2.5. *Buď* $\mathbf{R}$ *komutativní okruh. V okruhu* $\mathbf{R}[x]$ *existuje právě jedna derivace a platí*

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

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

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

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

Lemma 2.6. *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) $(f \cdot g)^{(n)} = \sum_{i=0}^{n} \binom{n}{i} \cdot f^{(i)} \cdot g^{(n-i)}$ [Leibnizova formule];
(3) $(f^n)' = n \cdot f^{n-1} \cdot f'$.

Důkaz je přímočarý 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). □

Vztah násobnosti kořene a hodnot derivací popisuje následující věta.

Věta 2.7 (násobnost kořene polynomu)**.** *Buď* $\mathbf{R}$ *obor integrity,* $0 \neq f \in R[x]$*,* $a \in R$ *a* $n \in \mathbb{N}$ *a uvažujme následující podmínky:*

(1) *a je alespoň n-násobný kořen polynomu f;*

(2) $f^{(0)}(a) = f^{(1)}(a) = \ldots = f^{(n-1)}(a) = 0$.

Podmínka (1) implikuje podmínku (2). Pokud je charakteristika oboru $\mathbf{R}$ *buď* 0*, nebo* $\geq n$*, pak jsou obě podmínky ekvivalentní.*

Důkaz. (1) $\Rightarrow$ (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 Leibnizovy 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) $\Rightarrow$ (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 Leibnizovy 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 je $\mathbf{R}$ obor 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 $\mathbf{R}$. □

Věta 2.7 umožňuje určit také přesnou násobnost. Je-li charakteristika 0 nebo $> n$, pak jsou ekvivalentní podmínky

(1') a je (přesně) n-násobný kořenem polynomu f,

(2') $f^{(0)}(a) = f^{(1)}(a) = \ldots = f^{(n-1)}(a) = 0$ a navíc $f^{(n)}(a) \neq 0$.

Důvod je jednoduchý: kdyby $f^{(n)}(a) = 0$, šlo by o alespoň $(n+1)$-násobný kořen.

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

Řešení. Nejprve ověříme, že lze použít Větu 2.7: polynom 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.) □

Pozor, v malé charakteristice derivace přesnou násobnost nedetekují!

Příklad. Uvažujme polynom $f = x^4 + x^3 = x^3(x+1)$ nad tělesem $\mathbb{Z}_3$. Pak $f' = 4x^3 + 3x^2 = x^3$, $f'' = 3x^2 = 0$, a tedy $f'' = f''' = f'''' = \ldots = 0$. Přitom násobnost kořene je omezena stupněm polynomu. Vidíme, že derivace detekují násobost kořene 2, ale nedetekují násobnost kořene 0.

Z Věty 2.7 plyne důležité a výpočetně efektivní kritérium existence vícenásobného kořene daného polynomu.

Důsledek 2.8. *Buď* **R** *obor integrity a $f \in R[x]$ stupně > 0. Jsou-li polynomy f, f' nesoudělné (viz sekce 5.2), pak f nemá žádný vícenásobný kořen v* **R**.

Důkaz. Podle Věty 2.7 je prvek $a \in R$ je vícenásobným kořenem polynomu f právě tehdy, když $f(a) = f'(a) = 0$. (Pro dvojnásobné kořeny je podmínka na charakteristiku triviální.) Tedy, kdyby měl polynom f vícenásobný kořen, oba polynomy f, f' by byly dělitelné nějakým polynomem $x - a$, a tedy soudělné. □

2.6. **Algebraická a transcendentní čísla.**

Jedním z hlavních problémů matematiky 18. a 19. století byly následující dvě otázky:

- Je dán polynom. Jak najít jeho kořeny? Lze je vyjádřit vzorci, které by používaly základní aritmetické operace na koeficientech?
- Je dáno číslo (reálné či komplexní). Existuje celočíselný polynom, jehož je toto číslo kořenem? Jak jej najít?

Odpověď na první otázku dává *Galoisova věta* (Věta 25.1), která charakterizuje polynomy, jejichž kořeny lze vyjádřit vzorci. Pro polynomy stupně ≤ 4 existují tzv. *Cardanovy vzorce* (sekce 25.1), ale pro některé polynomy stupně ≥ 5 žádné vzorce neexistují a v praxi se kořeny hledají pouze přibližně, pomocí numerických metod. V této podsekci se podíváme podrobněji na druhou otázku.

Definice. Komplexní 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 $bx - a$.
- Odmocniny jsou algebraické, například $\sqrt[n]{s}$ je kořenem polynomu $x^n - s$.
- Leckterá iracionální čísla jsou algebraická, i když to není na první pohled vidět. Například $\sqrt{2} + \sqrt{3}$, příslušným polynomem je $x^4 - 10x^2 + 1$.
- Pomocí teorie tělesových rozšíření si později dokážeme, že součet, rozdíl, součin a podíl algebraických čísel je algebraické číslo (Věta 22.8).

Příklad. Již Leonhard Euler zřejmě předpokládal, že ne každé číslo je algebraické, ale první prokazatelně transcendetní číslo bylo předvedeno mnohem později.

- V roce 1840 dokázal Joseph Liouville, že neracionální algebraická čísla nelze, v jistém smyslu, dobře aproximovat racionálními čísly. Z toho důvodu nemůže být algebraické například číslo $\sum_{i=1}^{\infty} 10^{-i!}$ (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* (ve smyslu pravděpodobnosti dané rovnoměrným rozdělením, tj. náhodné reálné číslo je s pravděpodobností 1 transcendentní).

Každý ze známých důkazů transcendence konkrétních čísel je poměrně komplikovaný. Nikoliv však argument Cantorův: poměrně jednoduchým způsobem dokázal, že existuje spousta transcedentních čísel, aniž by musel nějaké nalézt. Jeho argument je založen na počítání: transcedentních čísel je mnohem víc (nespočetně) než těch algebraických (těch je jen spočetně). Cantorův důkaz, který byl jednou z hlavních motivací vzniku teorie množin, nyní ukážeme.

Připomeňme, že nekonečná množina se nazývá *spočetná*, pokud lze její prvky seřadit do posloupnosti indexované přirozenými čísly (tj. jde o množinu stejně velkou jako $\mathbb{N}$). Všechny ostatní (tj. větší) nekonečné množiny nazýváme *nespočetné*.

Nejprve s všimněte, že sjednocení dvou spočetných množin je spočetné: je-li $A = \{a_1, a_2, \dots\}$ a $B = \{b_1, b_2, \dots\}$, pak $A \cup B = \{a_1, b_1, a_2, b_2, \dots\}$.

Tedy množina $\mathbb{Z}$ je spočetná (sjednocení kladných a záporných čísel). Dokonce i množina $\mathbb{Q}$ je spočetná: seřaďte kladná racionální čísla do posloupnosti podle součtu čitatele a jmenovatele (ty se stejným součtem seřaďte libovolně), proveďte to samé pro záporná, vezměte sjednocení a přidejte na začátek nulu.

Tvrzení 2.9. *Množina algebraických čísel je spočetná.*

Důkaz. Definujme *index polynomu* $f = a_0 + a_1x + \ldots + a_nx^n \neq 0$ jako součet $|a_0| + |a_1| + \ldots + |a_n| + n$. Všimněte si, že existuje jen konečně mnoho celočíselných polynomů daného indexu (např. index 1: $f = \pm 1$; index 2: $f = \pm 2$, $f = \pm x$; index 3: $f = \pm 3$, $f = \pm 2x$, $f = \pm x \pm 1$, $f = \pm x^2$), čili všechny polynomy 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 2.10. *Množina reálných čísel je nespočetná.*

Důkaz. Kdyby byla množina reálných čísel spočetná, byl by jistě spočetný i interval $[0, 1)$, 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 $b = 0, b_1b_2b_3\ldots$ 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 $[0, 1)$. (K tomu, aby byl tento argument korektní, je třeba se vyhnout rozvojům končícím samými devítkami.) □

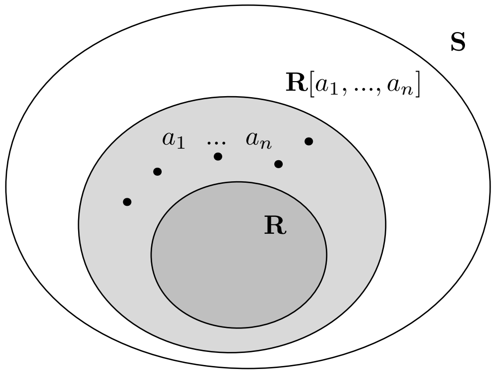
Každé reálné číslo je algebraické nebo transcendentní. Těch prvních je jen spočetně, čili těch druhých musí být nespočetně. Tedy, nejen že musí transcendentní čísla existovat, ale je jich *mnohem* více, než těch racionálních.

V sekci 22 dáme do souvislosti algebraičnost daného čísla se stupněm jistého tělesového rozšíření a také si řekneme, jak pro algebraická čísla hledat polynomy, jichž jsou kořenem.

3. Číselné obory

3.1. **Okruhová a tělesová rozšíření.**

Asi nejzajímavější využití abstraktní teorie dělitelnosti je v teorii čísel. Předmětem studia jsou různá rozšíření celých čísel, jako třeba Gaussova a Eisensteinova čísla představená v sekci 1.1. Jejich teorie je zajímavá sama o sobě, ale také se dá využít k řešení úloh o celých číslech, jak si ukážeme na příkladu diofantických rovnic v sekci 6.3.



Obrázek 4. Rozšíření $\mathbf{R}[a_1, \ldots, a_n] \leq \mathbf{S}$.

V této podsekci definujeme obecný pojem rozšíření a ukážeme, že jeho prvky lze vyjádřit pomocí hodnot polynomů. V další části se pak budeme věnovat specificky rozšířením celých čísel.

Definice. Buď $\mathbf{R} \leq \mathbf{S}$ komutativní okruhy a $a_1, \ldots, a_n \in S$. Definujeme

- $\mathbf{R}[a_1, \ldots, a_n]$ jako nejmenší podokruh okruhu $\mathbf{S}$ obsahující množinu R i prvky $a_1, \ldots, a_n$; říká se mu *okruhové rozšíření* $\mathbf{S}$ *o prvky* $a_1, \ldots, a_n$.

Jsou-li $\mathbf{R}, \mathbf{S}$ tělesa, pak definujeme

- $\mathbf{R}(a_1, \ldots, a_n)$ jako nejmenší podtěleso tělesa $\mathbf{S}$ obsahující množinu R i prvky $a_1, \ldots, a_n$; říká se mu *tělesové rozšíření* $\mathbf{S}$ *o prvky* $a_1, \ldots, a_n$.

Příklad (Gaussova čísla)**.** Gaussova celá čísla lze zapsat jako obor $\mathbb{Z}[i]$: jde o nejmenší podobor tělesa $\mathbb{C}$ obsahující jak celá čísla, tak číslo i. Analogicky, Gaussova racionální čísla lze zapsat jako $\mathbb{Q}[i]$: jde o nejmenší podobor tělesa $\mathbb{C}$ obsahující jak racionální čísla, tak číslo i. Obor $\mathbb{Q}[i]$ je tělesem, protože

$$(a+bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i \in \mathbb{Q}[i],$$

čili platí $\mathbb{Q}[i] = \mathbb{Q}(i)$. Pro reálná čísla pak platí $\mathbb{R}[i] = \mathbb{R}(i) = \mathbb{C}$.

Gaussova čísla lze přirozeně zobecnit dvěma způsoby: místo $i = \sqrt{-1}$ budeme přidávat druhé odmocniny z jiných čísel, anebo vyšší komplexní odmocniny z jedné.

Příklad (kvadratická rozšíření)**.** Pro libovolné celé číslo s uvažujme *kvadratické rozšíření*

$$\mathbb{Z}[\sqrt{s}] = \{a + b\sqrt{s} : \ a, b \in \mathbb{Z}\} \leq \mathbb{C},$$
$$\mathbb{Q}[\sqrt{s}] = \mathbb{Q}(\sqrt{s}) = \{a + b\sqrt{s} : a, b \in \mathbb{Q}\} \leq \mathbb{C}.$$

Není těžké nahlédnout, že množina na pravé straně je skutečně podoborem, resp. podtělesem, tělesa $\mathbb{C}$.

Příklad (cyklotomická rozšíření)**.** Pro $\zeta_n = e^{2\pi i/n}$ (tzv. primitivní n-tá odmocnina z jedné) uvažujme *n-té cyklotomické rozšíření*

$$\mathbb{Z}[\zeta_n] = \{a_0 + a_1\zeta_n + a_2\zeta_n^2 + \ldots + a_{n-1}\zeta_n^{n-1} : \ a_0, \ldots, a_{n-1} \in \mathbb{Z}\} \leq \mathbb{C},$$
$$\mathbb{Q}[\zeta_n] = \mathbb{Q}(\zeta_n) = \{a_0 + a_1\zeta_n + a_2\zeta_n^2 + \ldots + a_{n-1}\zeta_n^{n-1} : \ a_0, \ldots, a_{n-1} \in \mathbb{Z}\} \leq \mathbb{C}.$$

Pro $n = 3$ dostáváme Eisensteinova čísla, pro $n = 4$ Gaussova čísla. Vyjádření na pravé straně není jednoznačné, například pro $n = 3$ je $\zeta_3^2 = -1 - \zeta_3$. Důkaz, že $\mathbb{Q}[\zeta_n] = \mathbb{Q}(\zeta_n)$, není zdaleka tak jednoduchý jako pro kvadratická rozšíření. Mnohem později si ukážeme obecné Tvrzení 22.2, z něhož tento fakt ihned plyne.

Příklad. Můžeme uvažovat i rozšíření o více prvků, například

$$\mathbb{Z}[\sqrt{2}, \sqrt{3}] = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : \ a, b, c, d \in \mathbb{Z}\}.$$

Tvrzení 3.1 (struktura okruhových rozšíření). *Buď* $\mathbf{R} \leq \mathbf{S}$ *komutativní okruhy a* $a \in R$. *Pak*

$$\begin{aligned}\mathbf{R}[a] &= \{f(a):\ f \in R[x]\} \\ &= \{u_0 + u_1 a + \ldots + u_n a^n:\ n \in \mathbb{N},\ u_0, \ldots, u_n \in R\}.\end{aligned}$$

Jsou-li $\mathbf{R} \leq \mathbf{S}$ *tělesa, pak*

$$\mathbf{R}(a) = \left\{ \frac{f(a)}{g(a)}:\ f, g \in R[x],\ g(a) \neq 0 \right\}.$$

Důkaz. Označme $M = \{f(a):\ f \in R[x]\}$. Je potřeba dokázat, že množina M

(1) tvoří podokruh okruhu $\mathbf{S}$,
(2) obsahuje $R \cup \{a\}$,
(3) je nejmenší podmnožinou okruhu $\mathbf{S}$ splňující tyto podmínky.

(1) Mějme dva prvky $f(a), g(a) \in M$, kde $f, g \in R[x]$. Jejich součet $f(a) + g(a) = (f + g)(a)$ je také v M, protože $f + g \in R[x]$, a analogický argument lze použít pro součin i prvek $-f(a)$. Volbou $f = 0$ dostaneme $0 \in M$.

(2) Volbou konstantních polynomů dostaneme $R \subseteq M$. Volbou $f = x$ dostaneme $a \in M$.

(3) Uvažujme libovolný podokruh $\mathbf{U}$ obsahující $R \cup \{a\}$. Tento podokruh musí obsahovat všechny mocniny a^i, jejich libovolné násobky prvky z R, a také součty těchto násobků. Čili musí obsahovat všechny prvky tvaru $u_0 + u_1 a + \ldots + u_n a^n$, kde $u_0, \ldots, u_n \in R$, a tedy $M \subseteq U$.

Vyjádření tělesových rozšíření se dokáže analogicky, navíc musíme dát pozor na inverzní prvky. □

Příklad. Uvažujme kvadratická rozšíření $\mathbb{Z}[\sqrt{s}]$. Vzhledem k tomu, že $\sqrt{s}^2 = s$, $\sqrt{s}^3 = s\sqrt{s}$, atd., hodnota libovolného polynomu $f \in \mathbb{Z}[x]$ na prvku $\sqrt{s}$ bude rovna číslu tvaru $a + b\sqrt{s}$, $a, b \in \mathbb{Z}$. Čili $\mathbb{Z}[\sqrt{s}] = \{f(\sqrt{s}) : f \in \mathbb{Z}[x]\} = \{a + b\sqrt{s}:\ a, b \in \mathbb{Z}\}$.

Analogicky, pro cyklotomická rozšíření dostaneme vyjádření $\mathbb{Z}[\zeta_n] = \{a_0 + a_1\zeta_n + \ldots + a_{n-1}\zeta_n^{n-1}:\ a_0, \ldots, a_{n-1} \in \mathbb{Z}\}$, protože $\zeta_n^n = 1$, $\zeta_n^{n+1} = \zeta_n$ atd.

Pro každé těleso $\mathbf{T}$ platí $\mathbf{T}[a] \leq \mathbf{T}(a)$, protože v okruhu $\mathbf{T}(a)$ navíc požadujeme přítomnost inverzních prvků. Za jakých podmínek platí $\mathbf{T}[a] = \mathbf{T}(a)$? V sekci 22 si ukážeme, že to nastane právě tehdy, když je a tzv. *algebraický prvek*, tedy když je kořenem nějakého nenulového polynomu z $\mathbf{T}[x]$ (jde o přímočaré zobecnění pojmu algebraického čísla z předchozí sekce). Jednu implikaci si můžeme dokázat hned.

Tvrzení 3.2. *Buď* $\mathbf{T} \leq \mathbf{S}$ *tělesa a* $a \in S$ *prvek, který není kořenem žádného nenulového polynomu z* $\mathbf{T}[x]$. *Pak* $\mathbf{T}[a] \neq \mathbf{T}(a)$.

Důkaz. Podle Tvrzení 3.1 je $\mathbf{T}[a] = \{f(a):\ f \in T[x]\}$. Kdyby se v této množině nacházel prvek a^{-1}, pak by existoval polynom $f \in T[x]$ takový, že $f(a) = a^{-1}$, čili $af(a) = 1$, a tedy a by bylo kořenem nenulového polynomu $xf - 1 \in T[x]$, spor. □

3.2. **Kvadratická rozšíření celých čísel.**

V této učebnici se pro jednoduchost soustředíme na kvadratická rozšíření typu $\mathbb{Z}[\sqrt{s}]$, kde $|s|$ není druhou mocninou přirozeného čísla. (Obecněji bychom mohli studovat tzv. celistvá rozšíření stupně 2, kam by patřily také některé obory $\mathbb{Z}[\frac{1+\sqrt{s}}{2}]$, ale nebudeme věci komplikovat.) Základním nástrojem pro práci v kvadratických rozšířeních je norma.

Definice. *Norma* na oboru $\mathbb{Z}[\sqrt{s}]$ se definuje jako zobrazení

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

(V teorii čísel se zavádí norma bez absolutní hodnoty a značí se N, my se však přidržíme značení kompatibilního s definicí eukleidovské normy v sekci 7.1.)

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 použít geometrický náhled.

Prvek u komutativního okruhu $\mathbf{R}$ se nazývá *invertibilní*, pokud existuje $v \in R$ takové, že $uv = 1$; toto v značíme u^{-1}. Stěžejním pozorováním je fakt, že norma je multiplikativní a identifikuje invertibilní prvky.

Tvrzení 3.3 (vlastnosti normy kvadratických rozšíření)**.** *Pro každá $u, v \in \mathbb{Z}[\sqrt{s}]$ platí*

(1) $\nu(u \cdot v) = \nu(u) \cdot \nu(v)$,
(2) $\nu(u) = 1$ *právě tehdy, když je u invertibilní,*
(3) *pokud $u \mid v$ a $v \nmid u$, pak $\nu(u) \mid \nu(v)$ a $\nu(u) \neq \nu(v)$.*

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) ($\Rightarrow$) Označme $u = a + b\sqrt{s}$. 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 $u^{-1} = \pm(a - b\sqrt{s})$. ($\Leftarrow$) Je-li $uv = 1$, pak z (1) plyne $1 = \nu(1) = \nu(uv) = \nu(u)\nu(v)$, a tedy $\nu(u) = \nu(v) = 1$.

(3) Dělitelnost ihned plyne z (1): pokud $v = uw$, pak $\nu(v) = \nu(u)\nu(w)$, čili $\nu(u) \mid \nu(v)$. Pokud by se normy rovnaly, pak $\nu(w) = 1$, čili podle (2) by byl w invertibilní a mohli bychom psát $u = vw^{-1}$, tedy $v \mid u$, spor. $\square$

Invertibilní prvky v oboru $\mathbb{Z}[\sqrt{s}]$ tedy můžeme hledat jako řešení diofantické rovnice $x^2 - sy^2 = \pm 1$ (kde $s \in \mathbb{Z}$ je fixní a $x, y \in \mathbb{Z}$ jsou neznámé). Pro s záporné je řešení očividné.

Příklad. V oboru $\mathbb{Z}[i]$ je $\nu(u) = \nu(a + bi) = a^2 + b^2 = \pm 1$ právě tehdy, když $u \in \{\pm 1, \pm i\}$. V oborech $\mathbb{Z}[\sqrt{s}]$ pro $s < -1$ je $\nu(u) = \nu(a + b\sqrt{s}) = a^2 - sb^2 = \pm 1$ právě tehdy, když $u \in \{\pm 1\}$.

Pro s kladné jde o velmi zajímavý problém. Diofantické rovnice $x^2 - sy^2 = 1$ a $x^2 - sy^2 = -1$ se nazývají *Pellovy rovnice* a matematiky zajímaly z různých důvodů už od starověku. První postup na hledání netriviálních řešení vymyslel Brahmagupta v 7. století, moderní přístup používá k řešení řetězové zlomky. Obecné řešení je nad rámec tohoto textu, ale ukážeme si jeden konkrétní příklad.

Příklad. V oboru $\mathbb{Z}[\sqrt{2}]$ je $\nu(u) = \nu(a + b\sqrt{2}) = |a^2 - 2b^2|$. 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 je u^k také invertibilní pro libovolné $k \in \mathbb{N}$: inverzním prvkem bude $(u^{-1})^k$. Obor $\mathbb{Z}[\sqrt{2}]$ tedy obsahuje nekonečně mnoho invertibilních prvků $(1 + \sqrt{2})^k$, pro libovolné k.

Norma umožňuje definovat *dělení se zbytkem*: pro daná u, v hledáme q, r splňující rovnici $u = vq + r$ a požadavek, že zbytek je menší než dělitel, tj. $\nu(r) < \nu(v)$. Dělit se zbytkem lze například v oborech $\mathbb{Z}[i]$, $\mathbb{Z}[i\sqrt{2}]$ nebo $\mathbb{Z}[\sqrt{2}]$. Důkaz si ukážeme pro obor $\mathbb{Z}[i]$, algoritmus dělení lze vyčíst z důkazu.

Tvrzení 3.4 (dělení Gaussových čísel se zbytkem)**.** *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

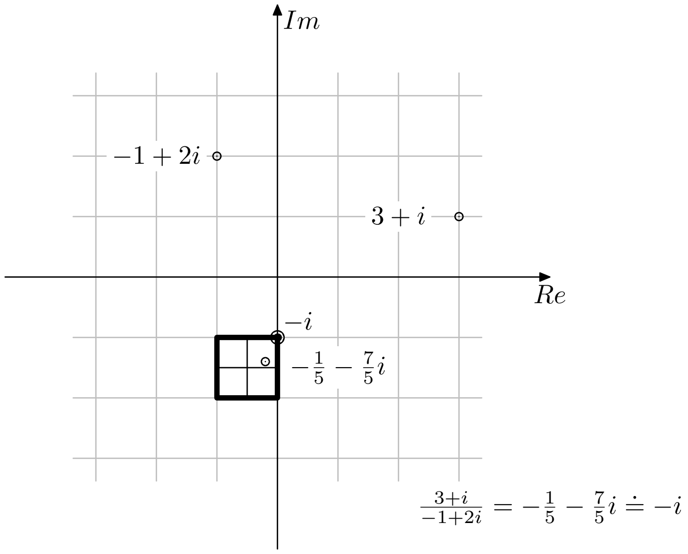
$$z = \frac{u}{v} \in \mathbb{C}$$

(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

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



OBRÁZEK 5. Dělení se zbytkem v $\mathbb{Z}[i]$.

□

Za pozornost stojí, že podíl a zbytek není určen jednoznačně: například $z = \frac{1}{2} + \frac{1}{2}i$ lze zaokrouhlit čtyřmi způsoby, každý z nich bude splňovat uvedené podmínky.

Pro obor $\mathbb{Z}[i\sqrt{2}]$ důkaz projde také, protože střed obdélníka, ve kterém se nachází přesný podíl, má vzdálenost od vrcholů méně než 1. Pro $\mathbb{Z}[i\sqrt{3}]$ už důkaz neprojde, protože střed obdélníka má vzdálenost od vrcholu rovnou 1. V tomto oboru, stejně jako třeba v oboru $\mathbb{Z}[\sqrt{5}]$, podíl a zbytek v uvedeném smyslu neexistuje.

Abstraktní teorie dělitelnosti

4. Elementární teorie čísel

Cílem této sekce je shrnout základní vlastnosti oboru celých čísel z hlediska dělitelnosti: prvočíselné rozklady, Eukleidův algoritmus, kongruence, Eulerovu větu a čínskou větu o zbytcích. Většinu z těchto principů později zobecníme (například na polynomy či větší číselné obory), ale přesto je důležité začít tímto speciálním případem.

Většina teorie této sekce byla v nějaké formě známa již ve starověku a v moderní podobě byly formulovány Carlem Friedrichem Gaussem v jeho slavné knize *Disquisitiones Arithmeticae* z roku 1801, která položila základ moderní teorie čísel.

Základním objektem v této sekci bude množina přirozených čísel $\mathbb{N} = \{1, 2, 3, 4, \ldots\}$, respektive množina celých čísel $\mathbb{Z} = \{\ldots, . - 2, -1, 0, 1, 2, \ldots\}$, a standardní aritmetické operace $+, -, \cdot$ na těchto číslech. Je řada způsobů, jak čísla definovat: můžeme je vybudovat v rámci teorie množin, můžeme je zavést axiomaticky (za standardní axiomatizaci se považují Peanovy axiomy, které jsou založeny na principu matematické indukce). Ani jedním z těchto postupů se zabývat nebudeme a budeme vycházet ze základních středoškolských poznatků.

Pokud napíšeme „číslo“, myslíme v této sekci celé číslo.

4.1. **Dělitelnost a základní věta aritmetiky.**

Buď a, b celá čísla. Řekneme, že číslo b *dělí* číslo a, píšeme $b \mid a$, pokud existuje 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.* Pokud b nedělí a, má smysl se ptát po zbytku po dělení.

Tvrzení 4.1 (dělení celých čísel se zbytkem)**.** *Buď $a, b \in \mathbb{Z}$, $b \neq 0$. Pak existuje právě jedna dvojice celých čísel q, r splňující*

$$a = q \cdot b + r \qquad a \qquad 0 \leq r < |b|.$$

Díky jednoznačnosti můžeme definovat *celočíselný podíl* a div $b = q$ a *zbytek* a mod $b = r$. Je vidět, že $b \mid a$ právě tehdy, když a mod $b = 0$.

Důkaz. Pro jednoduchost předpokládejme $a, b > 0$, ostatní případy se vyřeší analogicky. Buď q největší číslo splňující $q \cdot b \leq a$ a položme $r = a - q \cdot b$. Zřejmě $0 \leq r < b$ a platí výše uvedený vztah.

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 $p > 1$, 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 4.2 (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.*

Tuto „samozřejmost“ jako první dokázal Eukleides v 4. století př.n.l. a v dnešní době ji každý středoškolák dobře zná, nicméně, přiznejme si, kde z vás by to uměl dokázat? Tedy existenci rozkladu lze dokázat poměrně snadno indukcí.

Důkaz Věty 4.2, existence rozkladu. Buď a nejmenší přirozené číslo, pro nějž neexistuje prvočíselný rozklad. To nemůže být prvočíslem, jinak bychom měli rozklad $a = a^1$. Čili a je složené a můžeme jej rozložit jako $a = b \cdot c$ 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 existence prvočíselných rozkladů je fakt, že existuje nekonečně mnoho prvočísel. Kdyby jich bylo jenom konečně mnoho, označme je $p_1, \dots, p_n$ a 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. Nástrojem nám bude Bézoutova rovnost, která vyjadřuje největší společný dělitel jako lineární kombinaci.

4.2. **Eukleidův algoritmus a Bézoutova rovnost.**

Největší společný dělitel celých čísel a a b je největší přirozené číslo c splňující zároveň $c \mid a$ a $c \mid b$, značíme jej $\mathrm{NSD}(a, b)$. Čísla a, b nazveme *nesoudělná*, pokud $\mathrm{NSD}(a, b) = 1$. Podobně, *nejmenší společný násobek* čísel a a b je nejmenší číslo c splňující zároveň $a \mid c$ a $b \mid c$, značíme jej $\mathrm{NSN}(a, b)$. Použitím základní věty aritmetiky je snadné nahlédnout, že

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

(vzorec nebudeme k důkazu této věty potřebovat). Vzhledem k tomu, že $\mathrm{NSD}(a, b) = \mathrm{NSD}(\pm a, \pm b)$, budeme se dále zabývat výpočtem NSD pro kladná čísla.

Jeden známý postup výpočtu NSD je pomocí 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 výsledek, což je problém. Nejsme tedy schopni dokázat správnost této metody, aniž bychom dokončili důkaz základní věty aritmetiky. (Skutečným příkladem, že tato metoda v obecnosti nefunguje, je např. následující situace v oboru $\mathbb{Z}[\sqrt{5}]$: zde $4 = 2 \cdot 2 = (1 + \sqrt{5})(-1 + \sqrt{5})$. Z prvního rozkladu bychom vydedukovali $\mathrm{NSD}(2, 4) = 2$, z druhého $\mathrm{NSD}(2, 4) = 1$. Viz sekce 5.)

Druhým používaným postupem výpočtu NSD je *Eukleidův algoritmus.* Ten je založen na následujícím pozorování, nezávislém na základní větě aritmetiky.

Lemma 4.3. *Pro libovolná celá čísla a, b platí*

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

Důkaz. Označme $q = a \operatorname{div} b$. Pak

$$a = b \cdot q + (a \bmod b),$$

a tedy dané číslo c dělí obě čísla a, b právě tehdy, když c dělí obě čísla $b, a \bmod b$. Čili obě dvojice mají stejné společné dělitele, mají tedy stejného i toho největšího. □

Algoritmus vezme daná čísla $a \geq b \geq 0$ a buduje posloupnost tak, že vždy vezme zbytek po dělení předposledního čísla posledním. Odpovědí je poslední nenulová hodnota. Formálně, 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. Například pro $\mathrm{NSD}(168, 396)$ dostáváme posloupnost 396, 168, 60, 48, 12, 0, a tedy $\mathrm{NSD}(168, 396) = 12$. Z Lemmatu 4.3 vidíme, že

$$\mathrm{NSD}(a, b) = \mathrm{NSD}(a_0, a_1) = \mathrm{NSD}(a_1, a_2) = \ldots = \mathrm{NSD}(a_k, 0) = a_k,$$

což je poslední nenulová hodnota v posloupnosti.

Rozšířením Eukleidova algoritmu lze dokázat následující vlastnost.

Tvrzení 4.4 (Bézoutova rovnost)**.** *Pro každou dvojici celých čísel a, b existují celá čísla u, v (tzv.* Bézoutovy koeficienty*) splňující*

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

Důkaz. 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

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

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 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 tvrzení.

Lemma 4.5. *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 $2 \mid (1 + \sqrt{5})(-1 + \sqrt{5})$, přesto $2 \nmid \pm 1 + \sqrt{5}$; jednoznačnost tedy hraje roli.)

Důkaz. Předpokládejme, že $p \nmid a$. Pak $\mathrm{NSD}(a, p) = 1$, protože je p prvočíslo, a tedy podle Tvrzení 4.4 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 4.6. *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 z Věty 4.2.

Důkaz Věty 4.2, jednoznačnost rozkladu. 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. □

4.3. **Kongruence.**

Symbol $\equiv$ pro kongruenci, zavedený Gaussem ve zmiňované knize *Disquisitiones Arithmeticae*, usnadňuje zápis při 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 $m \mid a - b$.

Předně si všimněte, že $a \equiv b \pmod m$ právě tehdy, když a a b dávají stejný zbytek po dělení m: napišme si $a = mq_1 + r_1$ a $b = mq_2 + r_2$, čili máme $a - b = m(q_1 - q_2) + (r_1 - r_2)$ a ihned vidíme, že $m \mid a - b$ právě tehdy, když $m \mid r_1 - r_2$, čili když jsou zbytky stejné.

Z uvedené interpretace 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 $a \equiv c \pmod m$.

Druhou základní vlastností je invariance vůči základním operacím.

Tvrzení 4.7 (vlastnosti kongruence)**.** *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. Z předpokladu $m \mid a - b$ a $m \mid c - d$ plyne

- $m \mid (a - b) + (c - d) = = (a + c) - (b + d)$ a podobně pro operaci $-$;
- $m \mid (a - b) \cdot c$ 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 indukcí: $a^{i+1} = a^i \cdot a \equiv b^i \cdot b = b^{i+1} \pmod m$. □

Obě vlastnosti lze interpretovat tak, že symbol $\equiv$ můžeme používat stejným způsobem, jako rovnítko: reflexivita říká, že z $a = b$ plyne také $a \equiv b$, symetrie říká, že zápis je platný zleva doprava i zprava doleva, a tranzitivita říká, že máme-li sérii po sobě jdoucích kongruencí, pak je číslo úplně vlevo kongruentní číslu úplně vpravo. Invariance vůči operacím pak umožňuje ve výpočtu nahrazovat navzájem kongruentní čísla. Ukážeme si to na jednoduchém příkladu.

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

Řešení. Protože $66 \equiv 0$ a $77 \equiv -1 \pmod 6$, můžeme psát

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

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 4.8 (vlastnosti kongruence)**.** *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) Levá strana říká, že existuje q takové, že $a - b = mq$. Pravá strana říká, že existuje q takové, že $ca - cb = c(a - b) = mq$. Ekvivalence obout tvrzení je nyní zřejmá.

(2) Implikace ($\Rightarrow$) plyne z Tvrzení 4.7. K opačné implikaci potřebujeme základní větu aritmetiky. Nechť $m \mid ca - cb = c(a - b)$. Uvažujme libovolnou mocninu p^k z prvočíselného rozkladu čísla m. Protože jsou čísla c, m nesoudělná, musí $p^k \mid a - b$. Úvahou přes všechny mocniny dostaneme $m \mid a - b$. □

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

Řešení. Použijeme několikrát Tvrzení 4.8.

(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 $x = 11 + 21k$, $k \in \mathbb{Z}$. □

4.4. **Eulerova věta a kryptosystém RSA.**

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 $77^{123} \equiv 7^{123} \equiv (-3)^{123} \pmod{10}$, ale bez další teorie nám nezbývá, než mocnit sedmičku nebo trojku. Např. pro sedmičku dostáváme $7^1 = 7$, $7^2 = 49 \equiv 9$, $7^3 \equiv 7 \cdot 9 \equiv 3$, $7^4 \equiv 7 \cdot 3 \equiv 1$, $7^5 \equiv 7 \cdot 1 = 7$ a vidíme, že poslední cifry se opakují s periodou 4. Vzhledem k tomu, že $123 \bmod 4 = 3$, dostáváme $7^{123} \equiv 7^3 \equiv 3 \pmod{10}$. □

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 $k \in \{1, \dots, n-1\}$ nesoudělných s číslem n, tj. splňujících $\mathrm{NSD}(k, n) = 1$.

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 s ním nesoudělná jsou všechna menší čísla.

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

Tvrzení 4.9. *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. $\varphi(4056) = \varphi(2^3 \cdot 3^1 \cdot 13^2) = 2^2 \cdot 1 \cdot 3^0 \cdot 2 \cdot 13^1 \cdot 12 = 1248.$

Vzorec se celkem snadno dokáže pomocí čínské věty o zbytcích, se kterou se seznámíme v příští části. Teď se podíváme na samotnou Eulerovu větu.

Věta 4.10 (Eulerova věta). *Jsou-li a, m nesoudělná přirozená čísla, pak*

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

Leonhard Euler publikoval tuto větu v roce 1763. Již dříve, v roce 1736, pak dokázal speciální případ, kdy je m prvočíslo. Objev tohoto vztahu bývá připisován Pierre de Fermatovi, objevuje se v jednom z jeho dopisů z roku 1640, a proto bývá nazýván malá Fermatova věta.

Důsledek 4.11 (malá Fermatova věta). *Je-li p prvočíslo a $p \nmid a$, pak*

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

Označme pro účely této sekce

$$\Phi_m = \{k \in \{1, \dots, m-1\} : \ \mathrm{NSD}(k, m) = 1\}.$$

Eulerovu funkci pak můžeme zapsat jako $\varphi(m) = |\Phi_m|$. K důkazu Eulerovy věty se nám bude hodit jedno pomocné lemma.

Lemma 4.12. *Buďte a, m nesoudělná přirozená čísla a definujme zobrazení*

$$f_a : \Phi_m \to \Phi_m, \quad x \mapsto ax \bmod m.$$

Zobrazení f_a je dobře definované a je to bijekce.

Důkaz. Předně je třeba ověřit, že $ax \bmod m \in \Phi_m$. Jsou-li obě čísla a, x nesoudělná s m, pak je s m nesoudělné i číslo ax: kdyby existovalo prvočíslo p dělící m i ax, pak by p dělilo a nebo x (Lemma 4.5), spor s nesoudělností. Čili $1 = \mathrm{NSD}(ax, m) = \mathrm{NSD}(ax \bmod m, m)$ použitím Lemmatu 4.3.

Nyní dokážeme, že je zobrazení f_a prosté. Uvažujme $x, y \in \Phi_m$ taková, že $f_a(x) = f_a(y)$, tj. $ax \equiv ay \pmod{m}$. Podle Tvrzení 4.8 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.

Vzhledem k tomu, že je f_a zobrazením na konečné množině, musí být také na (Lemma 1.4). □

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

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

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

$$c = \prod_{b\in\Phi_m} b,$$

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 4.8 krátit a dostáváme $1 \equiv a^{\varphi(m)} \pmod{m}$. □

Poznámka. Malou Fermatovu větu je snadné dokázat přímo: indukcí podle a dokažte ekvivalentní tvrzení $a^p \equiv a \pmod{p}$. Alternativní důkaz Eulerovy věty bychom pak mohli vést takto: nejprve pomocí Fermatovy věty dokážeme indukcí podle k Eulerovu větu pro čísla tvaru $n = p^k$, 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 $u \equiv v \pmod{n}$.

Zřejmě nejpřirozenější důkaz Eulerovy věty pak uvidíme v sekci 14.2, kde ji dostaneme jako speciální případ Lagrangeovy věty aplikované na grupu $\mathbb{Z}_m^*$.

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 $7^{123} \equiv 7^3 = 3$.) □

Úloha. Spočtěte $10^{10^{10}} \bmod 21$.

Řešení. Použijeme Eulerovu větu: protože $\varphi(21) = 12$ a $\mathrm{NSD}(10, 21) = 1$, stačí zjistit zbytek po dělení 10^{10} číslem 12. Avšak čísla $10, 12$ jsou soudělná. Protože $\mathrm{NSD}(10^{10}, 12) = 4$, výsledek bude dělitelný 4. Napišme $10^{10} = 2^{10}\cdot 5^{10} \equiv 4k \pmod{12}$, podle Tvrzení 4.8 budeme řešit úlohu $2^8\cdot 5^{10} \equiv k \pmod{3}$. Nyní znovu použijeme Eulerovu větu: protože $\varphi(3) = 2$ a všechna zmíněná čísla jsou nesoudělná, máme $k = 2^8\cdot 5^{10} \equiv 2^0\cdot 5^0 = 1 \pmod{3}$, čili $10^{10} \equiv 4k = 4 \pmod{12}$, a tedy $10^{10^{10}} \equiv 10^4 = 4 \pmod{21}$. □

Poznámka. Lemma 4.12 říká, že pro každé a nesoudělné s m existuje právě jedno $b \in \{1, \dots, m-1\}$ takové, že

$$a\cdot b \equiv 1 \pmod{m}.$$

Toto b lze najít dvěma způsoby:

- podle Eulerovy věty lze vzít $b = a^{\varphi(m)-1} \bmod m$,
- Eukleidovým algoritmem spočteme Bézoutovy koeficienty $1 = \mathrm{NSD}(a, m) = ua + vm$ a vezmeme $b = u \mod m$.

Tato úloha má vyšší smysl: právě jsme popsali způsob nalezení inverzního prvku a^{-1} v okruhu $\mathbb{Z}_m$.

S Eulerovou větou je úzce spjata jedna významná aplikace: protokol *RSA* (pojmenovaný po trojici matematiků Rivest, Shamir, Adleman) na tzv. *šifrování s veřejným klíčem.* Problém je následující: Bob přijímá zprávy od řady klientů a je nepraktické, aby si s každým vyměňoval tajné heslo. Bob tedy publikuje tzv. *veřejný klíč*, pomocí něhož mu může každý poslat šifrovanou zprávu, a tajně bude držet *soukromý klíč*, pomocí něhož může pouze on zprávy dešifrovat. Popíšeme algoritmus, jak generovat klíče a jak šifrovat a dešifrovat zprávu.

Na začátku Bob inicializuje komunikační kanál. Zvolí dvě různá prvočísla p, q a spočte $N = pq$ a $\varphi(N) = (p-1)(q-1)$. Dále náhodně zvolí číslo e nesoudělné s $\varphi(N)$ a pomocí Eukleidova algoritmu spočte číslo d splňující

$$d \cdot e \equiv 1 \pmod{\varphi(N)}$$

(viz poznámka výše). Čísla N, e budou *veřejným klíčem*, ten Bob rozhlásí do světa. Číslo d bude *soukromým klíčem*, ten bude držet v tajnosti (prvočísla p, q může Bob zapomenout).

Nyní popíšeme, jak může Alice poslat Bobovi zprávu. Pro jednoduchost budeme předpokládat, že zprávu tvoří nějaké přirozené číslo $0 < x < N$ nesoudělné s N. Alice vezme veřejný klíč, vypočítá

$$y = x^e \bmod N$$

a výsledek y pošle Bobovi. Bob vezme soukromý klíč a spočítá

$$y^d \bmod N.$$

Podle Eulerovy věty vyjde

$$y^d \equiv \left(x^e\right)^d = x^{ed} \equiv x^1 = x \pmod{N},$$

protože $ed \equiv 1 \pmod{\varphi(N)}$.

Co vidí nepřítel? Zašifrovanou zprávu y a veřejný klíč N, e. Aby se dostal ke zprávě, potřebuje efektivní algoritmus, který z hodnoty $x^e \bmod N$ vypočte hodnotu x, tj. něco jako „e-tou odmocninu z x modulo N“. Očividným řešením je spočítat hodnotu $\varphi(N)$ a dále postupovat jako Bob při inicializaci, nicméně v současné době není znám algoritmus, který by uměl počítat Eulerovu funkci lépe než přes prvočíselné rozklady, ani efektivní algoritmus na hledání prvočíselného rozkladu — aspoň tedy za dodržení jistých předpokladů, např. že rozklad neobsahuje mnoho prvočísel, že jsou tato prvočísla zhruba stejně velká atd. (při stavu současné techniky stačí volit prvočísla p, q s řádově tisíci binárních cifer). Žádný jiný efektivní postup na výpočet odmocniny modulo N nebyl dosud nalezen.

Na podobném principu funguje také *digitální podpis*: Bob zprávu zašifruje svým soukromým klíčem a zveřejní původní i zašifrovanou zprávu, tedy dvojici (x, x^d). Každý může pomocí veřejného klíče ověřit, že je tato dvojice správná, výpočtem $(x^d)^e = x^{de} \equiv x \pmod{N}$. Tím je prokázáno, že autor musel znát Bobovo heslo, bez znalosti d nelze efektivně vytvářet dvojice (x, y) splňující $y^e = x$. Detaily viz libovolná moderní učebnice kryptografie.

4.5. **Čínská věta o zbytcích.**

Čínská věta o zbytcích hovoří o řešeních soustav lineárních kongruencí. Byla známa již starověkým Číňanům, je uvedena v knize Umění války autora Sun-c' ze 5. století př. n. l. Traduje se, že motivací čínské věty o zbytcích byl způsob, jakým Sun-c' počítal své vojáky. Věděl, že před bitvou měl 1000 vojáků, a chtěl je spočítat po bitvě. Velké množství lidí se špatně počítá, nechal je tedy seřadit do desetistupů, jedenáctistupů a třináctistupů a počítal, kolik mu jich zbyde mimo řady. Jinými slovy, zjistil, kolik je počet vojáků modulo 10, modulo 11 a modulo 13. Z čínské věty o zbytcích plyne, že z těchto zbytků lze jednoznačně zjistit celkový počet vojáků.

Věta 4.13 (čínská věta o zbytcích). *Buď $m_1, \ldots, m_n$ po dvou nesoudělná přirozená čísla, označme $M = m_1 \cdot \ldots \cdot m_n$. Buď $u_1, \ldots, u_n$ libovolná celá čísla. Pak 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 $x, y \in \{0, \dots, M-1\}$, 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 obě čísla x, y, a tedy i jejich rozdíl, jsou menší než M, takže nutně $x - y = 0$, čili $x = y$.

Nyní dokážeme, že nějaké řešení vůbec existuje. Uvažujme zobrazení

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

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 1.4 i na. Tedy ke každé n-tici $(u_1, \dots, u_n)$ existuje právě jedno x, které se na něj zobrazuje, a to je hledaným řešením 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 $x = 60m + 53$, $m \in \mathbb{Z}$. □

Čínská věta o zbytcích platí v mnohem obecnějším kontextu: pro polynomy (sekce 9.1) a v jistém smyslu ji lze formulovat v jakémkoliv okruhu. Pro polynomy se jednoznačnost dokáže analogicky (horní mez je daná součtem stupňů polynomů m_i), ale s existencí je to o něco složitější, místo velikosti množiny je třeba argumentovat dimenzí jistého vektorového prostoru.

Čínská věta o zbytcích, pro čísla i pro polynomy, má zásadní aplikace ve výpočetní algebře: úlohu pro jeden „velký objekt“ (velké číslo, polynom velkého stupně) umožňuje rozdělit na větší množství úloh s „malými objekty“. To je výhodné nejen pro paralelizaci, ale také v tom, že pro modulární aritmetiku jsou k dispozici lepší algoritmy. Čtenáře odkazujeme např. do skript Počítačová algebra.

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 4.9. 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) Zde je snadné spočítat *soudělná* čísla v intervalu $1, \dots, p$: jsou to právě čísla $p, 2p, 3p, \dots, p^{k-1} \cdot p$ a 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

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

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

(a) f zobrazuje množinu Φ_{ab} do množiny $\Phi_a \times \Phi_b$, tj. že $\mathrm{NSD}(x, ab) = 1$ implikuje $\mathrm{NSD}(x \bmod a, a) = 1 = \mathrm{NSD}(x \bmod b, b)$;

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

Obě části dokážeme zároveň: $\mathrm{NSD}(x, ab) = 1$ právě tehdy, když $\mathrm{NSD}(x, a) = 1 = \mathrm{NSD}(x, b)$ (uvažujte prvočíselného dělitele a použijte Lemma 4.5), což je právě tehdy, když $\mathrm{NSD}(x \bmod a, a) = 1 = \mathrm{NSD}(x \bmod b, b)$ použitím Lemmatu 4.3. □

5. Základní pojmy

V této sekci definujeme základní pojmy jako dělitelnost, asociovanost, největší společný dělitel a ireducibilní rozklady. Tyto pojmy definujeme pro oboecný obor integrity $\mathbf{R}$ a budeme je ilustrovat v následujících konkrétních případech: pro tělesa, pro obor $\mathbb{Z}$, pro obory polynomů a pro kvadratická rozšíření celých čísel.

5.1. **Dělitelnost a asociovanost.**

Řekneme, že a *dělí* b v oboru $\mathbf{R}$ a píšeme $a \mid b$, pokud existuje $c \in R$ takové, že $b = ac$. Pozor, při použití symbolu pro dělitelnost, i jakéhokoliv odvozeného pojmu, musíme vždy uvést (anebo musí být z kontextu zřejmé), v jakém oboru pracujeme! Například

- $3x + 6 \mid x + 2$ v oboru $\mathbb{Q}[x]$, protože $x + 2 = \frac{1}{3} \cdot (3x + 6)$; ale
- $3x + 6 \nmid x + 2$ v oboru $\mathbb{Z}[x]$, protože neexistuje $f \in \mathbb{Z}[x]$ splňující $x + 2 = f \cdot (3x + 6)$.

Řekneme, že prvky a a b jsou *asociované* a píšeme $a \parallel b$, pokud $a \mid b$ a $b \mid a$. Prvek a je invertibilní právě tehdy, když $a \parallel 1$.

Všimněte si, že relace dělitelnosti je reflexivní a tranzitivní: pokud $a \mid b$ a $b \mid c$, tedy pokud $b = ax$ a $c = by$ pro nějaká x, y, pak $c = axy$, tedy $a \mid c$. Z toho ihned plyne, že relace $\parallel$ je ekvivalencí.

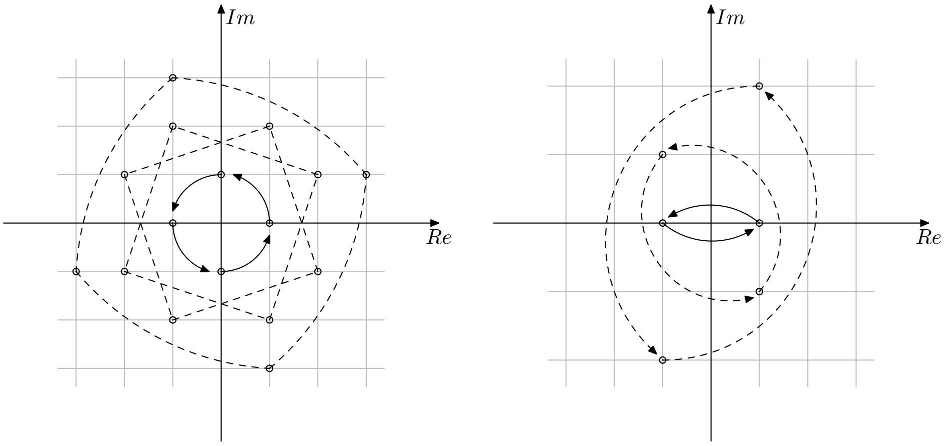
Tvrzení 5.1 (asociovanost vs. invertibilní prvky)**.** *Buď* $\mathbf{R}$ *obor integrity a* $a, b \in R$. *Pak* $a \parallel b$ *právě tehdy, když existuje invertibilní prvek* $q \in R$ *takový, že* $a = bq$.

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

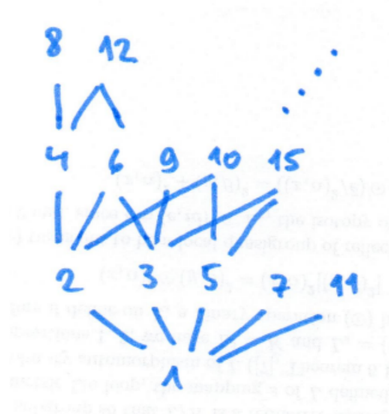
($\Rightarrow$) Pokud $a = 0$, pak $b = 0$ a tvrzení platí. Uvažujme $a \neq 0$. Protože $b \mid a$, můžeme psát $a = bu$, a protože $a \mid b$, můžeme psát $b = av$, pro nějaká u, v. Tedy $a = bu = avu$ a krácením dostáváme $uv = 1$, čili $u, v \parallel 1$. □

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 koeficient je invertibilní v oboru $\mathbf{R}$. Nad tělesem jsou tedy invertibilní všechny nenulové polynomy stupně 0. Například v $\mathbb{Z}[x]$ je $f \parallel g$ právě tehdy, když $g = \pm f$, zatímco v $\mathbb{Q}[x]$ je $f \parallel g$ právě tehdy, když $g = cf$ pro nějaké $0 \neq c \in \mathbb{Q}$.
- V oborech $\mathbb{Z}[\sqrt{s}]$ jsou invertibilní právě prvky normy 1, viz Tvrzení 3.3 a příklady pod ním. Z nich plyne (viz též obrázek 6), že
  - v oboru $\mathbb{Z}[i]$ jsou invertibilní právě prvky $\pm 1, \pm i$, a tedy $u \parallel v$ právě tehdy, když $u = \pm v$ nebo $u = \pm iv$;



Obrázek 6. Asociovanost v $\mathbb{Z}[i]$ a v $\mathbb{Z}[i\sqrt{2}]$.



Obrázek 7. Část Hasseova diagramu uspořádání dělitelností v oboru $\mathbb{Z}$.

– v oboru $\mathbb{Z}[i\sqrt{s}]$, $s > 1$, jsou invertibilní právě prvky ± 1, a tedy $u \parallel v$ právě tehdy, když $u = \pm v$.

K tomu, aby byla dělitelnost uspořádáním, 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 $\mathbb{Z}_2[x]$.) Ke grafickému znázornění vztahu dělitelnosti se používají tzv. *Hasseovy diagramy* (viz obrázek), kde jednotlivé třídy ekvivalence $\parallel$ jsou reprezentovány vrcholy a dělitelnost $a \mid b$ znázorňujeme tak, že vrchol odpovídající prvku b je výše než vrchol odpovídající prvku a a mezi těmito vrcholy vede hrana; hrany, jejichž existence plyne z tranzitivity, se vynechávají.

V obecných oborech integrity není možné definovat dělení se zbytkem, protože nemáme způsob, jak vyjádřit, že zbytek má být menší než dělitel. Přesto má smysl zavést symbol kongruence podmínkou

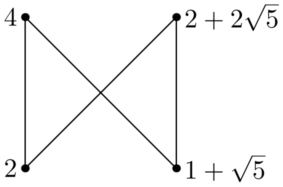
$$a \equiv b \pmod{m} \iff m \mid a - b.$$

Stejně jako pro čísla je snadné dokázat, že jde o ekvivalenci invariantní vzhledem k okruhovým operacím (důkaz Tvrzení 4.7 projde téměř doslova). S krácením je to složitější, k jeho důkazu jsme potřebovali prvočíselné rozklady; pohledem na důkaz je vidět, že analogie Tvrzení 4.8 platí v tzv. gaussovských oborech **R** (viz sekce 6).

5.2. 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);



Obrázek 8. V $\mathbb{Z}[\sqrt{5}]$ neexistuje $\mathrm{NSD}(4, 2+2\sqrt{5})$.

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

(Všimněte si, že definice $\mathrm{NSD}(a,b)$ odpovídá definici infima množiny $\{a,b\}$ vzhledem k relaci $|$.) Prvky a, b nazýváme *nesoudělné*, pokud $\mathrm{NSD}(a,b) = 1$.

Operátor NSN označující *nejmenší společný násobek* se definuje analogicky.

S definicí NSD je potřeba zacházet opatrně: hodnota c není určena jednoznačně. Například, v oboru $\mathbb{Z}$ bude platit $\mathrm{NSD}(4,6) = 2$, ale také $\mathrm{NSD}(4,6) = -2$, obě čísla splňují definici. Na NSD je potřeba nahlížet jako na relaci „c vyhovuje definici největšího společného dělitele prvků a, b“.

Situace naštěstí není tak špatná: NSD je určen jednoznačně *až na asociovanost*. Z jedné strany, pokud $\mathrm{NSD}(a,b) = c_1$ a $\mathrm{NSD}(a,b) = c_2$, pak c_1 i c_2 jsou společní dělitelé a, b, a tedy oba musí být největší, tj. $c_1 \mid c_2$ a zároveň $c_2 \mid c_1$, tedy $c_1 \parallel c_2$. Na druhou stranu, pokud $\mathrm{NSD}(a,b) = c_1$ a $c_1 \parallel c_2$, pak c_2 jistě také splňuje podmínky největšího společného dělitele.

Další problém spočívá v tom, že existence NSD není ničím garantovaná. A skutečně, jsou obory, v nichž NSD pro některé dvojice prvků neexistuje.

Příklad. Uvažujme obor $\mathbb{Z}[\sqrt{5}]$ a prvky

$$u = 4, \quad v = 2 + 2\sqrt{5}.$$

Čísla $r = 2$ a $s = 1 + \sqrt{5}$ jsou určitě společnými děliteli prvků u, v (viz obrázek 8), protože $u = 2 \cdot 2 = (1+\sqrt{5})(1-\sqrt{5})$ a $v = 2 \cdot (1+\sqrt{5})$. Přitom určitě $r \nmid s$ a $s \nmid r$. Uvažujme nějvětší společný dělitel $z = \mathrm{NSD}(u,v)$, tedy $r, s \mid z$ a $z \mid u, v$. Z Tvrzení 3.3 plyne, že $4 = \nu(r) = \nu(s) \mid \nu(z)$ a $\nu(z) \mid \nu(u) = \nu(v) = 16$, a protože jsou to vlastní dělitelé, z musí mít normu 8. Napíšeme-li $u = z \cdot w$, dostaneme $\nu(w) = 2$. Ale takový prvek v $\mathbb{Z}[\sqrt{5}]$ neexistuje: snadno ověříme, že číslo $\nu(a + b\sqrt{5}) = |a^2 - 5b^2|$ je buď liché (mají-li čísla a, b různou paritu), nebo dělitelné čtyřmi (mají-li stejnou paritu).

5.3. **Ireducibilní prvky a rozklady.**

Pro každý prvek a platí, že $1 \mid a$ a $a \mid a$. Dělitel prvku a se nazývá *vlastní*, jestliže není asociovaný ani s 1, ani s a.

Definice. Prvek a se nazývá *ireducibilní*, pokud $a \neq 0$, $a \not\parallel 1$ a a nemá vlastní dělitele. Jinými slovy, pokud pro každý rozklad $a = bc$ platí $b \parallel 1$ nebo $c \parallel 1$.

Příklad.

- V tělesech žádné ireducibilní prvky nejsou.
- V oboru $\mathbb{Z}$ jsou ireducibilní právě čísla $\pm p$, kde p je prvočíslo.

V oborech polynomů není obecně snadné určit, které polynomy jsou ireducibilní. V oboru $\mathbf{R}[x]$ jsou ireducibilní

- ty polynomy stupně 0, které jsou ireducibilní jako prvky $\mathbf{R}$;
- ty polynomy stupně 1, které nejsou dělitelné žádným neinvertibilním prvkem $\mathbf{R}$ (např. polynom $2x + 2$ je ireducibilní v $\mathbb{Q}[x]$, ale nikoliv v $\mathbb{Z}[x]$).

Je-li f polynom z $\mathbf{R}[x]$ stupně ≥ 2, který má kořen $a \in R$, pak nemůže být ireducibilní, protože má vlastního dělitele $x - a$ (Tvrzení 2.3). Pozor, opačná implikace neplatí: např. polynom $x^4 + 2x^2 + 1$ je rozložitelný v oboru $\mathbb{Z}[x]$, přestože tam nemá kořen. Pro polynomy vyšších stupňů není žádné obecné pravidlo.

Příklad.

- V oboru $\mathbb{C}[x]$ jsou ireducibilní právě polynomy stupně 1, jak praví základní věta algebry (Věta 12.1).
- V oboru $\mathbb{R}[x]$ jsou ireducibilní právě polynomy stupně 1 a ty polynomy stupně 2, které nemají reálný kořen (viz cvičení).
- V oboru $\mathbb{Q}[x]$ jsou ireducibilní i některé polynomy vyšších stupňů, např. všechny polynomy $x^n - 2$, jak plyne z Eisensteinova kritéria (Tvrzení 8.6).

K určení ireducibilních prvků v oborech $\mathbb{Z}[\sqrt{s}]$ lze použít Tvrzení 3.3. Je-li v vlastním dělitelem prvku u, pak $\nu(v) \mid \nu(u)$ a navíc $1 < \nu(v) < \nu(u)$. 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 $v \mid 3$, pak $\nu(v) = 3$. Označíme-li $v = a + bi$, hledáme a, b splňující $a^2 + b^2 = 3$, ale taková $a, b \in \mathbb{Z}$ neexistují.

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.

Ireducibilitu lze snadno prokázat pomocí Tvrzení 3.3. Důkaz, že ostatní prvky lze rozložit, je o něco složitější (viz cvičení).

Definice. *Ireducibilním rozkladem* prvku a 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. Řekneme, že prvek *a má jednoznačný ireducibilní rozklad*, pokud má právě jeden rozklad *až na pořadí a asociovanost*, tj. 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 například $12 = 2^2 \cdot 3^1 = 3^1 \cdot (-2)^2 \parallel (-2)^2 \cdot (-3)^1$. Formálně vzato, jde o tři různé rozklady, ale přesto je rozumné je považovat za „totožné“: 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, $-4 \parallel 2^2$.

Je důležité, v kterém oboru rozkládáme (nutno explicitně zmínit, nebo to musí být zřejmé z kontextu). Např. polynom $2x^2 + 2$ je ireducibilní v $\mathbb{Q}[x]$, ale má ireducibilní rozklad $2^1 \cdot (x^2+1)^1$ v $\mathbb{Z}[x]$, a dále třeba $(1+i)^2 \cdot (x^2+1)^1$ v $\mathbb{Z}[i][x]$.

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	$(x-\sqrt{2})(x+\sqrt{2})$	$(x^2+1)^2$
$\mathbb{C}[x]$	$(x-i)(x+i)$	$(2x-2i)(x+i)$	$(x-\sqrt{2})(x+\sqrt{2})$	$(x-i)^2(x+i)^2$
$\mathbb{Z}_5[x]$	$(x+2)(x+3)$	$(x+2)(2x+1)$	ireducibilní	$(x+2)^2(x+3)^2$

Existence ani jednoznačnost rozkladů není ničím garantovaná.

Příklad (obor bez rozkladů)**.** Uvažujme podokruh $\mathbf{R}$ oboru $\mathbb{Q}[x]$ sestávající z polynomů, jejichž absolutní člen je celé číslo. Polynom x není invertibilní, ale nemá ireducibilní rozklad: pokud $f \mid x$, tj. pokud existuje $g \in R$ takové, že $x = f \cdot g$, pak buď $f = a \in \mathbb{Z}$ a $g = \frac{1}{a}x$, nebo naopak. Přitom z těchto polynomů jsou ireducibilní pouze konstantní polynomy s prvočíselným koeficientem, protože každý polynom $\frac{1}{a}x$ má netriviální rozklad $\frac{1}{2a}x \cdot 2$. Součinem konstantních polynomů však nikdy nebude polynom x.

Příklad (obor s nejednoznačnými rozklady)**.** Uvažujme obor $\mathbb{Z}[\sqrt{5}]$. Prvek 4 lze rozložit dvěma způsoby:

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

Všechny tři prvky $2, \pm 1+\sqrt{5}$ jsou ireducibilní: jejich norma je 4 a jak jsme dokázali v sekci 5.2, žádné prvky normy 2 v $\mathbb{Z}[\sqrt{5}]$ nejsou. Přitom $2 \not\parallel \pm 1+\sqrt{5}$, protože všechny prvky dělitelné 2 mají sudé koeficienty.

5.4. **Prfvočinitelé.**

Na závěr definujeme pojem prvočinitele, který úzce souvisí s ireducibilitou a je motivovaný Lemmatem 4.5.

Definice. Prvek p splňující implikaci

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

se nazývá *prvočinitel.*

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. Opačná implikace v některých oborech platí (pro $\mathbb{Z}$ viz Lemma 4.5, pro gaussovské obory viz Důsledek 6.2), ale obecně ne.

Příklad. Uvažujme obor $\mathbb{Z}[\sqrt{5}]$. Prvek 2 je ireducibilní, $2 \mid (\sqrt{5}-1)(\sqrt{5}+1)$, ale přitom $2 \nmid (\sqrt{5} \pm 1)$, čili prvek 2 není prvočinitelem.

To, že jsou si příklady na neexistenci NSD, nejednoznačnost ireducibilního rozkladu a existenci ireducibilního neprvočinitele podobné, není náhoda. Podrobněji si to vysvětlíme v následující sekci.

6. Existence a jednoznačnost ireducibilních rozkladů

6.1. **Gaussovské obory.**

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

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 (Věta 4.2).
- Obory polynomů nad tělesem jsou gaussovské. Pro polynomy jedné proměnné funguje podobný důkaz jako pro celá čísla; obecná varianta tohoto důkazu je předmětem následujících dvou sekcí (Věty 6.3 a 7.3). Pro polynomy více proměnných nebo pro polynomy nad $\mathbb{Z}$ můžeme použít Gaussovu větu (Věta 8.4).
- Některé obory $\mathbb{Z}[\sqrt{s}]$ jsou gaussovské, např. pro $s = -1, \pm 2, 3$, některé ne, např. pro $s = -3, 5$. Na tyto případy se také vztahují Věty 6.3 a 7.3.

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

Tvrzení 6.1. *Buď* **R** *gaussovský obor, $a, b \in R$ a uvažujme ireducibilní rozklad*

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

Pak $b \mid a$ právě tehdy, když

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

pro nějaká $0 \le l_i \le k_i$.

Důkaz. ($\Leftarrow$) Pokud $a = qp_1^{k_1} \cdot \ldots \cdot p_n^{k_n}$ a $b = rp_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$ pro nějaké invertibilní prvky $q, r \in R$, definujme $c = qr^{-1}p_1^{k_1-l_1} \cdot \ldots \cdot p_n^{k_n-l_n}$ a vidíme, že $a = bc$, tedy $b \mid a$.

($\Rightarrow$) Uvažujme prvek c takový, že $a = b \cdot c$ a uvažujme ireducibilní rozklady

$$b \parallel q_1^{s_1} \cdot \ldots \cdot q_u^{s_u}, \qquad c \parallel r_1^{t_1} \cdot \ldots \cdot r_v^{t_v}.$$

Složením těchto dvou rozkladů (vynecháme ty prvky r_i, které jsou asociované s některým q_j a jejich exponenty sečteme) dostaneme druhý ireducibilní rozklad prvku a:

$$a \parallel p_1^{k_1} \cdot \ldots \cdot p_n^{k_n} \parallel q_1^{s_1'} \cdot \ldots \cdot q_u^{s_u'} r_{i_1}^{t_{i_1}} \cdot \ldots \cdot r_{i_w}^{t_{i_w}}.$$

Z jednoznačnosti rozkladů plyne, že ke každému $i = 1, \ldots, u$ existuje jednoznačně určené $j \in \{1, \ldots, n\}$ takové, že $q_i \parallel p_j$, přičemž $s_i \leq s_i' = k_j$. Z toho vyplývá, že $b \parallel p_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$ pro nějaká $0 \leq l_i \leq k_i$. $\square$

Snadným důsledkem Tvrzení 6.1 je několik vlastností dobře známých z celých čísel. V gaussovských oborech existují NSD: stačí srovnat rozklady obou prvků a vzít největší společný podrozklad, např.

$$\begin{aligned}\mathrm{NSD}(540, 336) &= \mathrm{NSD}((-2)^2 \cdot 3^3 \cdot 5,\ 2^4 \cdot (-3) \cdot 7) = \\ &= \mathrm{NSD}(2^2 \cdot 3^3 \cdot 5^1 \cdot 7^0,\ 2^4 \cdot 3^1 \cdot 5^0 \cdot 7^1) = 2^2 \cdot 3^1 \cdot 5^0 \cdot 7^0 = 12.\end{aligned}$$

V gaussovských oborech jsou ireducibilní prvky prvočiniteli: pokud ireducibilní prvek p dělí součin ab, pak jej musíme nalézt v rozkladu aspoň jednoho z činitelů. A žádný prvek nemá nekonečnou posloupnost vlastních dělitelů, neboť při každém dělení ztrácíme z rozkladu aspoň jeden činitel. Důkaz následujícího důsledku je formálním vyjádřením právě uvedených myšlenek.

Důsledek 6.2 (dělitelnost v gaussovských oborech)**.** *Buď* $\mathbf{R}$ *gaussovský obor. Pak*

(1) *pro každé* $a, b \in R$ *existuje* $\mathrm{NSD}(a, b)$*;*
(2) *každý ireducibilní prvek je prvočinitelem;*
(3) *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. (1) Uvažujme ireducibilní prvky $p_1, \ldots, p_n$, $p_i \nparallel p_j$ pro $i \neq j$, a $k_i, l_i \geq 0$ takové, že

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

(libovolné ireducibilní rozklady prvků a, b můžeme přepsat do této formy tak, že ze dvou asociovaných činitelů vybereme jeden a do rozkladu případně doplníme činitele v nulté mocnině.) Podle Tvrzení 6.1 $c \mid a, b$ právě tehdy, když $c \parallel p_1^{m_1} \cdot \ldots \cdot p_n^{m_n}$, kde $0 \leq m_i \leq k_i$ a $0 \leq m_i \leq l_i$, čili právě tehdy, když $0 \leq m_i \leq \min(k_i, l_i)$, pro všechna i. Největším z těchto společných dělitelů tedy bude ten, kde $m_i = \min(k_i, l_i)$.

(2) Uvažujme ireducibilní prvek p a prvky a, b takové, že $p \mid ab$. Podobně jako v (1), napišme $a \parallel p_1^{k_1} \cdot \ldots \cdot p_n^{k_n}$, $b \parallel p_1^{l_1} \cdot \ldots \cdot p_n^{l_n}$, kde $k_i, l_i \geq 0$, aspoň jeden nenulový. Pak $ab \parallel p_1^{k_1+l_1} \cdot \ldots \cdot p_n^{k_n+l_n}$ a podle Tvrzení 6.1 musí mít dělitel p rozklad, který obsahuje některé z prvků $p_1, \ldots p_n$. Z ireducibility plyne, že $p \parallel p_i$ pro některé i a tedy $p \mid a$ (pokud $k_i > 0$) nebo $p \mid b$ (pokud $l_i > 0$).

(3) Začneme obecnou úvahou. Každý nenulový neinvertibilní prvek a 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. Z jednoznačnosti rozkladů plyne, že číslo $\nu(a)$ je nezávislé na volbě rozkladu. Z Tvrzení 6.1 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. $\square$

Na závěr uveďme, že gaussovské obory nesdílejí všechny hezké vlastnosti oboru celých čísel: obecně nelze dělit se zbytkem a neplatí analogie Bézoutovy rovnosti.

Příklad. V oboru $\mathbb{Z}[x]$ neplatí Bézoutova rovnost. Platí $\mathrm{NSD}(x+1, x-1) = 1$, ale neexistují $f, g \in \mathbb{Z}[x]$ takové, že $f \cdot (x+1) + g \cdot (x-1) = 1$: pokud dosadíme číslo 1, vyjde nám $2f(1) = 1$, což není pro celočíselný polynom možné.

Příklad. V oboru $\mathbb{Q}[x,y]$ neplatí Bézoutova rovnost. Platí $\mathrm{NSD}(x,y)=1$, ale neexistují $f,g\in\mathbb{Q}[x,y]$ takové, že $f\cdot x+g\cdot y=1$, protože absolutní člen polynomu na levé straně je nutně nula.

6.2. **Zobecnění základní věty aritmetiky.**

Dostáváme se k slibované souvislosti ireducibilních rozkladů a existence NSD. Už víme, že v gaussovských oborech NSD existují. Opačná implikace bude sledovat důkaz základní věty artimetiky. Ten používal dvě základní ingredience: kromě existence NSD také matematickou indukci, vycházející z uspořádání přirozených čísel. Obecné obory integrity nemusí být dobře uspořádatelné, klasická indukce nám tedy nepomůže. Místo ní použijeme podmínku o neexistenci nekonečných posloupností vlastních dělitelů.

Věta 6.3 (zobecněná základní věta aritmetiky)**.** *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 Důsledku 6.2. Důkaz opačné implikace bude sledovat postup, kterým jsme dokázali základní větu aritmetiky v sekci 4.1. Opět jej rozdělíme do dvou částí: nejprve dokážeme existenci rozkladů, což je poměrně snadné, a pak se budeme věnovat jednoznačnosti, která vyžaduje jistá technická lemmata.

Důkaz existence rozkladů. Pro spor uvažujme prvek a, který nemá ireducibilní rozklad, $0\neq a\nparallel 1$. Rekurzí zkonstruujeme posloupnost, která protiřečí bodu (2).

- Položme $a_1=a$. Tedy $a_1\nparallel 1$ a nemá ireducibilní rozklad.
- Předpokládejme, že $a_i\nparallel 1$ a nemá ireducibilní rozklad. Speciálně, prvek a_i není sám ireducibilní, a tedy $a_i=b\cdot c$ pro nějaká $b,c\nparallel 1$. 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). □

K důkazu jednoznačnosti se nám bude hodit ještě jedna analogie Lemmatu 4.5, tentokrát dokázaná za předpokladu existence NSD. Protože obecně nemáme k dispozici Bézoutovu rovnost, budeme muset postupovat obezřetněji než v sekci 4.1.

Lemma 6.4. *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 $u=\mathrm{NSD}(ac,bc)$. Pro $c=0$ tvrzení platí triviálně, předpokládejme tedy $c\neq 0$.

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 6.5. *Uvažujme obor integrity, kde existují NSD všech dvojic prvků. Pak je každý ireducibilní prvek prvočinitelem.*

Důkaz. Buď p ireducibilní a a,b taková, že $p\mid ab$. Předpokládejme, že $p\nmid a$. Z ireducibility plyne, že $\mathrm{NSD}(a,p)=1$, a tedy podle Lemmatu 6.4

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

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

Důkaz jednoznačnosti rozkladů. Sporem. Mezi všemi prvky s různými ireducibilními rozklady 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 6.5 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, spor. □

Věta 6.3 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 **R**. Naopak druhou stranu charakterizace lze formulovat čistě v jazyce relace dělitelnosti: podmínka (1) říká, že vzhledem k „uspořádání“ $\mid$ existují „infima“ všech dvouprvkových množin, a podmínka (2) říká, že v něm neexistuje nekonečný ostře klesající řetězec.

6.3. **Řešení diofantických rovnic pomocí rozkladu v rozšíření.**

Jednou z původních motivací k vytvoření abstraktní teorie dělitelnosti bylo řešení *diofantických rovnic*, tj. polynomiálních rovnic v oboru celých čísel. Asi nejznámější takovou rovnici řeší velká Fermatova věta, tedy tvrzení, že pro žádné $n \geq 3$ neexistují nenulová celá čísla x, y, z splňující $x^n + y^n = z^n$. Historie tohoto problému je dávná: Pythagoras řešil případ $n = 2$ již v 6. století před letopočtem (pravoúhlé trojúhelníky s celočíselnými hranami) a obecná rovnice byla na stole přinejmenším od dob Pierra Fermata (17. století), který, jak známo, prohlásil, že žádné řešení neexistuje, ale své prohlášení nepodložil důkazem. Leonhard Euler v roce 1753 prokázal neexistenci nenulového řešení pro $n = 3$ pomocí počítání v Eisensteinových číslech. Jeho metodu rozvíjeli další matematici, na hranici jejích možností ji dotáhl Ernst Kummer v polovině 19. století, kdy dokázal neexistenci nenulových řešení pro všechna regulární prvočísla.

Základní myšlenka této metody je, že obě strany dané rovnice rozložíme a srovnáním ireducibilních rozkladů dojdeme k nějaké malé množině řešení (zkuste si takto vyřešit rovnice $x^2 - y^2 = 5$ a $x^2 + 3x + 2 = y^3$). Přitom rozklad nemusíme hledat v oboru $\mathbb{Z}$, ale v libovolném gaussovském nadoboru, ve kterém pak budeme provádět všechny úvahy. Metodu si ukážeme na konkrétní rovnici $x^2 + 1 = y^3$. Řešení využívá fakt, že je obor $\mathbb{Z}[i]$ gaussovský, což je okamžitým důsledkem Vět 3.4 a 7.3.

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 $x - i$, čili 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 $\mathbb{Z}[i]$ (třetí mocnina má všechny činitele ireducibilního rozkladu s exponentem dělitelným 3; přitom rozklady $x+i$, $x-i$ používají disjunktní sady ireducibilních prvků, čili samy musí být třetí mocninou). Uvažujme takové $a+bi$: z rovnosti $(a + bi)^3 = (a^3 - 3ab^2) + (3a^2b - b^3)i = x + i$ plyne $b(3a^2 - b^2) = 1$, což má jediné celočíselné řešení $b = -1$, $a = 0$. To dává jediné celočíselné řešení původní rovnice $x = 0$, $y = 1$. □

Pro Fermatovu rovnici se hodí rozklad

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

v cyklotomickém rozšíření $\mathbb{Z}[\zeta_n]$ a dojde se ke sporu s tím, že by tento rozklad měl být n-tou mocninou nějakého čísla. Wilesův důkaz velké Fermatovy věty z roku 1995 nakonec využil úplně jinou metodu, která převádí řešení jistého typu rovnic na problémy o eliptických křivkách a modulárních formách, ale to už je jiná historka.

7. Eukleidův algoritmus a Bézoutova rovnost

7.1. **Eukleidovské obory.**

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

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

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. Všimněte si, že $\nu(a) = 0$ právě tehdy, když $a = 0$: zbytek po dělení jakýmkoliv nenulovým prvkem musí mít menší normu, čili norma dělitele nemůže být 0.

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 $\mathbf{T}[x]$ je eukleidovský pro libovolné těleso $\mathbf{T}$. Normou je

  $$\nu(f) = 1 + \deg f.$$

  Vlastnost (1) je zřejmá a vlastnost (2) plyne z Tvrzení 2.2. (Norma je nezáporná, takže ke stupni musíme přičíst 1.)
- 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

  $$\nu(a + b\sqrt{s}) = |a^2 - sb^2|.$$

  Vlastnost (1) platí pro každé s díky Tvrzení 3.3. Vlastnost (2) pro Gaussova celá čísla plyne z Tvrzení 3.4.

Existují gaussovské obory, které nejsou eukleidovské, například obor $\mathbb{Z}[x]$ nebo obory polynomů více proměnných (i nad tělesem). Rozebereme případ oboru $\mathbb{Z}[x]$. Zobrazení $\nu(f) = 1 + \deg f$ není eukleidovskou normou: například 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 uvedeného 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. Naštěstí, tento fakt ihned plyne z Věty 7.1: v eukleidovských oborech platí Bézoutova rovnost, ale $\mathbb{Z}[x]$ ne, jak jsme si ukázali v minulé sekci.

Primárním důsledkem eukleidovské normy je Eukleidův algoritmus na výpočet NSD a Bézoutových koeficientů.

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 $\mathrm{NSD}(a,b) = u \cdot a + v \cdot b$.
- $a_0 = a, \quad u_0 = 1, \quad v_0 = 0.$
  $a_1 = b, \quad u_1 = 0, \quad v_1 = 1.$
- pro $i = 2, 3, ...$ prováděj následující:
  zvol q, r tak, aby $a_{i-1} = a_i q + r$ a $\nu(r) < \nu(a_i)$, a definuj
  $$a_{i+1} = r, \quad u_{i+1} = u_{i-1} - u_i q, \quad v_{i+1} = v_{i-1} - v_i q$$
  pokud $a_{i+1} = 0$, odpověz a_i, u_i, v_i

Věta 7.1 (správnost Eukleidova algoritmu)**.** *V eukleidovském oboru* $\mathbf{R}$ *najde Eukleidův algoritmus pro jakýkoliv vstup* $a, b \in R$ *hodnotu* $\mathrm{NSD}(a,b)$ *a tzv.* Bézoutovy koeficienty $u, v \in R$ *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 n číslo kroku, ve kterém se tak stane. Stačí dokázat následující dvě vlastnosti:

(1) NSD dvou po sobě jdoucích prvků se nemění, tj. pro každé $i = 1, \ldots, n$ platí $\mathrm{NSD}(a_{i-1}, a_i) = \mathrm{NSD}(a_i, a_{i+1})$;
(2) pro každé $i = 0, \ldots, n$ platí $a_i = u_i \cdot a + v_i \cdot b$.

Vzhledem k tomu, že $\mathrm{NSD}(u, 0) = u$ pro každé u, algoritmus správně odpoví
$$a_n = \mathrm{NSD}(a_n, 0) = \mathrm{NSD}(a_{n-1}, a_n) = \ldots = \mathrm{NSD}(a_0, a_1) = \mathrm{NSD}(a, b).$$
Obě vlastnosti 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 4.3). Indukcí ověříme (2). Pro $i = 0, 1$ výrok platí z definice. 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 pomocí Eukleidova algoritmu. Dělením 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. □

Řešení pomocí rozkladů na ireducibilní prvky. Platí $\nu(3 + 11i) = 130 = 2 \cdot 5 \cdot 13$ a $\nu(-2 + 9i) = 85 = 5 \cdot 17$, čili případný společný dělitel by měl normu 5. Jediné dva prvky normy 5, až na asociovanost, jsou $2 \pm i$. Snadno vyzkoušíme, že $2 + i \nmid 3 + 11i$, $2 - i \nmid -2 + 9i$, čili uvedená čísla jsou nesoudělná. □

Nyní již snadno dokážeme, že v eukleidovských oborech má každý prvek jednoznačný ireducibilní rozklad.

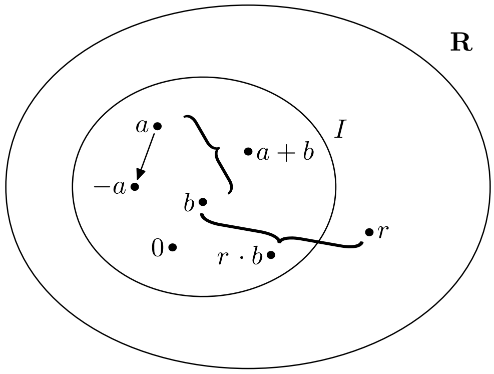
Lemma 7.2. *Buď* $\mathbf{R}$ *eukleidovský obor a* $a, b \in R$, $a, b \neq 0$. *Pokud* $a \mid b$ *a* $a \nparallel b$, *pak* $\nu(a) < \nu(b)$.

Důkaz. Napišme

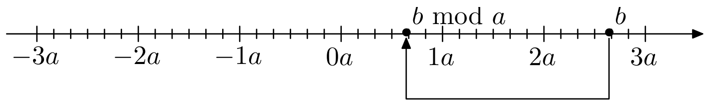
- $b = au$ pro nějaké $u \in R$,
- $a = bq + r$ pro nějaká $q, r \in R$, $\nu(r) < \nu(b)$.

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 $\nu(a) \leq \nu(r) < \nu(b)$. □

Věta 7.3. *Eukleidovské obory jsou gaussovské.*



OBRÁZEK 9. Ideál I v oboru $\mathbf{R}$.



OBRÁZEK 10. Ilustrace důkazu Věty 7.5 v případě $\mathbf{R} = \mathbb{Z}$.

Důkaz. Podle Věty 6.3 stačí dokázat, že v eukleidovských oborech existují NSD a neexistují nekonečné posloupnosti vlastních dělitelů. První fakt jsme dokázali ve Větě 7.1. Druhý plyne bezprostředně z předešlého lemmatu: taková posloupnost by měla ostře klesající normu, což nelze. □

Důsledkem Věty 7.3 je, že Gaussova celá čísla či obory polynomů nad tělesem jsou gaussovské.

7.2. **Obory hlavních ideálů.**

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

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

Příklad. Množiny $n\mathbb{Z} = \{nz : \ z \in \mathbb{Z}\} = \{u \in \mathbb{Z} : \ n \mid u\}$ jsou ideály v oboru $\mathbb{Z}$. Z Věty 7.5 plyne, že žádné jiné ideály v oboru $\mathbb{Z}$ nejsou.

Konstrukci ideálů z předchozího příkladu lze zobecnit.

Tvrzení 7.4 (definice hlavních ideálů). *Buď $\mathbf{R}$ komutativní okruh a $a \in R$. Pak*

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

je ideál v $\mathbf{R}$. Je to nejmenší ideál (nejmenší vzhledem k inkluzi) obsahující prvek a.

Důkaz. Součet a rozdíl dvou prvků dělitelných a je dělitelný a, a pokud $a \mid u$, pak $a \mid ru$ pro libovolné $r \in R$. Čili aR je ideál.

Buď I libovolný ideál obsahující prvek a. Pak I jistě obsahuje i všechny jeho násobky, čili $aR \subseteq I$, a tedy aR je nejmenším ideálem obsahujícím prvek a. □

Definice. Ideály z Tvrzení 7.4 se nazývají *hlavní*. Speciálně, $\{0\} = 0R$ a $R = 1R$ jsou hlavní ideály v libovolném komutativním okruhu; říká se jim *nevlastní*.

Hlavní ideály pěkně odrážejí dělitelnost: z tranzitivity relace dělitelnosti ihned plyne, že

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

Co jiné příklady ideálů? Následující věta říká, že se budeme muset poohlédnout jinde než v $\mathbb{Z}$ či v $\mathbf{T}[x]$.

Věta 7.5. *V eukleidovských oborech je každý ideál hlavní.*

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. □

Okamžitým důsledkem je fakt, že tělesa jsou právě ty komutativní okruhy, které nemají žádné vlastní ideály. Využijeme jej v sekci 20.3 při konstrukci těles jako faktorokruhů.

Tvrzení 7.6 (ideály v tělesech)**.** *Buď* $\mathbf{R}$ *komutativní okruh s jednotkou. Pak* $\mathbf{R}$ *je těleso právě tehdy, když má pouze nevlastní ideály.*

Důkaz. ($\Rightarrow$) Tělesa jsou eukleidovské obory, tedy každý ideál je hlavní. Pro každé $a \neq 0$ platí $a \parallel 1$, čili pro každý ideál aR platí $aR = 1R = R$.

($\Leftarrow$) Pro každý hlavní ideál aR, $a \neq 0$, platí $aR = R = 1R$, čili každý nenulový prvek a je invertibilní. □

Definice. Komutativní okruhy, které neobsahují jiné ideály než hlavní, nazýváme *okruhy hlavních ideálů*; v případě oborů integrity hovoříme o *oborech hlavních ideálů*.

Věta 7.5 tak říká že eukleidovské obory jsou obory hlavních ideálů. Opačná implikace neplatí, ale najít nějaký příklad není snadné. Asi nejjednodušším příkladem neeukleidovského oboru hlavních ideálů je $\mathbb{Z}[\frac{1+i\sqrt{19}}{2}]$ a důkaz tohoto faktu je poměrně obtížný.

V Sekci 7.1 jsme ukázali, že obory $\mathbb{Z}[x]$ ani obory polynomů více proměnných nejsou eukleidovské, protože v nich neplatí Bézoutova rovnost. Ukážeme, že v nich také existuje ideál, který není hlavní. Oba příklady jsou založené na následující myšlence. Je-li aR hlavní ideál, který obsahuje dva nesoudělné prvky u, v, pak $aR = R$: z $u, v \in aR$ plyne $a \mid u$ i $a \mid v$, čili $a \parallel 1$, a tedy $aR = R$.

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

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

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

Nyní si dokážeme jedno pomocné tvrzení o obecných ideálech.

Tvrzení 7.7 (průnik, součet a sjednocení ideálů)**.** *Buď* $\mathbf{R}$ *komutativní okruh.*

1. *Jsou-li* I, J *ideály v* $\mathbf{R}$*, pak* $I \cap J$ *je také ideál v* $\mathbf{R}$.
2. *Jsou-li* I, J *ideály v* $\mathbf{R}$*, pak* $I + J = \{a + b : \ a \in I, b \in J\}$ *je také ideál v* $\mathbf{R}$*. Tento ideál je nejmenší takový, že obsahuje* $I \cup J$.
3. *Jsou-li* I_j*,* $j \in \mathbb{N}$*, ideály v* $\mathbf{R}$ *takové, že* $I_1 \subseteq I_2 \subseteq I_3 \subseteq \dots$*, pak* $\bigcup_{j \in \mathbb{N}} I_j$ *je také ideál v* $\mathbf{R}$.

Důkaz. (1) Buď $a, b \in I \cap J$ a $r \in R$. Pak $a + b, -a, ra$ náleží do obou ideálů I, J, tedy i do jejich průniku.

(2) Buď $a + b, c + d \in I + J$, přičemž $a, c \in I$ a $b, d \in J$, a buď $r \in R$. Pak $(a + b) + (c + d) = (a + c) + (b + d) \in I + J$, $-(a + b) = (-a) + (-b) \in I + J$ a $r(a + b) = ra + rb \in I + J$. Oba ideály I, J jsou podmnožinou $I + J$ a naopak, pokud ideál K obsahuje I i J, pak jistě obsahuje i všechny součty prvků z I a prvků z J, tedy $I + J \subseteq K$.

(3) Buď $a, b \in \bigcup_{j \in \mathbb{N}} I_j$ a $r \in R$. Pak existují $j, k \in \mathbb{N}$ taková, že $a \in I_j$ a $b \in I_k$, čili $a, b \in I_{\max(j,k)}$, a tedy $a + b, -a, ra \in I_{\max(j,k)} \subseteq \bigcup_{j \in \mathbb{N}} I_j$. □

Vztaženo na hlavní ideály, nejmenším ideálem obsahujícím dva prvky a, b je ideál

$$aR + bR = \{ar + bs : \ r, s \in R\},$$

a dále indukcí, nejmenším ideálem obsahujícím prvky $a_1, \dots, a_n$, tzv. *ideál generovaný prvky* $a_1, \dots, a_n$, je

$$a_1R + \dots + a_nR = \{\sum a_i r_i : \ r_1, \dots, r_n \in R.\}$$

Těmto prvkům se také říká *báze ideálu* (bez nároku na nezávislost, jak je zvykem v lineární algebře). Výše uvedené nehlavní ideály pak můžeme napsat jako $2\mathbb{Z}[x] + x\mathbb{Z}[x]$, resp. $x_1\mathbf{R}[x_1, \dots, x_k] + \ldots + x_k\mathbf{R}[x_1, \dots, x_k]$. Hilbertova větu o bázi nicméně říká, že v oborech polynomů nad $\mathbb{Z}$ či nad tělesy má každý ideál konečnou bázi.

Nyní již můžeme dokázat, jak se obory hlavních ideálů zařazují do hierarchie oborů z hlediska teorie dělitelnosti.

Věta 7.8. *Obory hlavních ideálů jsou gaussovské a platí v nich Bézoutova rovnost.*

Důkaz. Buď $\mathbf{R}$ obor hlavních ideálů. Podle Věty 6.3 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 $I = aR + bR$. Každý ideál je hlavní, existuje tedy $c \in R$ takové, že $I = cR$. Protože $aR, bR \subseteq cR$, máme $c \mid a$ i $c \mid b$. Dále, pokud je d společným dělitelem a, b, pak $aR \subseteq dR$ a $bR \subseteq dR$, tedy $I = cR \subseteq dR$ a dostáváme $d \mid c$. Vidíme, že $c = \mathrm{NSD}(a, b)$ a navíc $c \in aR + bR$, tedy $c = ar + bs$ pro nějaká $r, s \in R$, což je Bézoutova rovnost.

(2) Pro spor předpokládejme, že v $\mathbf{R}$ existuje nekonečná posloupnost vlastních dělitelů $a_1, a_2, \dots$ (tj. $a_{i+1} \mid a_i$ a $a_i \nmid a_{i+1}$). Pak $a_1R \subset a_2R \subset a_3R \subset \dots$ a označme $I = \bigcup_{i \in \mathbb{N}} a_iR$. Tato množina také tvoří ideál, takže $I = bR$ pro nějaké $b \in I$. Ovšem toto b musí být prvkem nějakého a_iR, pro nějaké $i \in R$. Ale pak $bR \subseteq a_iR \subset a_{i+1}R \subset \dots \subset I = bR$, spor. □

Na závěr uvedeme historickou motivaci pojmu ideál. Myšlenku lze nastínit následujícím způsobem: obor $\mathbb{Z}[\sqrt{5}]$ není gaussovský, protipříkladem je např. rozklad $4 = 2^2 = (1+\sqrt{5})(-1+\sqrt{5})$. Kdyby ovšem existovaly nějaké „ideální ireducibilní prvky“ (ideální ve smyslu hypotetické) p, q takové, že $2 = pq$, $1 + \sqrt{5} = p^2$ a $-1 + \sqrt{5} = q^2$, najednou bychom měli tentýž rozklad. Těmito „ideálními prvky“ se nakonec ukázaly být tzv. *prvoideály*, tj. ideály splňující dodatečnou podmínku „$ab \in I \Rightarrow a \in I$ nebo $b \in I$“. Moderní algebraická teorie čísel pak vychází z poznatku, že v mnoha oborech, včetně $\mathbb{Z}[\sqrt{5}]$, lze každý ideál rozložit jednoznačně na součin prvoideálů. Více viz libovolná učebnice algebraické teorie čísel.

7.3. Hierarchie oborů integrity.

Ve Větách 7.3, 7.5 a 7.8 jsme dokázali následující hierarchii oborů integrity:

$$\textit{eukleidovský obor} \Longrightarrow \textit{obor hlavních ideálů} \Longrightarrow \textit{gaussovský obor}$$

Základní vlastnosti těchto tříd jsou shrnuty v následující tabulce:

obory	ireducibilní rozklady	existence NSD	Bézoutova rovnost	Eukleidův algoritmus
eukleidovské	✓	✓	✓	✓
hlavních ideálů	✓	✓	✓	×
gaussovské	✓	✓	×	×
obecné	×	×	×	×

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

Algebra polynomů

8. Polynomy nad gaussovskými obory

8.1. Gaussova věta.

Polynomy jedné proměnné nad tělesem tvoří eukleidovský obor, a podle Věty 7.3 jsou gaussovské. Polynomy více proměnných, nebo třeba polynomy nad $\mathbb{Z}$, eukleidovské nejsou. K důkazu jejich gaussovskosti potřebujeme jinou techniku: dělitelnost polynomů nad oborem $\mathbf{R}$ úzce souvisí s dělitelností nad jeho podílovým tělesem $\mathbf{Q}$, viz Lemma 8.2 a Věta 8.3.

Definice. Polynom f z $\mathbf{R}[x]$ se nazývá *primitivní*, pokud jsou jeho koeficienty nesoudělné, tj. kdykoliv nějaký prvek $c \in R$ dělí všechny jeho koeficienty, pak $c \parallel 1$.

Například $x^2y + x$ není primitivní v oboru $(\mathbb{Z}[x])[y]$ (máme netriviálního dělitele x), ale je primitivní v oboru $(\mathbb{Z}[y])[x]$.

Matematický princip vztahu mezi dělitelností v $\mathbf{R}[x]$ a $\mathbf{Q}[x]$ je obsažen v následujícím tvrzení, známém jako Gaussovo lemma, které říká, že součin primitivních polynomů je primitivní, ovšem pouze za předpokladu, že je obor $\mathbf{R}$ gaussovský.

Lemma 8.1 (Gaussovo lemma)**.** *Buď $\mathbf{R}$ gaussovský obor a f, g primitivní polynomy z $\mathbf{R}[x]$. Pak součin fg je také 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. Díky existenci ireducibilních rozkladů existuje ireducibilní prvek $p \in R$, který dělí všechny koeficienty součinu fg. Zvolme nejmenší j takové, že $p \nmid a_j$, a nejmenší k takové, že $p \nmid b_k$ (protože jsou polynomy f, g primitivní, p 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 $p \mid a_i$ pro všechna $i < j$, máme

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

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

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

Tedy p dělí všechny členy součtu vlevo i vpravo od $a_j b_k$. Tento člen naopak p dělitelný není, protože p je ireducibilní, tedy prvočinitel (Důsledek 6.2), a přitom nedělí ani a_j, ani b_k. Čili $p \nmid c_{j+k}$, spor. □

Důsledkem je slibovaná souvislost dělitelnosti nad oborem a nad jeho podílovým tělesem.

Lemma 8.2 (dělitelnost v oboru vs. podílovém tělese)**.** *Buď $\mathbf{R}$ gaussovský obor, $\mathbf{Q}$ jeho podílové těleso a f, g primitivní polynomy z $\mathbf{R}[x]$. Pak $f \mid g$ v $\mathbf{R}[x]$ právě tehdy, když $f \mid g$ v $\mathbf{Q}[x]$.*

Důkaz. $f \mid g$ v $\mathbf{R}[x]$ znamená, že existuje $h \in R[x]$ splňující $g = fh$. Podobně, $f \mid g$ v $\mathbf{Q}[x]$ znamená, že existuje $h \in Q[x]$ splňující $g = fh$. Čili implikace $(\Rightarrow)$ je triviální a musíme dokázat tu opačnou. Mějme takový polynom $h \in Q[x]$ a zvolme $q \in Q$ tak, aby byl qh primitivní polynom z $\mathbf{R}[x]$ (stačí vzít $q = \frac{a}{b} \in Q$, kde a je NSN jmenovatelů všech koeficientů, a b je NSD čitatelů všech koeficientů polynomu h). Dostáváme $qg = f \cdot qh$. Na pravé straně je součin primitivních polynomů z $\mathbf{R}[x]$, takže podle Gaussova lemmatu 8.1 je qg také primitivní polynom z $\mathbf{R}[x]$. Protože je g primitivní, jmenovatel q musí být invertibilní. Protože je qg primitivní, čitatel q musí být také invertibilní. Čili q je invertibilní prvek $\mathbf{R}$, a tedy polynom h musel náležet $\mathbf{R}[x]$. □

Pro účely následující věty nám bude hodit následující značení. Pro polynom $f = \sum_{i=0}^{n} a_i x^i \neq 0$ definujeme

$$\mathrm{c}(f) = \mathrm{NSD}(a_0, \dots, a_n) \qquad \text{a} \qquad \mathrm{pp}(f) = \frac{1}{\mathrm{c}(f)} \cdot f$$

(pokud uvedený NSD neexistuje, pak $\mathrm{c}(f)$ a $\mathrm{pp}(f)$ nejsou definovány), hovoříme o *obsahu* a *primitivní části* polynomu f. Všimněte si, že polynom f je primitivní právě tehdy, když $\mathrm{c}(f) = 1$. Polynom $\mathrm{pp}(f)$ je vždy primitivní.

Věta 8.3 (NSD a ireducibilita v oboru vs. podílovém tělese)**.** *Buď* $\mathbf{R}$ *gaussovský obor,* $\mathbf{Q}$ *jeho podílové těleso a* f, g *polynomy z* $\mathbf{R}[x]$*. Pak*

(1) $\mathrm{NSD}_{\mathbf{R}[x]}(f, g)$ *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* $h = \mathrm{NSD}_{\mathbf{Q}[x]}(\mathrm{pp}(f), \mathrm{pp}(g))$.

(2) 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* $\mathbf{Q}[x]$*.*

Předně je třeba vyjasnit, proč je formulace bodu (1) tak složitá, proč nemůžeme rovnou psát vzorec ve tvaru

$$\mathrm{NSD}_{\mathbf{R}[x]}(f, g) = \mathrm{NSD}_{\mathbf{R}}(\mathrm{c}(f), \mathrm{c}(g)) \cdot \mathrm{NSD}_{\mathbf{Q}[x]}(\mathrm{pp}(f), \mathrm{pp}(g)).$$

Je to kvůli nejednoznačnosti operátoru NSD. Následující tvrzení jsou platná v $\mathbb{Q}[x]$: $\mathrm{NSD}_{\mathbb{Q}[x]}(x^2 + 2x + 1, x^2 - 1) = 2x + 2$, $\mathrm{NSD}_{\mathbb{Q}[x]}(x^2 + 2x + 1, x^2 - 1) = \frac{3}{4}x + \frac{3}{4}$. První odpověď nemůžeme v $\mathbb{Z}[x]$ použít, protože výsledek je dělitelný 2, ale ani jeden z polynomů číslem 2 dělitelný není. Druhou odpověď nemůžeme použít, protože výsledek ani neleží v $\mathbb{Z}[x]$. Věta říká, že použít máme primitivní výsledek, tj. v našem případě $x + 1$ nebo $-x - 1$. Takový polynom jistě existuje: stačí vzít libovolný $h = \mathrm{NSD}_{\mathbb{Q}[x]}(f, g)$ a přenásobit jej vhodným zlomkem, podobně jako v důkazu Lemmatu 8.2.

Důkaz Věty 8.3. (1) Nejprve dokážeme, že pro primitivní polynomy f, g z $\mathbf{R}[x]$ existuje $\mathrm{NSD}_{\mathbf{R}[x]}(f, g)$ a je roven primitivnímu polynomu h z $\mathbf{R}[x]$ splňujícímu $h = \mathrm{NSD}_{\mathbf{Q}[x]}(f, g)$. Polynom h dělí f, g v $\mathbf{Q}[x]$ a je primitivní, tedy díky Lemmatu 8.2 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, $d \mid f, g$ v $\mathbf{Q}[x]$, tedy $d \mid h$ v $\mathbf{Q}[x]$, a opět podle Lemmatu 8.2 $d \mid h$ v $\mathbf{R}[x]$.

Nyní odvodíme obecný vztah. Protože $c = \mathrm{NSD}_{\mathbf{R}}(\mathrm{c}(f), \mathrm{c}(g))$ dělí $\mathrm{c}(f)$ i $\mathrm{c}(g)$, a zároveň $h = \mathrm{NSD}_{\mathbf{R}[x]}(\mathrm{pp}(f), \mathrm{pp}(g))$ dělí $\mathrm{pp}(f)$ i $\mathrm{pp}(g)$, tak jejich součin ch dělí oba polynomy f, g, čili ch je společný dělitel. Dokážeme, že je to největší společný dělitel: pokud nějaký d dělí f i g, pak $\mathrm{c}(d)$ dělí $\mathrm{c}(f)$ i $\mathrm{c}(g)$, tedy $\mathrm{c}(d) \mid c$; analogicky $\mathrm{pp}(d) \mid h$ a dostáváme $d \mid ch$.

(2) Rozložme $f = c(f) \cdot pp(f)$. Je-li f ireducibilní, pak $c(f) \parallel 1$ nebo $pp(f) \parallel 1$. V druhém případě je f konstantní a musí být ireducibilní v $\mathbf{R}$. V prvním případě je f primitivní. Zbývá si uvědomit, že primitivní polynom je ireducibilní v $\mathbf{R}[x]$ právě tehdy, když v $\mathbf{Q}[x]$. Pokud by měl v $\mathbf{Q}[x]$ vlastního dělitele g, pak uvažujme $q \in Q$ takové, že qg je primitivní polynom z $\mathbf{R}[x]$, a tento bude díky Lemmatu 8.2 vlastním dělitelem v $\mathbf{R}[x]$. □

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 $\mathrm{NSD}_{\mathbb{Z}[x]}(f, g) = 2 \cdot (x + 1)$.

Příklad. Z bodu (2) plyne, že primitivní polynom je ireducibilní v $\mathbf{R}[x]$ právě tehdy, když v $\mathbf{Q}[x]$, ale obecně to neplatí:

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

Z bodu (2) ihned plyne existence ireducibilních rozkladů v $\mathbf{R}[x]$, ale s jednoznačností to je složitější, k ní potřebujeme mašinerii zobecněné základní věty aritmetiky.

Věta 8.4 (Gaussova věta)**.** *Je-li* $\mathbf{R}$ *gaussovský obor, pak je* $\mathbf{R}[x]$ *také gaussovský obor.*

Důkaz. Použijeme charakterizaci z Věty 6.3. Existenci NSD v $\mathbf{R}[x]$ jsme dokázali ve Větě 8.3. Buď $f_1, f_2, f_3, \ldots$ nekonečná posloupnost vlastních dělitelů v $\mathbf{R}[x]$. 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 u_i vedoucí koeficient polynomu f_i, pak $u_n, u_{n+1}, u_{n+2} \ldots$ tvoří nekonečnou 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 $\mathbf{R}[x_1, \ldots, x_n] = (\mathbf{R}[x_1, \ldots, x_{n-1}])[x_n]$.

8.2. **Racionální kořeny a Eisensteinovo kritérium.**

Připomeňte si důkaz Gaussova lemmatu (Lemma 8.1). Na podobném principu jsou založena dvě užitečná kritéria, jedno na existenci racionálních kořenů, druhé na ireducibilitu.

Tvrzení 8.5 (kritérium existence racionálního kořene)**.** *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 $a_1 r s^{n-1}, \ldots, a_n r^n$ i pravou stranu, musí dělit i první člen $a_0 s^n$. Protože jsou r, s nesoudělné, musí r dělit a_0 (zde využíváme gaussovskost, konkrétně Důsledek 6.2(2) aplikovaný na všechny ireducibilní prvky 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 8.5 jsou jedinými kandidáty čísla ± 1, ± 3, $\pm\frac{1}{2}$ a $\pm\frac{3}{2}$. Dosazením zjistíme, že vyhovuje pouze číslo $-\frac{3}{2}$.

Příklad. Racionálními kořeny polynomu $x^n - p$, p prvočíslo, mohou být pouze čísla $\pm 1, \pm p$ a ani jedno očividně nevyhovuje (pro $n \geq 2$). Důsledkem je, že všechny odmocniny prvočísel $\sqrt[n]{p}$ jsou iracionální.

Tvrzení 8.6 (Eisensteinovo kritérium)**.** *Buď* $\mathbf{R}$ *obor integrity a* $f = \sum_{i=0}^{n} a_i x^i$ *primitivní polynom z* $\mathbf{R}[x]$*. Pokud existuje prvočinitel* $p \in R$ *splňující* $p \mid a_0$, $p \mid a_1$, $\ldots$, $p \mid a_{n-1}$ *a* $p^2 \nmid a_0$*, pak je polynom* f *ireducibilní v* $\mathbf{R}[x]$*.*

Připomeňme, že podle Důsledku 6.2(2) jsou v gaussovských oborech prvočinitelé totéž, co ireducibilní prvky.

Důkaz. Pro spor 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 prvočinitel p dělí $a_0 = b_0 c_0$, 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. Podobně, protože $p \mid a_1 = b_0 c_1 + b_1 c_0$ a $p \nmid c_0$, 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 $p \mid b_2$. Postupně zjistíme, že p dělí všechny koeficienty b_i, tedy $p \mid g \mid f$, což je spor s primitivitou. □

Příklad. Z Eisensteinova kritéria plyne ireducibilita polynomů $x^n \pm a$ v oboru $\mathbb{Z}[x]$ pro každé a takové, že existuje prvočíslo p splňující $p \mid a$, $p^2 \nmid a$.

9. Počítání modulo polynom

9.1. **Čínská věta o zbytcích a interpolace.**

Čínská věta o zbytcích hovoří o tom, jak vypadají řešení soustav lineárních kongruencí. V sekci 4.5 jsme viděli její variantu pro obor celých čísel, nicméně tato věta platí daleko obecněji. V této sekci si ukážeme další speciální případ, pro polynomy.

Věta 9.1 (čínská věta o zbytcích pro polynomy)**.** *Buď* $\mathbf{T}$ *těleso. Buď* $m_1, \ldots, m_n \in T[x]$ *po dvou nesoudělné polynomy, označme* $d = \sum \deg m_i$*. Buď* $u_1, \ldots, u_n \in T[x]$ *libovolné polynomy. Pak existuje právě jeden polynom* $f \in T[x]$ *stupně* $< d$*, který řeší soustavu kongruencí*

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

Důkaz. Nejprve dokážeme jednoznačnost. Předpokládejme, že soustava má dvě řešení f, g stupně $< d$, tj. pro každé i platí

$$f \equiv g \equiv u_i \pmod{m_i}.$$

Polynom $f-g$ je také stupně $< d$, je dělitelný každým m_i, a protože jsou polynomy m_i navzájem nesoudělné, dostáváme (díky gaussovskosti oboru $\mathbf{T}[x]$)

$$m_1 \cdot \ldots \cdot m_n \mid f - g.$$

Čili polynom stupně d dělí polynom stupně $< d$, což je možné pouze v tom případě, že $f-g = 0$, tj. $f = g$.

Nyní dokážeme, že nějaké řešení existuje. Označme

$$P_k = \{f \in T[x] : \deg f < k\}$$

a uvažujme tuto množinu jako vektorový prostor dimenze k nad tělesem $\mathbf{T}$ (jeho bází jsou polynomy $1, x, x^2, \ldots, x^{k-1}$, každý polynom stupně $< k$ je lineární kombinací těchto polynomů s koeficienty z $\mathbf{T}$). Označme $d_i = \deg m_i$ a uvažujme zobrazení

$$\begin{aligned} \varphi : P_d &\to P_{d_1} \times \cdots \times P_{d_n} \\ f &\mapsto (f \bmod m_1, \ldots, f \bmod m_n). \end{aligned}$$

Uvědomte si, že jde o homomorfismus vektorových prostorů, neboť $(f + g) \bmod m = (f \bmod m) + (g \bmod m)$ a $af \bmod m = a(f \bmod m)$ pro libovolné polynomy f, g, m a každé $a \in T$. V předchozím odstavci jsme ukázali, že zobrazení φ je prosté. Přitom definiční obor i obor hodnot mají stejnou dimenzi $d = \sum d_i$, a podle jisté věty z lineární algebry (v jistém smyslu jde o analogii Lemmatu 1.4) je prosté zobrazení mezi vektorovými prostory stejné konečné dimenze také *na*. Tedy ke každé n-tici $(u_1, \ldots, u_n)$ existuje právě jedno f, které se na něj zobrazuje, a to je hledaným řešením soustavy. □

Stejně jako pro celá čísla, i tento důkaz je nekonstruktivní a využívá konečnosti, tentokráte dimenze jistého vektorového prostoru.

Poznámka. Vzhledem k aplikacím je na místě uvést návod, jak se řešení hledá. V jednom kroku snížíme počet kongruencí o jedna, a tento krok opakujeme tak dlouho, než zbyde jedna kongruence, jejíž řešení je očividné. Podobný postup funguje i pro číselnou verzi.

Uvažujme soustavu dvou kongruencí. Z kongruence $f \equiv u_2 \pmod{m_2}$ vyjádříme $f = gm_2 + u_2$ pro nějaký polynom $g \in T[x]$ a dosadíme do první kongruence

$$f = gm_2 + u_2 \equiv u_1 \pmod{m_1}.$$

Označme $\widetilde{m_2}$ inverz polynomu m_2 modulo m_1, tj. takový polynom, pro který platí $m_2\widetilde{m_2} \equiv 1 \pmod{m_1}$. Ten najdeme pomocí Bézoutovy rovnosti: napíšeme $1 = \mathrm{NSD}(m_1, m_2) = um_1 + vm_2$ a vidíme, že $vm_2 \equiv 1 \pmod{m_1}$. Přenásobením původní kongruence polynomem $\widetilde{m_2}$ dostáváme

$$g \equiv gm_2\widetilde{m_2} \equiv (u_1 - u_2)\widetilde{m_2} \pmod{m_1},$$

řešením tedy je každý polynom $g = hm_1 + (u_1 - u_2)\widetilde{m_2}$, pro libovolné $h \in T[x]$. Zpětným dosazením dostaneme obecné řešení

$$f = gm_2 + u_2 = hm_1m_2 + (u_1 - u_2)\widetilde{m_2}m_2 + u_2$$

Původní dvojice kongruencí je tedy ekvivalentní jedné kongruenci $f \equiv u \pmod{m_1m_2}$, pro jisté u (pokud chceme řešení stupně $< d$, stačí vzít $u \bmod m_1m_2$). Podmínka nesoudělnosti je zachovaná: jsou-li oba polynomy m_1, m_2 nesoudělné se všemi m_i, pak je s nimi nesoudělný i polynom m_1m_2 (opět se využije gaussovskost).

Úloha. Najděte polynom $f \in \mathbb{Q}[x]$ stupně < 4 splňující

$$f \equiv 1 \pmod{x^2-1} \quad \text{a} \quad f \equiv x+1 \pmod{x^2+1}.$$

Řešení. Z druhé kongruence vyjádříme $f = g \cdot (x^2+1) + x + 1$ a dosadíme do první kongruence: budeme hledat $g \in \mathbb{Q}[x]$ splňující $g \cdot (x^2+1) \equiv -x \pmod{x^2-1}$. Všimněte si, že $\widetilde{x^2+1} = \frac{1}{2}$ (protože $x^2+1 \equiv 2$), takže dostaneme vyjádření $g \equiv -\frac{1}{2}x \pmod{x^2-1}$, čili řešením je každý polynom $g = h \cdot (x^2-1) - \frac{1}{2}x$, $h \in \mathbb{Q}[x]$. Zpětným dosazením dostaneme obecné řešení

$$f = h \cdot (x^2-1)(x^2+1) - \frac{1}{2}x(x^2+1) + (x+1),\ h \in \mathbb{Q}[x],$$

a hledaný polynom $f = -\frac{1}{2}x^3 + \frac{1}{2}x + 1$ dostaneme volbou $h = 0$. □

Speciálním případem čínské věty o zbytcích je *věta o interpolaci*: ta říká, že pokud předepíšeme hodnoty v n různých bodech, pak existuje právě jeden polynom stupně $< n$, který těchto hodnot v daných bodech nabývá.

Důsledek 9.2 (věta o interpolaci)**.** *Buď* $\mathbf{T}$ *těleso. Mějme po dvou různé body* $a_1, \dots, a_n \in T$ *a libovolné hodnoty* $u_1, \dots, u_n \in T$*. Pak existuje právě jeden polynom* $f \in T[x]$ *stupně* $< n$ *splňující* $f(a_i) = u_i$ *pro všechna* $i = 1, \dots, n$*.*

Důkaz. Připomeňme, že $f \equiv f(u) \pmod{x-u}$. Řešíme tedy soustavu kongruencí $f \equiv u_i \pmod{x - a_i}$. □

Na rozdíl od obecné věty o zbytcích není těžké nalézt vzorec, který určuje řešení interpolační úlohy: je jím tzv. *Lagrangeův interpolační polynom*

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

Dosazením do vzorce snadno zjistíme, že

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

Jednoznačnost pak plyne z věty o počtu kořenů vztažené na rozdíl $f - g$ dvou řešení.

Důsledek 9.3 (zobrazení na konečných tělesech jsou polynomiální)**.** *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* $\varphi(a) = f(a)$ *pro každé* $a \in T$*.*

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

Pro *nekonečná* tělesa nic takového platit nemůže, přesto polynomy hrají důležitou roli i v reálné analýze: každou spojitou reálnou funkci lze libovolně přesně aproximovat polynomiální funkcí, v různých smyslech. Například, lokální aproximaci (na okolí daného bodu) popisují Taylorovy polynomy, globální (na intervalu) třeba Weierstrassova věta, která říká, že pro každou spojitou reálnou funkci $\varphi : [u,v] \to \mathbb{R}$ na omezeném uzavřeném intervalu a pro každé $\varepsilon > 0$ existuje polynom $f \in \mathbb{R}[x]$ takový, že $|\varphi(a) - f(a)| < \varepsilon$ pro každé $a \in [u,v]$.

9.2. **Faktorokruh modulo polynom.**

Připomeňme konstrukci okruhů $\mathbb{Z}_m$. Začali jsme s oborem celých čísel a uvažovali všechny možné zbytky po dělení m, tj. čísla $0, \dots, m-1$, a na nich operace modulo m. Pokud bylo m prvočíslo, dostali jsme těleso. Podobný postup lze provést i s polynomy, dostaneme tzv. *faktorokruhy*. Aby se nepletla proměnná v polynomech s prvky faktorokruhu, obvykle se v konstrukci používá proměnná α.

Definice. Buď $\mathbf{T}$ těleso a zvolme polynom $m \in T[\alpha]$ stupně $n \geq 1$. *Faktorokruhem* $\mathbf{T}[\alpha]/(m)$ rozumíme množinu všech polynomů stupně $< n$ se standardními operacemi sčítání a odčítání a s operací násobení modulo m. Ve zkratce,

$$\mathbf{T}[\alpha]/(m) = (\{f \in T[\alpha] : \deg f < n\}, +, -, \odot, 0, 1),$$

kde $f \odot g = f \cdot g \bmod m$.

Předně je potřeba dokázat, že to je skutečně komutativní okruh. Axiomy obsahující pouze sčítání a odčítání jsou zřejmé, protože tyto operace jsou totožné jako v $\mathbf{T}[x]$. Pro úvahy s násobením je třeba si připomenout, že $f \equiv g \pmod m \Leftrightarrow f \bmod m = g \bmod m$, a že $f \bmod m \equiv f \pmod m$. Tímto způsobem lze všechny identity přeložit do kongruencí, kde je platnost zřejmá. Například pro asocitivitu dokazujeme

$$(f \odot g) \odot h = f \odot (g \odot h),$$

tj.

$$(f \cdot g \bmod m) \cdot h \bmod m = f \cdot (g \cdot h \bmod m) \bmod m,$$

což je ekvivalentní kongruenci

$$(f \cdot g) \cdot h \equiv f \cdot (g \cdot h) \pmod m,$$

což je pravda pro všechny polynomy f, g, h. Podobně můžeme ověřit distributivitu.

Příklad. Uvažujme faktorokruh $\mathbb{R}[\alpha]/(\alpha^2 + 1)$. Jeho prvky jsou polynomy $a + b\alpha$, $a, b \in \mathbb{R}$. Sčítání probíhá po složkách, tj. $(a + b\alpha) + (c + d\alpha) = (a + c) + (b + d)\alpha$. Násobení vypadá takto:

$$\begin{aligned}(a + b\alpha) \odot (c + d\alpha) &= ac + (ad + bc)\alpha + bd\alpha^2 \bmod (\alpha^2 + 1)\\ &= (ac - bd) + (ad + bc)\alpha.\end{aligned}$$

Všimněte si, že jsme dostali stejné vzorce jako pro sčítání a násobení komplexních čísel. Při ztotožnění symbolů i a α bychom mohli psát, že $\mathbb{R}[\alpha]/(\alpha^2 + 1) = \mathbb{C}$ (formálně, zobrazení $a + b\alpha \mapsto a + bi$ je izomorfismus). Vysvětlení je prosté: při počítání modulo $\alpha^2 + 1$ vlastně zaměňujeme α^2 za -1, neboť $\alpha^2 \equiv -1 \pmod{\alpha^2 + 1}$. Čili pracujeme přesně s vlastností, která definuje komplexní jednotku.

Podobně lze nahlédnout, že faktorokruh $\mathbb{Q}[\alpha]/(\alpha^2 + 1)$ je izomorfní tělesu $\mathbb{Q}(i)$.

Příklad. Nad tělesy $\mathbb{Z}_p$ jsou vlastnosti faktorokruhu závislé na p.

- Faktorokruh $\mathbb{Z}_2[\alpha]/(\alpha^2+1)$ má čtyři prvky, ale není to těleso, dokonce ani obor integrity, protože
$$(\alpha + 1) \odot (\alpha + 1) = \alpha^2 + 1 \bmod (\alpha^2 + 1) = 0.$$
- Faktorokruh $\mathbb{Z}_3[\alpha]/(\alpha^2 + 1)$ má devět prvků. Je to těleso, ale na první pohled to vidět není.

Kdy dostaneme těleso vysvětluje následující tvrzení.

Tvrzení 9.4 (faktor podle ireducibilního prvku)**.** *Buď* $\mathbf{T}$ *těleso a* $m \in T[\alpha]$ *stupně* ≥ 1. *Následující tvrzení jsou ekvivalentní:*

(1) $\mathbf{T}[\alpha]/(m)$ *je těleso,*
(2) $\mathbf{T}[\alpha]/(m)$ *je obor integrity,*
(3) m *je ireducibilní prvek v* $\mathbf{T}[\alpha]$.

Důkaz. (1) $\Rightarrow$ (2) viz Tvrzení 1.3.

(2) $\Rightarrow$ (3). Pro spor předpokládejme, že v $\mathbf{T}[\alpha]$ existuje rozklad $m = f \cdot g$, kde $\deg f, \deg g < \deg m$. Pak ale v $\mathbf{T}[\alpha]/(m)$ platí $f \odot g = m \bmod m = 0$, spor.

(3) $\Rightarrow$ (1). Uvažujme polynom $f \neq 0$ stupně menšího než $\deg m$. Protože je m ireducibilní, platí $1 = \mathrm{NSD}(f, m) = uf + vm$ pro nějaké polynomy $u, v \in T[\alpha]$. Označme $\tilde{u} = u \bmod m$. Pak v $\mathbf{T}[\alpha]/(m)$ platí $\tilde{u} \odot f = \tilde{u}f \bmod m \equiv uf \equiv 1 \pmod m$, čili $\tilde{u}$ je hledaný inverzní prvek k f. $\square$

V dalším textu budeme místo symbolu $\odot$ psát standardní symbol násobení. Z kontextu bude vždy jasné, že jde o násobení ve faktorokruhu, tedy modulo m (stejně se používá standardní symbol pro násobení v okruzích $\mathbb{Z}_m$).

9.3. **Kořenová a rozkladová nadtělesa.**

Nyní si ukážeme jednu důležitou aplikaci konstrukce faktorokruhu: každé těleso lze rozšířit tak, aby v něm měl daný polynom kořen. Pro racionální polynomy to zní triviálně, každý racionální polynom má přece komplexní kořen, ale tento fakt je předmětem Základní věty algebry (Věta 12.1), kterou zatím nemáme dokázanou. Naopak, existence rozkladového nadtělesa je stěžejním krokem k jejímu důkazu. A pro konečná tělesa žádnou analogii použít nelze.

Tvrzení 9.5. *Buď* $\mathbf{T}$ *těleso a* $f \in T[x]$ *stupně* ≥ 1. *Pak existuje těleso* $\mathbf{S} \geq \mathbf{T}$, *kde má polynom* f *kořen.*

Důkaz. Buď m nějaký ireducibilní dělitel polynomu f, označme $m = \sum_{i=0}^{n} a_i x^i$. Stačí najít nadtěleso, kde má kořen polynom m, ten bude kořenem i pro f. Uvažujme faktorokruh $\mathbf{S} = \mathbf{T}[\alpha]/(m(\alpha))$. Podle Tvrzení 9.4 je $\mathbf{S}$ těleso. Vyhodnotíme-li v $\mathbf{S}$ polynom m na prvku α, dostaneme
$$m(\alpha) = \sum_{i=0}^{n} a_i(\alpha^i \bmod m(\alpha)) = \sum_{i=0}^{n-1} a_i\alpha^i + a_n(\alpha^n \bmod m(\alpha)),$$
ovšem $a_n\alpha^n \bmod m(\alpha) = -\sum_{i=0}^{n-1} a_i\alpha^i$, takže se to odečte na nulu. Prvek α je tedy kořenem obou polynomů m, f v nadtělese $\mathbf{S}$. □

Příklad.

- Pro polynom $x^3 - 2$ nad tělesem $\mathbb{Q}$ dostaneme těleso $\mathbb{Q}[\alpha]/(\alpha^3 - 2)$, které lze pomocí úvah v předchozí podsekci ztotožnit s tělesem $\mathbb{Q}(\sqrt[3]{2})$.
- Pro polynom $x^3 - 2$ nad tělesem $\mathbb{Z}_7$ dostaneme těleso $\mathbb{Z}_7[\alpha]/(\alpha^3 - 2)$, což je těleso s 7^3 prvky, které jste ještě asi nepotkali.

Indukcí snadno dokážeme, že existuje také nadtěleso, kde má daný polynom všechny kořeny, tj. kde se rozkládá na součin lineárních činitelů (polynomů stupně 1).

Věta 9.6. *Buď* $\mathbf{T}$ *těleso a* $f \in T[x]$ *stupně* ≥ 1. *Pak existuje těleso* $\mathbf{S} \geq \mathbf{T}$, *kde se polynom* f *rozkládá na součin polynomů stupně 1.*

Důkaz. Budeme postupovat indukcí podle stupně polynomu f. Je-li f stupně 1, $f = ax - b$, pak má kořen $a^{-1}b$ již v tělese $\mathbf{T}$. V opačném případě uvažujme nadtěleso $\mathbf{U} \geq \mathbf{T}$, kde má polynom f kořen u, a uvažujme polynom $g \in U[x]$ takový že $f = g \cdot (x - u)$. Protože $\deg g < \deg f$, podle indukčního předpokladu existuje nadtěleso $\mathbf{S} \geq \mathbf{U}$, kde se g rozkládá na součin polynomů stupně 1, čili se tam rozkládá i f. □

Minimální nadtělesa, kde má daný polynom kořen, resp. kde se rozkládá na lineární činitele, mají své jméno. Jak si později ukážeme, hrají stěžejní roli v Galoisově teorii.

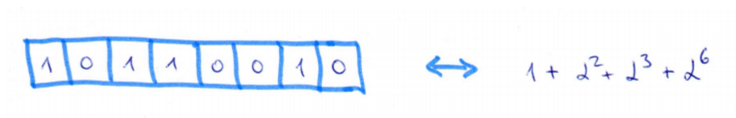
Definice. Buď $\mathbf{T}$ těleso a $f \in T[x]$ stupně ≥ 1.

(1) *Kořenovým nadtělesem* polynomu f rozumíme těleso $\mathbf{S} \geq \mathbf{T}$, ve kterém existuje $a \in S$ takové, že $\mathbf{S} = \mathbf{T}(a)$ a $f(a) = 0$.
(2) *Rozkladovým nadtělesem* polynomu f rozumíme těleso $\mathbf{S} \geq \mathbf{T}$, ve kterém existují $a_1, \dots, a_n \in S$ taková, že $\mathbf{S} = \mathbf{T}(a_1, \dots, a_n)$ a $f \parallel (x - a_1) \cdot \ldots \cdot (x - a_n)$ v $\mathbf{S}[x]$.

Důsledek 9.7 (existence kořenových a rozkladových nadtěles)**.** *Buď* $\mathbf{T}$ *těleso a* $f \in T[x]$ *stupně* ≥ 1. *Pak existuje kořenové i rozkladové nadtěleso polynomu* f *nad* $\mathbf{T}$.

Důkaz. Podle Tvrzení 9.5 existuje nadtěleso $\mathbf{S}$, kde má f kořen a. Kořenovým nadtělesem bude podtěleso $\mathbf{T}(a) \leq \mathbf{S}$.

Analogicky, podle Věty 9.6 existuje nadtěleso $\mathbf{S}$ takové, že $f \parallel (x - a_1) \cdots (x - a_n)$ v $\mathbf{S}[x]$. Kořenovým nadtělesem bude podtěleso $\mathbf{T}(a_1, \dots, a_n) \leq \mathbf{S}$. □



Obrázek 11. Bitvektor a jeho reprezentace polynomem.

Problém jednoznačnosti budeme řešit v sekci 24.2. Věta 24.2 říká, že všechna rozkladová nadtělesa daného polynomu jsou izomorfní, a že kořenová nadtělesa jsou izomorfní za předpokladu, že je daný polynom ireducibilní.

10. Konečná tělesa a jejich aplikace

10.1. Konečná tělesa a počítačová reprezentace dat.

Důležitou aplikací faktorokruhů je konstrukce konečných těles. Buď p prvočíslo a uvažujme ireducibilní polynom $m \in \mathbb{Z}_p[\alpha]$ stupně k. Faktorokruh $\mathbb{Z}_p[\alpha]/(m)$ je podle Tvrzení 9.4 tělesem, jeho prvky jsou polynomy stupně $< k$ nad $\mathbb{Z}_p$, čili toto těleso má právě p^k prvků. Například,

- čtyřprvkové těleso můžeme zkonstruovat jako $\mathbb{Z}_2[\alpha]/(\alpha^2+\alpha+1)$,
- osmiprvkové jako $\mathbb{Z}_2[\alpha]/(\alpha^3+\alpha+1)$ nebo $\mathbb{Z}_2[\alpha]/(\alpha^3+\alpha^2+1)$,
- devítiprvkové jako $\mathbb{Z}_3[\alpha]/(\alpha^2+1)$ nebo $\mathbb{Z}_3[\alpha]/(\alpha^2 \pm \alpha + 2)$.

Pozor, pro $k > 1$ je p^k-prvkové těleso něco jiného než okruh $\mathbb{Z}_{p^k}$!

V sekci 26 si dokážeme Věty 26.3 a 26.4, které říkají, že

1. pro každé p, k existuje ireducibilní polynom stupně k nad $\mathbb{Z}_p$,
2. každé konečné těleso lze sestrojit jako faktorokruh $\mathbb{Z}_p[\alpha]/(m)$,
3. na volbě m nezáleží, tj. jsou-li m_1, m_2 dva ireducibilní polynomy stupně k nad tělesem $\mathbb{Z}_p$, pak jsou tělesa $\mathbb{Z}_p[\alpha]/(m_1)$ a $\mathbb{Z}_p[\alpha]/(m_2)$ izomorfní.

Důkaz uvedených vlastností je složitější, než se teď může zdát. Dokázat samotnou existenci konečného tělesa velikosti p^k je poměrně snadné: dostaneme jej jako rozkladové nadtěleso polynomu $x^{p^k} - x$ nad tělesem $\mathbb{Z}_p$ (důkazu Lemmatu 26.1 by čtenář porozuměl již nyní). Avšak k reprezentaci pomocí faktorokruhů je potřeba jednoznačnost rozkladových nadtěles (sekce 24.2), teorie tělesových rozšíření konečného stupně (sekce 22) a také některá fakta z teorie grup (Lagrangeova věta 14.9 a sekce 16 o struktuře cyklických grup).

Konečné těleso velikosti p^k budeme značit $\mathbb{F}_{p^k}$ (používá se také označení $\mathbf{GF}(p^k)$, jako *Galois field*). Vzhledem k tomu, že jsou stejně velká tělesa izomorfní, na konkrétní reprezentaci (zpravidla) nezáleží.

Konečná tělesa, zejména ta velikosti 2^k, mají zásadní využití v informatice. Jejich pomocí lze reprezentovat počítačová data a provádět s nimi různé operace, anebo v nich datové operace analyzovat. Reprezentaci dat si nyní vysvětlíme.

Základním datovým objektem, se kterým pracují počítače, jsou tzv. *bitvektory*, tedy k-tice nul a jedniček, tzv. *bitů*. Délka k bývá mocnina dvojky, na starých počítačích byl standard $k = 8$ (osmicím se říká *bajty*), na moderních strojích je obvykle $k = 32$ nebo $k = 64$. Bitvektory délky k lze přirozeně reprezentovat pomocí konečného tělesa $\mathbb{F}_{2^k} = \mathbb{Z}_2[\alpha]/(m)$, kde m je ireducibilní stupně k: vektor $(a_0, \dots, a_{k-1})$ se reprezentuje polynomem $a_0 + a_1\alpha + \ldots + a_{k-1}\alpha^{k-1}$.

Běžné operace na bitvektorech mají často přirozenou tělesovou interpretaci. Například posun bitvektoru vlevo či vpravo odpovídá násobení a dělení prvkem α. Logická spojka XOR po bitech odpovídá tělesovému sčítání, logická spojka AND po bitech odpovídá násobení po koeficientech (pozor, to je něco jiného, než násobení v tělese). Konečná tělesa přinášejí navíc dvě zajímavé operace, které pracují se všemi bity najednou: tělesové násobení a invertování. Tyto operace mají zajímavé vlastnosti, které nacházejí uplatnění v konstrukci šifer.

Příklad. V současnosti nejpoužívanější symetrická šifra AES (*Advanced Encryption Standard*, též známá jako *Rijndael*) pracuje s bitvektory délky 8, které reprezentuje pomocí prvků tělesa

$$\mathbb{F}_{256} = \mathbb{Z}_2[\alpha]/(\alpha^8 + \alpha^4 + \alpha^3 + \alpha + 1).$$

Text je rozdělen na bloky po 128 bitech a každý blok je reprezentován jako matice 4×4 prvků tělesa $\mathbb{F}_{256}$. Šifra opakuje pro každý blok několikrát za sebou čtyři fáze (první a poslední průběh je trochu jiný, ale to teď není důležité). V první se provádí pro každý prvek matice následující operace:

$$u \mapsto u^{-1} \cdot (1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4) + (1 + \alpha + \alpha^5 + \alpha^8) \bmod (\alpha^8 + 1),$$

kde inverz se bere v tělese $\mathbb{F}_{256}$ a zbytek výpočtu probíhá v $\mathbb{Z}_2[\alpha]$. V druhé fázi se rotuje každý řádek o jistý počet pozic. Ve třetí se mixuje každý sloupec tak, že se interpretuje jako polynom z $\mathbb{F}_{256}[x]$ stupně < 4, na který se provede operace

$$f \mapsto f \cdot (\alpha + x + x^2 + (\alpha + 1)x^3) \bmod (x^4 + 1).$$

Ve čtvrté fázi se pak přičítá jistým způsobem vybraná část klíče (po bitech). Smysl prvních tří fází je rozprostřít změnu provedenou přičítáním klíče do celé tabulky (míchání prvků, řádků, sloupců), a to tak, aby se co nejlépe ztratily slabiny opakování klíče. Pro praxi je zásadní, že lze velmi rychle nejen šifrovat, ale také dešifrovat: není těžké odvodit, že operace inverzní k těm výše uvedeným mají podobně jednoduchý algebraický zápis.

Velký význam má také Důsledek 9.3 a jeho zobecnění pro zobrazení více proměnných (viz cvičení v sekci 9.1), které říká, že každou operaci na datech lze interpretovat jako polynomiální zobrazení nad příslušným tělesem. Tento fakt nachází uplatnění například v kryptoanalýze. Mezi další aplikace věty o interpolaci patří samoopravné kódy, kterými se na přednášce zabývat nebudeme. Naopak, ukážeme si algoritmus na sdílení tajemnství (sekce 10.2).

Další oblastí aplikace konečných těles jsou konečné geometrie (afinní a projektivní prostory nad konečnými tělesy). Konečné geometrie jsou zdrojem zajímavých kombinatorických objektů, na proseminaři si ukážeme konstrukci navzájem ortogonálních latinských čtverců pomocí afinních zobrazení. Konečné geometrie jsou také zdrojem zajímavých výpočetních problémů, jako je například počítání na eliptických křivkách nad konečnými tělesy, opět s aplikací v návrhu šifer.

10.2. **Sdílení tajemství.**

Motivační úloha je následující: armáda má tajný kód, který umožňuje odpálit jaderné rakety. Zřejmě není dobré, aby jeden šílenec mohl odpálit rakety o své vůli. Ani dva šílenci by neměli mít možnost odpálit rakety. Prezident nařídil, že k odpálení raket je potřeba souhlas aspoň tří šílenců ze sedmičlenného generálního štábu. Jak to zařídit?

Obecně hovoříme o (k, n)-*schématu sdílení tajemství*, pokud se n účastníků dělí o tajemství, k jehož odhalení je potřeba přítomnost alespoň k z nich. V celém odstavci budeme uvažovat, že sdílíme tajemství t z nějakého tělesa $\mathbf{T}$. V praxi se sdílí bitvektor délky m, interpretovaný buď jako m tajemství z tělesa $\mathbf{T} = \mathbb{Z}_2$, nebo jedno tajemství z $\mathbf{T} = \mathbb{F}_{2^m}$.

Pro případ $k = n$ lze použít jednoduché schéma založené na *maskování hodnot*. Vlastník tajemství vydá každému účastníku náhodný prvek $a_i \in T$ a zveřejní hodnotu $c = t + \sum a_i$. Pokud má dojít k odhalení, každý účastník sdělí své a_i a společně spočtou $t = c - \sum a_i$. Pokud se sejde účastníků méně, byť jen $n - 1$, o hodnotě t nemohou říci vůbec nic: chybějící prvek může hodnotu součtu změnit na libovolnou jinou hodnotu, s pravděpodobností přesně $\frac{1}{|T|}$ (protože zobrazení $x \mapsto a + x$ je permutace). V praxi, pro bitvektor délky m, se schéma použije m-krát pro těleso $\mathbb{Z}_2$. Pravděpodobnost uhodnutí jednoho bitu je $\frac{1}{2}$, čili pro m-bitový klíč je pravděpodobnost $(\frac{1}{2})^m$.

Klasickým řešením obecného (k, n)-schématu je tzv. *Shamirův protokol*. Vlastník tajemství náhodně zvolí polynom $f \in T[x]$ stupně $< k$ takový, že $f(0) = t$ (tj. tajemství je absolutní člen f), vybere n po dvou různých prvků $0 \neq a_1, \dots, a_n \in T$ (ta mohou být veřejná) a jednotlivým účastníkům rozdá hodnoty $f(a_1), \dots, f(a_n)$. Pokud se sejde libovolných k účastníků, vezmou své hodnoty, provedou interpolaci ve svých bodech a spočtou (ten jediný) polynom stupně $< k$, který vyhovuje jejich podmínkám; tajemství je jeho absolutní člen. Naopak, pokud se jich sejde méně, byť jen $k-1$, o absolutním členu nezjistí nic: $k-1$ nenulovými body lze proložit polynom s libovolnou hodnotou v bodě 0, a navíc rozložení hodnot v 0 je rovnoměrné. V praxi se pro

m-bitový klíč používá těleso s 2^m prvky (musí být $2^m > n$), které zajistí pravděpodobnost náhodného uhodnutí $\frac{1}{2^m}$.

Schéma lze snadno modifikovat pro sofistikovanější úlohy. Například, co kdyby prezident rozhodl, že rakety mohou odpálit buď aspoň tři ze sedmi šílených generálů, nebo on sám? Snadná pomoc: vyrobíme $(3, 1O)$-schéma, každému z generálů dáme po jednom dílu a prezidentu dáme tři. A tak podobně. V reálném životě se schéma používá například pro rozhodnutí komisí, tajemstvím je klíč k elektronickému podpisu.

11. Symetrické polynomy a Vìètovy vztahy

Definice. Buď $\mathbf{R}$ libovolný komutativní okruh. Polynom $f \in R[x_1, \dots, x_n]$ nazveme *symetrický*, pokud po libovolném přeuspořádání proměnných dostaneme ten samý polynom. Formálně, pokud

$$f(x_1, \dots, x_n) = f(x_{\pi(1)}, \dots, x_{\pi(n)})$$

pro libovolnou permutaci π na množině indexů $\{1, \dots, n\}$.

Příklad. Polynomy $x^k+y^k+z^k$ a $x^ky^kz^k$ třech proměnných x, y, z jsou symetrické, pro libovolné k.

Příklad. Roznásobme součin $(y - x_1)(y - x_2)(y - x_3)$ a podívejme se na něj jako na polynom v proměnné y, jehož koeficienty jsou z $\mathbf{R}[x_1, x_2, x_3]$:

$$(y - x_1)(y - x_2)(y - x_3) = y^3 - (x_1 + x_2 + x_3)y^2 + (x_1x_2 + x_1x_3 + x_2x_3)y - (x_1x_2x_3).$$

Vidíme, že všechny koeficienty jsou symetrické polynomy vzhledem k proměnným x_1, x_2, x_3.

Předchozí příklad lze zobecnit na libovolný počet činitelů. Označme

$$\text{(V)} \qquad (y - x_1) \cdot \ldots \cdot (y - x_n) = y^n - s_1y^{n-1} + s_2y^{n-2} - s_3y^{n-3} + \ldots + (-1)^n s_n.$$

Vzhledem k tomu, že v součinu nezáleží na pořadí, všechny koeficienty budou symetrické polynomy, přičemž je snadné dopočítat, že

$$\begin{aligned}
s_1 &= x_1 + x_2 + \dots + x_n = \sum_i x_i,\\
s_2 &= x_1x_2 + x_1x_3 + \ldots + x_{n-1}x_n = \sum_{i<j} x_ix_j,\\
&\ldots\\
s_k &= \sum_{i_1<\ldots<i_k} x_{i_1} \cdots x_{i_k},\\
&\ldots\\
s_n &= x_1x_2 \cdots x_n.
\end{aligned}$$

Těmto polynomům se říká *elementární symetrické polynomy* v proměnných $x_1, \dots, x_n$. Z rovnosti (V) pak plynou známé *Vìètovy vztahy.*

Tvrzení 11.1 (Vìètovy vztahy). *Buď $\mathbf{T}$ těleso a $f = \sum a_ix^i$ polynom z $\mathbf{T}[x]$ stupně ≥ 1. Uvažujme jeho rozklad $f = a_n(x - u_1) \cdot \ldots \cdot (x - u_n)$ v nějakém nadtělese $\mathbf{S} \geq \mathbf{T}$. Pak*

$$\frac{a_{n-i}}{a_n} = (-1)^i \cdot s_i(u_1, \dots, u_n).$$

Důkaz. Uvažujme polynom $g = a_n^{-1}f$. Do rovnosti (V) dosadíme za proměnné x_i prvky $u_i \in S$ a dostaneme

$$g = \sum_{i=0}^{n} \frac{a_i}{a_n}x^i = (x - u_1) \cdot \ldots \cdot (x - u_n) = x^n + \sum_{i=1}^{n}(-1)^i s_i(u_1, \dots, u_n)x^{n-i}.$$

Porovnáním koeficientů dostaneme Vìètovy vztahy. □

Díky Viètovým vztahům můžeme určit některé vlastnosti kořenů daného polynomu, aniž bychom znali jejich konkrétní hodnoty. Například víme, že jejich součet je $-\frac{a_{n-1}}{a_n}$ (dosaďte do s_1), jejich součin je $(-1)^n\frac{a_0}{a_n}$ (dosaďte do s_n).

Příklad. Všimněte si, že

$$x_1^2+\ldots+x_n^2=s_1^2-2s_2$$

(ověřte roznásobením!). Z Viètových vztahů plyne, že součet čtverců všech kořenů daného polynomu $\sum a_ix^i$ je roven

$$u_1^2+\ldots+u_n^2=s_1(u_1,\ldots,u_n)^2-2s_2(u_1,\ldots,u_n)=\left(\frac{a_{n-1}}{a_n}\right)^2-2\cdot\frac{a_{n-2}}{a_n}.$$

Všimněte si, že součet, rozdíl a součin symetrických polymů je symetrický polynom (čili symetrické polynomy tvoří podokruh okruhu všech polynomů). Speciálně, různé součty a součiny elementárních symetrických polynomů jsou symetrické. Je pravda i opačné trvrzení? V předchozím příkladu jsme viděli, že součet čtverců, což je symetrický polynom, lze takto vyjádřit. Důležitým poznatkem je, že tuto vlastnost má každý symetrický polynom.

Věta 11.2 (základní věta o symetrických polynomech)**.** *Buď* **R** *obor integrity a $f\in R[x_1,\ldots,x_n]$ symetrický polynom. Pak existuje* právě jeden *polynom $g\in R[z_1,\ldots,z_n]$ takový, že $f=g(s_1,\ldots,s_n)$.*

Polynom g ve znění věty popisuje, jak získat f pomocí součtů a součinů elementárních polynomů. Například pro $f=x_1^2+\ldots+x_n^2=s_1^2-2s_2$ bude $g=z_1^2-2z_2$.

Ve zbytku sekce dokážeme Větu 11.2. Předvedeme si Gaussův důkaz obsahující algoritmus, který vyjádření daného symetrického polynomu najde. Nejprve si však musíme vysvětlit, jak uspořádat členy v polynomech více proměnných, abychom mohli pracovat s pojmem vedoucího členu.

Termem v proměnných $x_1,\ldots,x_m$ rozumíme výraz $x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$, kde $k_1,\ldots,k_n\geq 0$. Definujeme relaci na termech:

$$x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}<x_1^{l_1}x_2^{l_2}\cdots x_n^{l_n},$$

pokud existuje $i\geq 0$ takové, že $k_1=l_1$, …, $k_i=l_i$ a $k_{i+1}<l_{i+1}$. Definujeme $t\leq s$ právě tehdy, když $t<s$ nebo $t=s$. Jak si ukážeme, jde o uspořádání, tzv. *lexikografické uspořádání*, velmi podobné uspořádání slov ve slovníku. *Vedoucím členem* polynomu $f\in R[x_1,\ldots,x_n]$ se pak rozumí ten člen, který má lexikograficky největší term; značíme jej $\ell(f)$.

Příklad. Pro polynomy tří proměnných se zpravidla používají proměnné x,y,z, implicitně se rozumí v tomto pořadí.

- $\ell(x+y+z)=x$, neboť $x=x^1y^0z^0$ je větší než $y=x^0y^1z^0$, a to je větší než $z=x^0y^0z^1$.
- $\ell(100z^{10}+2x^2y-5x^2z^2)=2x^2y$, neboť $x^2y>x^2z^2>z^{10}$ (koeficienty nerozhodují).

Lemma 11.3. *Relace $\leq$ má nasledující vlastnosti:*

(1) *je to lineární uspořádání,*
(2) *pro libovolné termy platí, že $t_1>t_2$ a $s_1>s_2$ implikuje $t_1s_1>t_2s_2$,*
(3) *neexistuje nekonečný klesající řetězec termů $t_1>t_2>t_3>\ldots$*

Důkaz lemmatu přenecháváme čtenáři jako snadné, byť poněkud pracné cvičení.

Lemma 11.4. *Buď* **R** *obor integrity a $f,g\in R[x_1,\ldots,x_n]$. Pak*

(1) $\ell(fg)=\ell(f)\ell(g)$,
(2) *je-li f symetrický a $\ell(f)=ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$, pak $k_1\geq k_2\geq\ldots\geq k_n$.*

Důkaz. (1) Díky Lemmatu 11.3(2) platí, že součin termů vedoucích členů je větší než součin libovolných jiných termů. Protože jsme v oboru integrity, koeficient součinu nebude nulový.

(2) Kdyby $k_i<k_j$, mohli bychom prohodit proměnné x_i,x_j, ze symetrie bychom dostali ten samý polynom, ale člen s prohozenými proměnnými by byl větší. □

Lemma 11.5. *Buď $k_1\geq k_2\geq\ldots\geq k_n$ nezáporná čásla. Pak existuje* právě jedna *n-tice nezáporných čísel $l_1,\ldots,l_n$ taková, že $\ell(s_1^{l_1}s_2^{l_2}\cdots s_n^{l_n})=x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$.*

Důkaz. Nejprve spočteme vedoucí člen $\ell(s_1^{l_1} s_2^{l_2} \cdots s_n^{l_n})$. Díky Lemmatu 11.4(1) je roven součinu

$$\begin{aligned}\ell(s_1)^{l_1}\ell(s_2)^{l_2}\cdots\ell(s_n)^{l_n} &= x_1^{l_1}\cdot(x_1x_2)^{l_2}\cdot(x_1x_2x_3)^{l_3}\cdot\ldots\cdot(x_1\cdots x_n)^{l_n}\\ &= x_1^{l_1+\ldots+l_n}x_2^{l_2+\ldots+l_n}\cdots x_{n-1}^{l_{n-1}+l_n}x_n^{l_n}.\end{aligned}$$

Máme dány exponenty $k_1,\ldots,k_n$, hledáme $l_1,\ldots,l_n$ splňující soustavu rovnic

$$k_1 = l_1+\ldots+l_n,\ k_2 = l_2+\ldots+l_n,\ \ldots,\ k_{n-1} = l_{n-1}+l_n,\ k_n = l_n.$$

Odečtením dvou po sobě jdoucích rovnic zjistíme, že existuje právě jedno řešení, a to

$$l_1 = k_1-k_2,\ l_2 = k_2-k_3,\ \ldots,\ l_{n-1} = k_{n-1}-k_n,\ l_n = k_n.$$

Tato řešení jsou nezáporná, protože $k_i \geq k_{i+1}$ pro všechna i. □

Nyní zformulujeme Gaussův algoritmus na výpočet vyjádření daného symetrického polynomu pomocí elementárních.

Gaussův algoritmus. Buď **R** obor intergrity.

- `VSTUP:` $f \in R[x_1,\ldots,x_n]$ symetrický
- `VÝSTUP:` $g \in R[z_1,\ldots,z_n]$ takový, že $f = g(s_1,\ldots,s_n)$
- $f_1 = f$, $g_1 = 0$
- pro $i = 2, 3, \ldots$ prováděj následující:
  najdi $l_1,\ldots,l_n$ takové, že $\ell(f_i) = c\cdot\ell(s_1^{l_1}\cdots s_n^{l_n})$ pro nějaké $c \in R$
  $f_{i+1} = f_i - c\cdot s_1^{l_1}\cdots s_n^{l_n}$
  $g_{i+1} = g_i + c\cdot z_1^{l_1}\cdots z_n^{l_n}$
  pokud je $f_{i+1} \in R$ (konstantní polynom), odpověz $g_{i+1}+f_{i+1}$

Dokážeme správnost algoritmu. Všimněte si, že pro všechna i platí

- f_i je symetrický polynom,
- $g_i \in R[z_1,\ldots,z_n]$,
- $f_i + g_i(s_1,\ldots,s_n) = f$.

(Pro $i = 1$ to platí triviálně a dále postupujeme indukcí.) Z těchto tří pozorování plyne správnost odpovědi i fakt, že taková $l_1,\ldots,l_n$ vždy najdeme (Lemmata 11.4(2) a 11.5, resp. algoritmus výpočtu skrytý v důkazu). Na závěr zbývá zpozorovat, že se termy vedoucích členů polynomů $f_1, f_2, \ldots$ zmenšují, čili podle Lemmatu 11.3(3) se algoritmus musí zastavit.

Příklad. Mějme na vstupu polynom $f = x_1^3+\ldots+x_n^3$.

- $f_1 = x_1^3+\ldots+x_n^3$, $g_1 = 0$.
- Vidíme, že $\ell(f_1) = x_1^3 = \ell(s_1^3)$, čili
$$f_2 = f_1 - s_1^3 = -3\sum_{i\neq j} x_i^2x_j - 6\sum_{i<j<k} x_ix_jx_k, \quad g_2 = g_1 + z_1^3 = z_1^3.$$
- Vidíme, že $\ell(f_2) = -3x_1^2x_2 = -3\ell(s_1s_2)$, čili:
$$f_3 = f_2 - (-3)s_1s_2 = 3\sum_{i<j<k} x_ix_jx_k, \quad g_3 = g_2 + (-3)z_1z_2 = z_1^3 - 3z_1z_2.$$
- Vidíme, že $\ell(f_3) = 3x_1x_2x_3 = 3\ell(s_3)$, čili
$$f_4 = f_3 - 3s_3 = 0, \quad g_4 = g_3 + 3z_3 = z_1^3 - 3z_1z_2 + 3z_3.$$

Odpovědí je polynom g_4, čili $f = g_4(s_1,\ldots,s_n) = s_1^3 - 3s_1s_2 + 3s_3$.

Důkaz Věty 11.2. Existence byla prokázána Gaussovým algoritmem. Jednoznačnost dokážeme sporem. Uvažujme dvě vyjádření

$$f = g_1(s_1,\ldots,s_n) = g_2(s_1,\ldots,s_n),$$

$g_1 \neq g_2$, a označme $g = g_1 - g_2 = \sum a_it_i$, kde t_i jsou jednotlivé termy. Tyto termy jsou různé, a tedy díky jednoznačnosti v Lemmatu 11.5 mají polynomy $t_i(s_1,\ldots,s_n)$ různé vedoucí členy.

Uvažujme ten lexikograficky největší z nich. Tím, že je striktně větší než všechny ostatní členy, se v součtu $\sum a_i t_i(s_1, \ldots, s_n)$ nemůže pokrátit, a tedy $g(s_1, \ldots, s_n) \neq 0$, spor. $\square$

Základní věta o symetrických polynomech má zajímavý důsledek, který použijeme v důkazu Základní věty algebry. Uvažujme celočíslený polynom a jeho komplexní kořeny. To mohou být komplexní čísla, která nelze nijak pěkně vyjádřit. Přesto, pokud je dosadíme do symetrického polynomu, vyjde racionální číslo (dokonce celé, pokud byl tento polynom monický).

Důsledek 11.6 (hodnota symetrického polynomu na kořenech)**.** *Buď* $\mathbf{T}$ *těleso a* f *polynom z* $\mathbf{T}[x]$ *stupně* $n \geq 1$*. Buď* $u_1, \ldots, u_n$ *jeho kořeny (včetně násobnosti) v nějakém nadtělese. Pak pro každý symetrický polynom* $s \in T[x_1, \ldots, x_n]$ *platí* $s(u_1, \ldots, u_n) \in T$*.*

Důkaz. Označme $f = \sum a_i x^i$. Díky Viètovým vztahům platí $s_i(u_1, \ldots, u_n) = (-1)^i \frac{a_{n-i}}{a_n} \in T$. Díky Větě 11.2 existuje polynom $g \in T[z_1, \ldots, z_n]$ splňující $s = g(s_1, \ldots, s_n)$, čili $s(u_1, \ldots, u_n)$ je rovno hodnotě polynomu g na n-tici prvků z T, což je opět prvek T. $\square$

12. Základní věta algebry

Cílem této sekce je dokázat, že každý komplexní polynom má komplexní kořen. Tomuto faktu se říká Základní věta algebry, i když název je poněkud zastaralý a odpovídá době svého vzniku, tedy přelomu 18. a 19. století, kdy se algebra zabývala především řešením polynomiálních rovnic. Trochu zavádějící je i samotný odkaz na algebru: důkaz nutně musí využít nějaké analytické metody, neboť se principiálně týká vlastností reálných funkcí, resp. jejich komplexních rozšíření.

Věta 12.1 (základní věta algebry)**.** *Každý komplexní polynom stupně* ≥ 1 *má nějaký komplexní kořen.*

Důsledkem je, že polynom $f \in \mathbb{C}[x]$ stupně n má právě n komplexních kořenů (včetně násobnosti) a rozkládá se v $\mathbb{C}[x]$ na součin

$$f \parallel (x - u_1) \cdot \ldots \cdot (x - u_n),$$

kde $u_1, \ldots, u_n \in \mathbb{C}$. Důkaz provedeme snadno indukcí: máme-li jeden komplexní kořen, u, vydělíme f polynomem $x - u$ a použijeme větu znovu, na polynom menšího stupně.

Jiným očividným důsledkem je, že každé polynomiální zobrazení $\mathbb{C} \to \mathbb{C}$ je *na*: řešíme-li rovnici $f(u) = a$, řešením je kořen polynomu $f - a$.

Důkazů základní věty algebry existuje celá řada, ať už čistě analytické (pomocí komplexní analýzy), geometrické, či algebraické, snažící se minimalizovat potřebné množství vlastností reálných čísel. Ukážeme si Gaussův důkaz z roku 1816, který patří do poslední rodiny, je poměrně jednoduchý a pěkně odděluje analytické a algebraické principy potřebné k důkazu. Z algebry jsou stěžejním nástrojem

- existence rozkladových nadtěles (Věta 9.6),
- teorie symetrických polynomů (klíčový krok důkazu je založen na úvaze podobné Důsledku 11.6).

Z reálné analýzy pak využijeme

- spojitost polynomiálních funkcí,
- větu o středním bodě,

které implikují, že reálné polynomy lichého stupně mají aspoň jeden reálný kořen. Začneme užitečným pozorováním, že problém lze zredukovat na reálné polynomy. K důkazu stačí pár elementárních výpočtů s komplexně sdruženými čísly.

Lemma 12.2. *Předpokládejme, že každý reálný polynom stupně* ≥ 1 *má nějaký komplexní kořen. Pak má každý komplexní polynom stupně* ≥ 1 *nějaký komplexní kořen.*

Důkaz. Buď $f \in \mathbb{C}[x]$ stupně ≥ 1. Označme $g = f \cdot \bar{f}$, kde $\bar{f}$ značí komplexně sdružený polynom, tj. pro $f = \sum a_i x^i$ definujeme $\bar{f} = \sum \bar{a}_i x^i$. Všimněte si, že $g \in \mathbb{R}[x]$: součin má tvar

$$f \cdot \bar{f} = \sum a_i x^i \cdot \sum \bar{a}_i x^i = \sum_k \left(\sum_{i+j=k} a_i \bar{a}_j \right) x^k.$$

Všechny koeficienty jsou reálné, protože pro $i = j$ máme $a_i \bar{a}_i \in \mathbb{R}$, a pro $i \neq j$ máme $a_i \bar{a}_j + a_j \bar{a}_i \in \mathbb{R}$ (všimněte si, že $r\bar{s} + \bar{r}s \in \mathbb{R}$ pro každé $r, s \in \mathbb{C}$). Čili podle předpokladu má polynom g komplexní kořen u, a tedy $f(u) = 0$ nebo $\bar{f}(u) = 0$. V prvním případě jsme hotovi a v druhém případě si všimneme, že $0 = \overline{\bar{f}(u)} = f(\bar{u})$, a tedy f má kořen u nebo $\bar{u}$. □

Lemma 12.3. *Komplexní polynom stupně 2 má komplexní kořen.*

Důkaz. Kořeny polynomu $ax^2 + bx + c$ lze spočítat vzorcem $\frac{-b \pm \sqrt{b^2 - 4ac}}{-2a}$ (snadno ověříme roznásobením, odvození viz sekce 25.1), výsledkem je komplexní číslo. Poněkud skrytý je fakt, že v komplexních číslech lze odmocňovat: pro $z = re^{i\alpha}$ platí $\sqrt{z} = \sqrt{r} e^{i\alpha/2}$, přičemž odmocnina z kladného reálného čísla existuje díky větě o středním bodě aplikované na spojitou funkci $x \mapsto x^2$. □

Lemma 12.4. *Reálný polynom lichého stupně má reálný kořen.*

Důkaz. Reálný polynom f určuje spojitou reálnou funkci. Má-li lichý stupeň, pak v závislosti na znaménku vedoucího koeficientu buď $\lim_{x \to -\infty} f(x) = -\infty$ a $\lim_{x \to \infty} f(x) = \infty$, nebo naopak, tedy existují body a, b takové, že $f(a) < 0$ a $f(b) > 0$. Z věty o středním bodě plyne, že existuje bod $u \in \mathbb{R}$, kde $f(u) = 0$. □

Důkaz Věty 12.1. Díky Lemmatu 12.2 stačí uvažovat reálný polynom f. Označme jeho stupeň $n = 2^k m$, kde m je liché. Budeme postupovat indukcí podle k. Je-li $k = 0$, odpověď dává Lemma 12.4. V indukčním kroku použijeme Větu 9.6 a budeme uvažovat nadtěleso $\mathbf{S}$, nad kterým se f rozkládá na součin linárních polynomů, tj. $f \parallel (x - u_1) \cdot \ldots \cdot (x - u_n)$. Chceme dokázat, že aspoň jeden z kořenů $u_1, \ldots, u_n$ je v $\mathbb{C}$. Pro každý parametr $a \in \mathbb{Z}$ definujeme polynom

$$h_a = \prod_{i<j} (x - (u_i + u_j + a u_i u_j)) \in S[x].$$

Klíčovým krokem je ukázat, že to jsou ve skutečnosti reálné polynomy. Uvažujte polynom

$$\tilde{h}_a = \prod_{i<j} (x - (y_i + y_j + a y_i y_j)) \in (\mathbb{Z}[x])[y_1, \ldots, y_n].$$

Ten je symetrický v proměnných $y_1, \ldots, y_n$ (koeficienty jsou ze $\mathbb{Z}[x]$) a

$$h_a = \tilde{h}_a(u_1, \ldots, u_n).$$

Podle Věty 11.2 existuje polynom $g \in (\mathbb{Z}[x])[z_1, \ldots, z_n]$ takový, že $\tilde{h}_a = g(s_1, \ldots, s_n)$. Z Viètových vztahů plyne, že $s_i(u_1, \ldots, u_n) \in \mathbb{R}$, a tedy

$$h_a = g(s_1(u_1, \ldots, u_n), \ldots, s_n(u_1, \ldots, u_n)) \in \mathbb{R}[x].$$

Přitom stupeň polynomu h_a je

$$\deg h_a = \binom{n}{2} = \frac{2^k m \cdot (2^k m - 1)}{2} = 2^{k-1} m (2^k m - 1),$$

takže můžeme použít indukční předpoklad a dostáváme, že h_a má kořen v $\mathbb{C}$. Shrnuto, dokázali jsme, že pro každé $a \in \mathbb{Z}$ existují nějaká $i < j$ taková, že $u_i + u_j + a u_i u_j \in \mathbb{C}$. Takových a je nekonečně mnoho, ale dvojic indexů je jen konečně mnoho, musí tedy existovat dvojice $i < j$, která se opakuje aspoň dvakrát (dokonce nekonečněkrát). Označme příslušná čísla a, b, tj.

$$u_i + u_j + a u_i u_j \in \mathbb{C} \quad \text{a} \quad u_i + u_j + b u_i u_j \in \mathbb{C}.$$

Odečtením obou výrazů vidíme, že $(a-b)u_iu_j \in \mathbb{C}$, čili také $u_iu_j \in \mathbb{C}$ a $u_i + u_j \in \mathbb{C}$. Z toho plyne, že oba kořeny u_i, u_j jsou komplexní, neboť

$$(x-u_i)(x-u_j) = x^2 - (u_i+u_j)x + u_iu_j$$

a podle Lemmatu 12.3 víme, že komplexní kvadratický polynom má nutně komplexní kořeny. $\square$

Grupy

13. Pojem grupy

13.1. **Základní vlastnosti permutací.**

Než se dostaneme k samotné definici grupy, shrneme základní poznatky o permutacích, které by měl typický čtenář znát ze základního kurzu lineární algebry nebo diskrétní matematiky, a doplníme je o pojem konjugace.

Permutací na množině X rozumíme bijekci (vzájemně jednoznačné zobrazení) $X \to X$. Pro permutace π, σ na X definujeme operace $\circ$, $^{-1}$, id předpisy

- $\pi \circ \sigma : x \mapsto \pi(\sigma(x))$,
- $\pi^{-1} : x \mapsto$ (ten jediný) prvek y splňující $\pi(y) = x$,
- $id : x \mapsto x$.

Cyklus v permutaci π je posloupnost $a_1, \dots, a_k$ navzájem různých prvků množiny X splňující $\pi(a_1) = a_2$, $\pi(a_2) = a_3$, ..., $\pi(a_k) = a_1$. *Rozkladem na cykly* se rozumí zápis

$$(a_{11}\ a_{12}\ \dots\ a_{1k_1})(a_{21}\ a_{22}\ \dots\ a_{2k_2}) \cdots (a_{m1}\ a_{m2}\ \dots\ a_{mk_m}),$$

kde $a_{i1}, a_{i2}, \dots, a_{ik_i}$, $i = 1, \dots, m$, jsou pod dvou různé prvky. Cykly délky 1 se ze zápisu zpravidla vynechávají. (Je-li X konečná množina, rozklad na cykly jistě existuje; pro nekonečné množiny bychom museli povolit „nekonečné cykly“.)

Transpozicí rozumíme permutaci tvaru $(x\ y)$. Permutace na konečné množině se nazývá *sudá*, pokud se skládá ze sudého počtu transpozic, *lichá* v opačném případě (máme-li dva různé rozklady jedné permutace, mohou mít různé délky, ale, jak lze snadno nahlédnout, stejnou paritu). Definujeme *znaménko permutace*: $\operatorname{sgn}\pi = 1$, je-li π sudá, a $\operatorname{sgn}\pi = -1$, je-li π lichá. Z definice snadno plyne, že

$$\operatorname{sgn}(\pi \circ \sigma) = \operatorname{sgn}\pi \cdot \operatorname{sgn}\sigma \quad \text{a} \quad \operatorname{sgn}\pi^{-1} = \operatorname{sgn}\pi.$$

Z rozkladu $(a_1\ a_2\ \dots\ a_k) = (a_1\ a_k) \circ \dots \circ (a_1\ a_3) \circ (a_1\ a_2)$ vidíme, že cykly sudé délky jsou liché a naopak, a že

$$\operatorname{sgn}\pi = (-1)^{n - \text{počet cyklů v } \pi} = (-1)^{\text{počet cyklů v } \pi \text{ sudé délky}}.$$

Konjugace permutací je analogií pojmu podobnosti matic, který znáte z lineární algebry.

Definice. Permutace π, σ nazýváme *konjugované*, pokud existuje permutace ρ taková, že $\sigma = \rho \circ \pi \circ \rho^{-1}$.

Konjugace má velmi přirozenou interpretaci: pro

$$\pi = (a_{11}\ a_{12}\ \dots\ a_{1k_1}) \cdots (a_{m1}\ a_{m2} \dots\ a_{mk_m})$$

dostáváme

$$\rho \circ \pi \circ \rho^{-1} = (\rho(a_{11})\ \rho(a_{12})\ \dots\ \rho(a_{1k_1})) \cdots (\rho(a_{m1})\ \rho(a_{m2})\ \dots\ \rho(a_{mk_m})),$$

neboť pro každé i, j platí

$$(\rho \circ \pi \circ \rho^{-1})(\rho(a_{ij})) = \rho(\pi(a_{ij})) = \rho(a_{i(j \oplus 1)}),$$

kde $j \oplus 1 = j + 1$ pro $j < k_j$ a $k_j \oplus 1 = 1$. Konjugace podle ρ tedy funguje jako „kopírování“ zápisu podle pravidel daných permutací ρ, každý prvek a v zápise permutace π se přepíše na $\rho(a)$, přičemž struktura cyklů zůstane zachována.

Tvrzení 13.1 (konjugace pro permutace)**.** *Permutace π, σ jsou konjugované právě tehdy, když mají stejný počet cyklů každé délky (říká se také, že mají* stejný typ*).*

Důkaz. ($\Rightarrow$) Plyne bezprostředně z výše uvedeného výpočtu.

($\Leftarrow$) Jsou-li

$$\pi = (a_{11}\ a_{12}\ \dots\ a_{1k_1})(a_{21}\ a_{22}\ \dots\ a_{2k_2})\cdots(a_{m1}\ a_{m2}\dots\ a_{mk_m}),$$
$$\sigma = (b_{11}\ b_{12}\ \dots\ b_{1k_1})(b_{21}\ b_{22}\ \dots\ b_{2k_2})\cdots(b_{m1}\ b_{m2}\dots\ b_{mk_m}),$$

dvě permutace stejného typu, definujeme $\rho(a_{ij}) = b_{ij}$ a výše uvedeným výpočtem dostaneme $\sigma = \rho \circ \pi \circ \rho^{-1}$. □

13.2. Definice a příklady grup.

Hlavní motivací teorie grup je studium nejrůznějších typů symetrií a transformací matematických objektů. Pojem pochází z Galoisovy teorie a původně označoval množinu (skupinu, v Galoisově jazyce *le groupe*) permutací G uzavřenou na skládání, tj. splňující $\pi \circ \sigma \in G$ pro všechna $\pi, \sigma \in G$. Abstrakcí tohoto pojmu vznikla rozsáhlá větev algebry, zvaná teorie grup. Aplikace nachází všude, kde se vyskytuje pojem symetrie či transformace, především v kombinatorice (konečné grupy) a geometrii (maticové grupy).

Definice. *Grupou* rozumíme čtveřici $\mathbf{G} = (G, *, ', e)$, kde G je množina, na které jsou definovány binární operace $*$, unární operace $'$ a konstanta e splňující pro každé $a, b, c \in G$ následující podmínky:

$$a * (b * c) = (a * b) * c, \qquad a * e = e * a = a, \qquad a * a' = a' * a = e.$$

Grupu nazýváme *abelovskou*, pokud navíc pro všechna $a, b \in G$ platí

$$a * b = b * a.$$

Prvku e se říká *jednotka*, prvku a' *inverzní prvek* k prvku a.

Formálně rozlišujeme mezi množinou G, tzv. *nosnou množinou*, a čtveřicí $\mathbf{G} = (G, *, ', e)$, která navíc obsahuje informaci o algebraické struktuře definované na množině G. V konkrétních příkladech bývá typickou trojicí operací buď $+, -, 0$, pak hovoříme o *aditivním zápise* (a místo $x + (-y)$ píšeme $x - y$), anebo trojice $\cdot, ^{-1}, 1$, čemuž říkáme *multiplikativní zápis*.

Definice. Buď $\mathbf{G} = (G, *, ', e)$ grupa a $H \subseteq G$ podmnožina její nosné množiny taková, že $e \in H$ a pro každé $a, b \in H$ platí

$$a' \in H \quad \text{a} \quad a * b \in H.$$

Říkáme, že H je *uzavřena na grupové operace* a že *tvoří podgrupu* grupy $\mathbf{G}$. Čtveřici $\mathbf{H} = (H, *|_H, '|_H, e)$ pak nazýváme *podgrupou*, přičemž $|_H$ značí restrikci operací na množinu H. Značíme $\mathbf{H} \le \mathbf{G}$. Podgrupy $\mathbf{G}$ a $\{e\}$ nazýváme *nevlastní*.

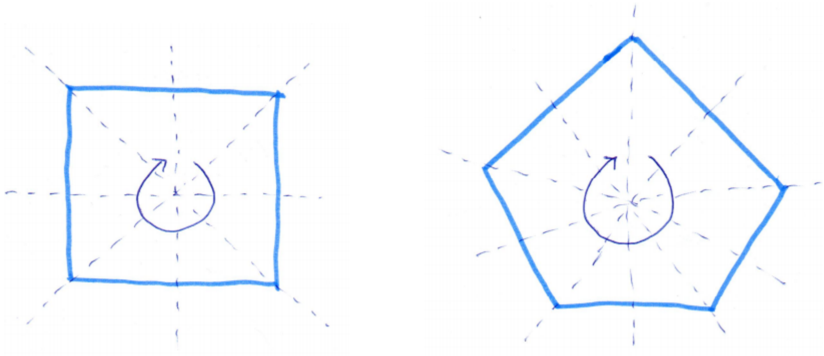
V matematice se vyskytují nejrůznější příklady grup, nicméně je možné identifikovat čtyři základní rodiny, které nacházejí asi největší využití: permutační grupy, maticové grupy, grupy geometrických zobrazení a číselné grupy.

Příklad. *Permutační grupy.* Základním příkladem je *symetrická grupa* sestávající z permutací na dané neprázdné množině X s operacemi $\circ$ skládání permutací, $^{-1}$ invertování permutací a konstantou $id : x \mapsto x$ (identické zobrazení), tj.

$$\mathbf{S}_X = (\{\pi : \ \pi \text{ je permutace na množině } X\}, \circ, ^{-1}, id).$$

Je-li $X = \{1, \dots, n\}$, pak místo $\mathbf{S}_X$ píšeme $\mathbf{S}_n$. Podgrupy symetrických grup se nazývají *permutační grupy*, např.

- *alternující grupa* $\mathbf{A}_n \le \mathbf{S}_n$ všech sudých permutací na n prvcích;
- *dihedrální grupa* $\mathbf{D}_{2n} \le \mathbf{S}_n$ všech permutací, které odpovídají symetriím pravidelného n-úhelníka vztaženým na jeho vrcholy očíslované po směru hodinových ručiček. Tyto permutace odpovídají n rotacím a n reflexím, proto značení $\mathbf{D}_{2n}$.
- nejrůznější grupy symetrií geometrických těles, automorfismů grafů a dalších matematických struktur, viz sekce 17.



Obrázek 12. Grupy $\mathbf{D}_8$ a $\mathbf{D}_{10}$.

Příklad. Speciálním případem permutačních grup jsou grupy zobrazení na různých typech geometrických prostorů (eukleidovské, afinní, projektivní apod.) zachovávajících jisté vlastnosti (afinní zobrazení, projektivní zobrazení apod.). Základním příkladem je *eukleidovská grupa* $\mathbf{E}_n$ sestávající ze všech izometrií (tj. zobrazení zachovávajících vzdálenosti) eukleidovského prostoru $\mathbb{R}^n$. Tzv. *Erlangenský program* formulovaný Felixem Kleinem v roce 1872 klasifikuje různé typy geometrií pomocí odpovídajících grup geometrických zobrazení.

Na grupy symetrií geometrických objektů daných konečně mnoha body lze nahlížet dvojím způsobem: jako na podgrupu eukleidovské grupy, sestávající z izometrií zachovávajících daný objekt, nebo jako na permutační grupu na bodech, které tento objekt určují. Typickým příkladem jsou dihedrální grupy $\mathbf{D}_{2n}$, na které lze nahlížet jako na podgrupy $\mathbf{E}_2$ nebo jako na podgrupy $\mathbf{S}_n$ (formálně jde o dvě izomorfní kopie téže grupy, viz sekce 15.2).

Příklad. *Maticové grupy.* Základním příkladem je *obecná lineární grupa* nad tělesem $\mathbf{T}$ sestávající z regulárních matic dané velikosti s operacemi $\cdot$ maticového násobení, $^{-1}$ maticového invertování a jednotkovou maticí jako jednotkou, tj.

$$\mathbf{GL}_n(\mathbf{T}) = (\{A : \ A \text{ je regulární matice } n \times n \text{ nad tělesem } \mathbf{T}\}, \cdot, ^{-1}, I),$$

Podgrupy lineárních grup se nazývají *maticové grupy*, např.

- *speciální lineární grupa* $\mathbf{SL}_n(\mathbf{T})$ všech matic s determinantem 1;
- *ortogonální grupa* $\mathbf{O}_n(\mathbf{T})$ všech ortogonálních matic, tj. takových A, které splňují $AA^T = I$ (nad tělesem $\mathbb{R}$ jde o matice, jejichž řádky, resp. sloupce, jsou ortonormální vektory vzhledem k standardnímu skalárnímu součinu).

Na maticové grupy lze také nahlížet jako na grupy geometrických zobrazení. Typickým příkladem je grupa $\mathbf{SO}_3(\mathbb{R})$, která sestává z ortogonálních matic s determinantem 1, a jak víme z lineární algebry, tyto matice vzájemně jednoznačně odpovídají rotacím eukleidovského prostoru $\mathbb{R}^3$ (formálně půjde opět o dvě izomorfní kopie téže grupy, jedna kopie je podgrupa $\mathbf{GL}_3(\mathbb{R})$, druhá je podgrupa $\mathbf{E}_3$).

V sekci 15.5 si ukážeme, že permutační a maticové grupy jsou v jistém smyslu univerzální příklady: každou grupu lze reprezentovat jako permutační grupu (*Cayleyova reprezentace*, Věta 15.10) a každou konečnou grupu lze reprezentovat jako maticovou grupu (Věta 15.11). Pro některé grupy je taková reprezentace přirozená, ale leckdy ne: příkladem je osmiprvková kvaternionová grupa.

Příklad. *Kvaternionová grupa* $\mathbf{Q}_8$ je definovaná na množině $\{\pm 1, \pm i, \pm j, \pm k\}$. Násobení je dáno vzorci

$$i^2 = j^2 = k^2 = -1, \quad ij = k, \quad jk = i, \quad ki = j,$$

a dále pravidly $xy = -(yx)$ a $(-x)y = x(-y) = -(xy)$ pro všechna $x, y \in \{i, j, k\}$.

Základním zdrojem příkladů abelovských grup jsou grupy odvozené od komutativních okruhů, zejména pak *číselné grupy*, odvozené od číselných oborů.

Příklad. Buď $\mathbf{R}$ okruh. Pak $(R, +, -, 0)$ je abelovská grupa, tzv. *aditivní grupa* okruhu $\mathbf{R}$. Důležité jsou zejména číselné grupy $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, a také grupy $\mathbb{Z}_n$ sestávající z čísel $0, \dots, n-1$ s operacemi modulo n.

Příklad. Buď $\mathbf{R}$ komutativní okruh s jednotkou, označme R^* množinu všech invertibilních prvků v $\mathbf{R}$. Pak $\mathbf{R}^* = (R^*, \cdot, ^{-1}, 1)$ je abelovská grupa, tzv. *multiplikativní grupa* okruhu $\mathbf{R}$. Skutečně jde o grupu: inverz invertibilního prvku je invertibilní (protože $(a^{-1}) \cdot a = 1$), součin dvou invertibilních prvků a, b je invertibilní (protože $(ab)(b^{-1}a^{-1}) = 1$) a grupové axiomy jsou obsaženy v definici okruhu.

- Je-li $\mathbf{R}$ těleso, pak $\mathbf{R}^* = (R \setminus \{0\}, \cdot, ^{-1}, 1)$.
- Pro polynomiální okruhy platí $\mathbf{R}[x]^* = \mathbf{R}^*$, protože invertibilní jsou právě konstantní polynomy invertibilní v $\mathbf{R}$.
- $\mathbb{Z}^* = (\{1, -1\}, \cdot, ^{-1}, 1)$.
- Prvky grupy $\mathbb{Z}_n^*$ jsou právě všechna čísla $a \in \{1, \dots, n-1\}$ nesoudělná s n. Soudělná čísla invertibilní nejsou: je-li $d \nparallel 1$ společný dělitel a, n, pak $d \mid (ab \bmod n)$ pro libovolné b, takže součin ab nikdy nemůže být 1. Naopak, jsou-li a, n nesoudělná, uvažujme Bézoutovy koeficienty u, v splňující $1 = \mathrm{NSD}(a, n) = ua + vn$. Podíváme-li se na rovnost modulo n, dostaneme $1 \equiv ua \pmod n$, a tedy $a^{-1} = u \bmod n$.

Zajímavé podgrupy číselných grup poskytuje jednotková kružnice v komplexní rovině.

Příklad. Komplexní jednotky, tj. množina $\{z \in \mathbb{C} : |z| = 1\}$, tvoří podgrupu grupy $\mathbb{C}^*$. Mezi jejími podgrupami dále jmenujme např.

- cyklotomické grupy $\mathbb{C}_n$ sestávající ze všech kořenů polynomu $x^n - 1$,
- *Prüferovu p-grupu* $\mathbb{C}_{p^\infty} = \bigcup_{k=1}^{\infty} \mathbb{C}_{p^k}$ sestávající ze všech komplexních čísel z splňujících $z^{p^k} = 1$ pro nějaké k.

Prüferovy grupy jsou oblíbeným protipříkladem na řadu vlastností.

Existuje řada dalších geometrických i algebraických konstrukcí abelovských grup, například grupy odvozené od eliptických křivek nebo třídové grupy prvoideálů v číselných tělesech. Některé z těchto konstrukcí mají významné aplikace v kryptografii (sekce 16.3), v praxi se hojně využívá například Diffie-Hellmanův protokol s grupami na eliptických křivkách nad konečnými tělesy.

Důležitou konstrukcí grup je direktní součin.

Definice. *Direktním součinem* grup $\mathbf{G}_i = (G_i, *_i, '^i, e_i)$, $i = 1, \dots, n$, rozumíme grupu

$$\prod_{i=1}^{n} \mathbf{G}_i = \mathbf{G}_1 \times \dots \times \mathbf{G}_n = (G_1 \times \dots \times G_n, *, ', e),$$

jejíž operace jsou definovány po složkách, tj.

$$(a_1, \dots, a_n) * (b_1, \dots, b_n) = (a_1 *_1 b_1, \dots, a_n *_n b_n),$$
$$(a_1, \dots, a_n)' = ((a_1)'^1, \dots, (a_n)'^n),$$
$$e = (e_1, \dots, e_n).$$

pro všechna $(a_1, \dots, a_n), (b_1, \dots, b_n) \in G_1 \times \ldots \times G_n$. Je snadné ověřit, že direktní součin splňuje všechny axiomy grup.

V případě, kdy $\mathbf{G}_1 = \ldots = \mathbf{G}_n = \mathbf{G}$, hovoříme o *direktní mocnině* a značíme ji $\mathbf{G}^n$.

Analogicky bychom mohli definovat také direktní součin okruhů, či jakéhokoliv jiného typu algebraických struktur.

Podobně jako v případě komutativních okruhů, definice grupy obsahuje pouze minimální množství podmínek. Následující tvrzení ukazuje několik aritmetických pravidel (mj. krácení, jednoznačnost jednotky a jednoznačnost inverzních prvků), které z definice snadno plynou a v dalším textu je budeme volně používat.

Tvrzení 13.2 (základní vlastnosti grup)**.** *Buď* $\mathbf{G} = (G, *, ', e)$ *grupa a* $a, b, c \in G$. *Pak*

(1) *jestliže* $a * c = b * c$ *nebo* $c * a = c * b$, *pak* $a = b$;
(2) *jestliže* $a * u = a$ *nebo* $u * a = a$ *pro nějaké* $u \in G$, *pak* $u = e$;
(3) *jestliže* $a * u = e$ *nebo* $u * a = e$ *pro nějaké* $u \in G$, *pak* $u = a'$;
(4) $(a')' = a$;
(5) $(a * b)' = b' * a'$.

Důkaz. (1) Je-li $a * c = b * c$, pak také $(a * c) * c' = (b * c) * c'$ a použitím všech tří axiomů dostaneme $(a * c) * c' = a * (c * c') = a * e = a$ a podobně $(b * c) * c' = b$. Tedy $a = b$. Analogicky pro $c * a = c * b$.

(2) Je-li $a * u = a = a * e$, krácením dostáváme $u = e$. Analogicky pro $u * a = a$.

(3) Je-li $a * u = e = a * a'$, krácením dostáváme $u = a'$. Analogicky pro $u * a = e$.

(4) Protože $a' * a = e$, z jednoznačnosti inverzních prvků dostáváme $a = (a')'$.

(5) Protože $(a*b)*(b'*a') = a*(b*b')*a' = a*e*a' = a*a' = e$, z jednoznačnosti inverzních prvků dostáváme $(a * b)' = b' * a'$. □

13.3. **Mocniny a řád prvku.**

Čtenář si snad již zažil, že se grupové operace značí nejrůznějšími způsoby. Nadále se budeme držet multiplikativního zápisu. Nebude-li výslovně uvedeno jinak, uvažujeme-li grupu $\mathbf{G}$, implicitně rozumíme $\mathbf{G} = (G, \cdot, ^{-1}, 1)$. Ze začátku je dobré si u všech výrazů rozmýšlet, jak bychom je přepsali do ostatních značení.

Nyní definujeme *mocniny*. Buď $\mathbf{G}$ grupa, $a \in G$, $n \in \mathbb{Z}$. Označme

$$a^n = \begin{cases} 1 & n = 0 \\ \underbrace{a \cdot a \cdot \ldots \cdot a}_{n} & n > 0 \\ \underbrace{a^{-1} \cdot a^{-1} \cdot \ldots \cdot a^{-1}}_{-n} & n < 0 \end{cases}$$

Tvrzení 13.3 (mocniny). *Buď* $\mathbf{G}$ *grupa,* $a, b \in G$ *a* $k, l \in \mathbb{Z}$. *Pak*

$$a^{k+l} = a^k \cdot a^l, \qquad a^{kl} = (a^k)^l = (a^l)^k$$

a je-li $\mathbf{G}$ *abelovská, pak navíc* $(ab)^k = a^k b^k$.

Důkaz. Pokud $k, l > 0$, ihned vidíme, že počet prvků a ve výrazech na obou stranách každé rovnosti je stejný. V případě záporných exponentů je třeba vzít v úvahu, že a a a^{-1} se navzájem pokrátí. Např. v první rovnosti, pro $k > 0 > l$, $|l| < |k|$, máme na levé straně součin $k + l$ prvků a, zatímco na pravé straně součin k prvků a a $-l$ prvků a^{-1}. Po vykrácení dostaneme rovnost obou výrazů. Ostatní případy se rozeberou podobně. □

V aditivním značení je mocninou výraz $a + \ldots + a$, resp. $(-a) + \ldots + (-a)$; tyto výrazy zkracujeme jako $n \cdot a$. Tvrzení 13.3 se pak přepíše jako

$$(k + l) \cdot a = k \cdot a + l \cdot a, \qquad (kl) \cdot a = k \cdot (l \cdot a), \qquad k \cdot (a + b) = k \cdot a + k \cdot b,$$

poslední rovnost samozřejmě platí pouze pro abelovské grupy. Pokud vám tyto podmínky připomínají definici vektorového prostoru, jste na správně stopě. Teorie abelovských grup je do značné míry teorií „vektorových prostorů nad $\mathbb{Z}$", neboli $\mathbb{Z}$-modulů, s řadou aplikací v teorii čísel. Tímto směrem se však v úvodním kurzu ubírat nebudeme.

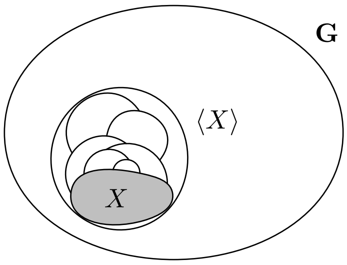
Definice. *Řádem grupy* $\mathbf{G}$ se rozumí počet prvků její nosné množiny, značíme jej $|\mathbf{G}|$ (tj., formálně vzato, $|\mathbf{G}| = |G|$).

Řádem prvku a v grupě $\mathbf{G}$ se rozumí nejmenší $n \in \mathbb{N}$ takové, že $a^n = 1$, pokud takové n existuje, resp. ∞ v opačném případě. Značíme jej $\mathrm{ord}(a)$.

V Tvrzení 14.6 si ukážeme, že řád prvku je roven řádu jisté podgrupy, ale zatím si vystačíme s definicí pomocí mocnin.

Příklad. Pokud mluvíme o řádu jistého prvku, je třeba říci, v které grupě!

- $\mathrm{ord}(2) = 7$ v grupě $\mathbb{Z}_7$, protože $7 \cdot 2 \equiv 0 \pmod 7$, ale $n \cdot 2 \not\equiv 0 \pmod 7$ pro $n = 1, \ldots, 6$;



Obrázek 13. Ilustrace generování podgrupy $\langle X\rangle_{\mathbf{G}}$.

- $\mathrm{ord}(2) = 3$ v grupě $\mathbb{Z}_7^*$, protože $2^3 \equiv 1 \pmod 7$, ale $2^n \not\equiv 1 \pmod 7$ pro $n = 1, 2$.

Příklad. V nekonečných grupách mohou řády vycházet všelijak:
- v grupě $\mathbb{Q}$ je $\mathrm{ord}(0) = 1$ a $\mathrm{ord}(a) = \infty$ pro všechna $a \neq 0$;
- v grupě $\mathbb{Q}^*$ je $\mathrm{ord}(1) = 1$, $\mathrm{ord}(-1) = 2$ a $\mathrm{ord}(a) = \infty$ pro všechna $a \neq \pm 1$;
- v grupě $\mathbb{C}^*$ existuje prvek libovolného řádu: $\mathrm{ord}(e^{2\pi i/k}) = k$.

Příklad. V konečných grupách řády nevycházejí všelijak:
- v grupě $\mathbb{Z}_6$ je $\mathrm{ord}(0) = 1$, $\mathrm{ord}(1) = 6$, $\mathrm{ord}(2) = 3$, $\mathrm{ord}(3) = 2$, $\mathrm{ord}(4) = 3$ a $\mathrm{ord}(5) = 6$, čili vyskytují se řády $1, 2, 3, 6$;
- v grupě $\mathbf{S}_3$ je $\mathrm{ord}(id) = 1$, $\mathrm{ord}((i\ j)) = 2$, $\mathrm{ord}((i\ j\ k)) = 3$, čili vyskytují se řády $1, 2, 3$.

Všimněte si, že řád každého prvku dělí řád celé grupy. To není náhoda, nýbrž pravidlo, které je speciálním případem Lagrangeovy věty (Věta 14.9), která je náplní příští sekce.

Tvrzení 13.4 (řád permutace). *Řád permutace v grupě $\mathbf{S}_n$ je roven nejmenšímu společnému násobku délek jejích cyklů.*

Důkaz. Cyklus délky n má řád n. Jsou-li $C_1, \ldots, C_m$ disjunktní cykly v π, pak $\pi^k = (C_1 \circ \ldots \circ C_m)^k = C_1^k \circ \ldots \circ C_m^k$. Z toho plyne, že $(C_1 \circ \ldots \circ C_m)^k = id$ právě tehdy, když je k násobkem všech délek cyklů. Čili řád je roven nejmenšímu společnému násobku. □

14. Podgrupy

14.1. **Generátory.**

Lemma 14.1. *Průnik podgrup je podgrupa.*

Důkaz. Buď $\mathbf{G}$ grupa, uvažujme podgrupy $\mathbf{H}_i$, $i \in I$, a označme $H = \bigcap_{i \in I} H_i$. Dokážeme, že je množina H uzavřená na grupové operace. Protože $1 \in H_i$ pro všechna $i \in I$, bude 1 náležet i jejich průniku. Nyní uvažujme $a, b \in H$. Tyto leží v každém H_i a díky uzavřenosti na operace tam leží také prvky a^{-1} a $a \cdot b$. Takže tyto prvky leží i v průniku všech H_i, čili v H. □

Definice. Uvažujme podmnožinu $X \subseteq G$ grupy $\mathbf{G}$. Podgrupou *generovanou množinou* X rozumíme nejmenší podgrupu (vzhledem k inkluzi) grupy $\mathbf{G}$ obsahující podmnožinu X, značíme ji $\langle X\rangle_{\mathbf{G}}$.

Taková podgrupa jistě existuje: stačí vzít průnik všech podgrup obsahujících množinu X, tj.

$$\langle X\rangle_{\mathbf{G}} = \bigcap\{H : X \subseteq H,\ \mathbf{H} \leq \mathbf{G}\}.$$

Podle předchozího lemmatu jde skutečně o podgrupu, mezi všemi podgrupami obsahujícími množinu X bude jistě nejmenší.

Jak najít podgrupu generovanou danou množinou? Pro konečné grupy lze v principu použít následující postup: začneme s prvky množiny X a postupně přidáváme všemožné součiny a inverzy. Ve chvíli, kdy nejsme schopni získat žádné nové prvky, naše množina je uzavřená na grupové operace a podgrupa je nalezena (viz obrázek). Leckdy je však efektivnější použít následující tvrzení.

Tvrzení 14.2. *Buď* $\mathbf{G}$ *grupa a* $\emptyset \neq X \subseteq G$*. Pak*

$$\langle X\rangle_{\mathbf{G}} = \{a_1^{k_1} \cdot a_2^{k_2} \cdot \ldots \cdot a_n^{k_n} : \ n \in \mathbb{N},\ a_1, \ldots, a_n \in X,\ k_1, \ldots, k_n \in \mathbb{Z}\}.$$

Důkaz. Důkaz je podobný důkazu Tvrzení 3.1. Označme M množinu na pravé straně rovnosti. Je potřeba dokázat, že množina M

(1) tvoří podgrupu,
(2) obsahuje X,
(3) je nejmenší podmnožinou grupy $\mathbf{G}$ splňující tyto podmínky.

(1) Součin dvou prvků z M je jistě v M, jednotka $1 = a^0$ je tam také, a uzavřenost na inverzy plyne ze vztahu $(a_1^{k_1} \cdot \ldots \cdot a_n^{k_n})^{-1} = a_n^{-k_n} \cdot \ldots \cdot a_1^{-k_1} \in M$.

(2) Volbou $n = 1, k_1 = 1$ dostaneme libovolný prvek X.

(3) Uvažujme libovolnou podgrupu $\mathbf{H}$ obsahující X. Tato podgrupa musí obsahovat všechny mocniny a^i, $a \in X$, i jejich libovolné násobky, čili celé M. □

Obecné tvrzení o tvaru podgrup generovaných danou podmnožinou má dva důležité speciální případy.

Důsledek 14.3. *Buď* $\mathbf{G}$ *grupa a* $a \in G$*. Pak* $\langle a\rangle_{\mathbf{G}} = \{a^k : \ k \in \mathbb{Z}\}$.

Důsledek 14.4. *Buď* $\mathbf{G}$ *abelovská grupa a* $u_1, \ldots, u_n \in G$*. Pak*

$$\langle u_1, \ldots, u_n\rangle_{\mathbf{G}} = \{u_1^{k_1} \cdot u_2^{k_2} \cdot \ldots \cdot u_n^{k_n} : \ k_1, \ldots, k_n \in \mathbb{Z}\}.$$

Vidíme, že v abelovských grupách je generování podgrup podobné jako generování vektorových prostorů: v aditivním zápise, tj. pro abelovskou grupu $\mathbf{G} = (G, +, -, 0)$, dostáváme

$$\langle u_1, \ldots, u_n\rangle_{\mathbf{G}} = \{k_1u_1 + k_2u_2 + \ldots + k_nu_n : \ k_1, \ldots, k_n \in \mathbb{Z}\}.$$

(S nezávislostí a jednoznačností zápisu je to složitější, ale tím se v této učebnici zabývat nebudeme.)

Příklad. Důležitým typem úlohy je zjistit, jakou podgrupu generuje daná podmnožina. Například:

- $\langle \frac{3}{4}, \frac{1}{3}\rangle_{\mathbb{Q}} = \{k\frac{3}{4} + l\frac{1}{3} : \ k, l \in \mathbb{Z}\} = \{\frac{k}{12} : \ k \in \mathbb{Z}\} = \langle \frac{1}{12}\rangle_{\mathbb{Q}}$. První a poslední rovnost plynou z Důsledku 14.4. K důkazu prostřední je potřeba si uvědomit, že na jednu stranu $\frac{3}{4}, \frac{1}{3} \in \langle \frac{1}{12}\rangle_{\mathbb{Q}}$, a na druhou stranu $\frac{1}{12} = \frac{3}{4} - 2 \cdot \frac{1}{3}$, a tedy $\frac{1}{12} \in \langle \frac{3}{4}, \frac{1}{3}\rangle_{\mathbb{Q}}$.
- $\langle \frac{3}{4}, \frac{1}{3}\rangle_{\mathbb{Q}^*} = \{(\frac{3}{4})^k \cdot (\frac{1}{3})^l : \ k, l \in \mathbb{Z}\} = \{3^k \cdot 4^l : \ k, l \in \mathbb{Z}\}$.

Příklad. Jiným důležitým typem úlohy je, najít k dané grupě $\mathbf{G}$ co nejmenší množinu generátorů, tj. podmnožinu $X \subseteq G$ takovou, že $\mathbf{G} = \langle X\rangle_{\mathbf{G}}$. Například:

- $\mathbb{Z} = \langle 1\rangle = \langle -1\rangle$, $\mathbb{Z}^* = \langle -1\rangle$, $\mathbb{Q}^* = \langle -1, \text{prvočísla}\rangle$.
- $\mathbb{Z}_n = \langle 1\rangle$, ale najít malou generující množinu grupy $\mathbb{Z}_n^*$ není obecně snadné. Například, $\mathbb{Z}_7^* = \langle 3\rangle$, ale $\mathbb{Z}_8^* = \langle 3, 5\rangle$ a nelze ji nagenerovat jedním prvkem.
- Pro některé grupy neexistuje minimální množina generátorů. Např. $\mathbb{Q} = \langle \frac{1}{n} : n \in \mathbb{N}\rangle$, kde lze z každé generující podmnožiny nějaký prvek vypustit.

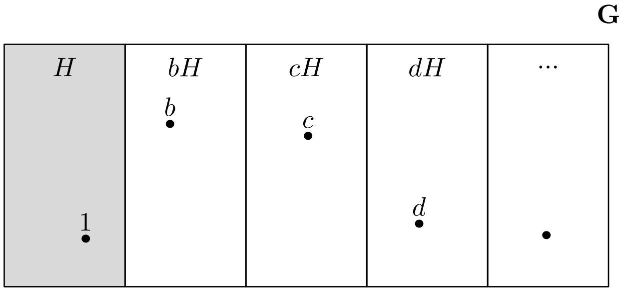
Tvrzení 14.5 (generátory permutačních grup)**.**

(1) *Grupa* $\mathbf{S}_n$ *je generovaná množinou všech transpozic.*
(2) *Grupa* $\mathbf{A}_n$ *je generovaná množinou všech trojcyklů.*

Důkaz. (1) Danou permutaci nejprve napíšeme jako složení svých cyklů a každý cyklus pak rozložíme podle následujícího vzoru:

$$(a_1\ a_2\ \ldots\ a_k) = (a_1\ a_k) \circ \ldots \circ (a_1\ a_3) \circ (a_1\ a_2).$$

(2) Danou sudou permutaci nejprve napíšeme jako složení sudého počtu transpozic a ty seskupíme do sousedících dvojic. Pokud jsou sousedící transpozice stejné, můžeme je vypustit. Pokud mají společný jeden prvek, pak $(i\ j) \circ (j\ k) = (i\ j\ k)$. A jsou-li disjunktní, pak $(i\ j) \circ (k\ l) = (k\ i\ l) \circ (i\ j\ k)$. Tímto způsobem přepíšeme rozklad na transpozice na složení trojcyklů. □



OBRÁZEK 14. Rozklad grupy **G** podle podgrupy **H** a jeho transverzála.

Uvedené množiny generátorů nejsou optimální. Například grupu $\mathbf{S}_n$ je možné generovat jednou transpozicí a jedním n-cyklem. Více příkladů najdete v cvičeních.

Příklad. Ukážeme, že

$$\mathbf{S}_n = \langle (1\ 2), (1\ 2\ \dots\ n) \rangle.$$

Díky Tvrzení 14.5 stačí dokázat, že lze nagenerovat všechny transpozice. Nejprve nagenerujeme transpozice $(k\ k+1)$, $k = 1, \dots, n-1$: induktivně

$$(k+1\ k+2) = (1\ 2\ \dots\ n) \circ (k\ k+1) \circ (1\ 2\ \dots\ n)^{-1}.$$

Dále, pro každé k nagenerujeme ostatní transpozice $(k\ k+i)$, $i > 0$: opět induktivně

$$(k\ k+i+1) = (k+i\ k+i+1) \circ (k\ k+i) \circ (k+i\ k+i+1)^{-1}.$$

Na závěr dokážeme, že řád prvku (definovaný pomocí mocnin) je roven řádu podgrupy jím generované.

Tvrzení 14.6 (řád prvku vs. řád podgrupy)**.** *Buď* **G** *grupa a* $a \in G$*. Pak*

$$\mathrm{ord}(a) = |\langle a \rangle_{\mathbf{G}}|.$$

Důkaz. Podle Důsledku 14.3 je $\langle a \rangle_{\mathbf{G}} = \{a^k : k \in \mathbb{Z}\}$. Všimněte si, že $a^i = a^j$ právě tehdy, když $a^{i-j} = 1$. Je-li $\mathrm{ord}(a) = \infty$, pak žádné $n \neq 0$ s vlastností $a^n = 1$ neexistuje, čili mocniny a^k jsou po dvou různé a podgrupa $\langle a \rangle$ je nekonečná. Je-li $\mathrm{ord}(a) = n < \infty$, pak jsou mocniny $a^0, a^1, \dots, a^{n-1}$ po dvou různé, ovšem další mocniny nové prvky nepřidají: $a^n = a^0 = 1$, $a^{n+1} = a^n \cdot a^1 = a^1$, $a^{n+2} = a^n \cdot a^2 = a^2$ atd., obecně $a^{qn+r} = (a^n)^q \cdot a^r = a^r$. Tedy $\langle a \rangle_{\mathbf{G}} = \{a^0, a^1, \dots, a^{n-1}\}$ obsahuje přesně n prvků. □

14.2. **Lagrangeova věta.**

Základní aritmetickou vlastností konečných grup je fakt, že řády podgrup dělí řád celé grupy, tj.

$$\mathbf{H} \leq \mathbf{G} \quad \Rightarrow \quad |\mathbf{H}| \text{ dělí } |\mathbf{G}|.$$

Speciálně, díky Tvrzení 14.6, řád prvku dělí řád celé grupy.

Myšlenka důkazu Lagrangeovy věty není složitá: celou grupu **G** rozložíme na několik podmnožin, které jsou po dvou disjunktní a stejně velké jako daná podgrupa **H**. Počet prvků grupy **G** tak bude roven počtu prvků **H** krát počet těchto podmnožin. Nesamozřejmou částí důkazu je konstrukce tohoto rozkladu.

Definice. Buď **G** grupa a **H** její podgrupa:

- množiny $aH = \{ah : h \in H\}$, kde $a \in G$, se nazývají *rozkladové třídy* podgrupy **H**;
- podmnožina $T \subseteq G$ s vlastností $|T \cap aH| = 1$ pro každé $a \in G$ se nazývá *transverzála* rozkladu **G** podle **H**;
- počet rozkladových tříd se nazývá *index* podgrupy **H** v grupě **G** a značí se

$$[\mathbf{G} : \mathbf{H}] = |\{aH : a \in G\}|.$$

Pojmy, které jsme definovali, se někdy používají s přívlastkem *levý*, tj. levé rozkladové třídy, levá transverzála, levý index. *Pravými rozkladovými třídami* pak rozumíme množiny $Ha = \{ha : h \in H\}$ a ostatní pojmy se definují analogicky. Levé a pravé varianty mohou být stejné či různé (viz příklady níže), ale jak uvidíme, počet rozkladových tříd, tj. index, vyjde z obou stran stejně.

Příklad. Buď $\mathbf{G} = \mathbb{Z}$ a $\mathbf{H} = \{h \in \mathbb{Z} : \ n \mid h\}$. Rozkladovou třídu určenou prvkem $a \in \mathbb{Z}$ můžeme vyjádřit

$$aH = \{a + h : \ h \in H\} = \{a + nk : \ k \in \mathbb{Z}\} = \{u \in \mathbb{Z} : \ u \equiv a \pmod{n}\}.$$

Dvě rozkladové třídy aH, bH jsou buď stejné, nebo disjunktní, přičemž $aH = bH$ právě tehdy, když $a \equiv b \pmod{n}$. Dostáváme tak n různých po dvou disjunktních rozkladových tříd, $[\mathbf{G} : \mathbf{H}] = n$. Jako transverzálu lze zvolit např. $T = \{0, \dots, n-1\}$, množinu všech možných zbytků po dělení n.

Příklad. Buď $\mathbf{G} = \mathbf{S}_n$ a $\mathbf{H} = \mathbf{A}_n$. Pak $\pi A_n = A_n \pi = A_n$ pro libovolnou π sudou a $\pi A_n = A_n \pi$ sestává ze všech lichých permutací pro libovolnou π lichou. Grupa $\mathbf{S}_n$ se tedy rozkládá na dvě disjunktní rozkladové třídy (levé i pravé jsou stejné), $[\mathbf{S}_n : \mathbf{A}_n] = 2$ a jako transverzálu lze zvolit např. $T = \{id, (1\ 2)\}$.

Levé a pravé rozkladové třídy nemusí být vždy stejné, nejmenším příkladem je následující situace.

Příklad. Buď $\mathbf{G} = \mathbf{S}_3$ a $\mathbf{H} = \{id, (1\ 2)\}$. Snadno spočteme, že levý i pravý rozklad obsahuje tři dvouprvkové třídy, avšak

$$(1\ 3)H = \{(1\ 3), (1\ 2\ 3)\}, \quad \text{ale} \quad H(1\ 3) = \{(1\ 3), (1\ 3\ 2)\}.$$

Lagrangeovu větu dokážeme pomocí dvou základních vlastností rozkladů: za prvé, různé rozkladové třídy jsou disjunktní, a za druhé, všechny rozkladové třídy jsou stejně velké. Analogická tvrzení platí i pro pravé rozkladové třídy.

Lemma 14.7 (disjunkce rozkladových tříd)**.** *Buď* $\mathbf{G}$ *grupa a* $\mathbf{H}$ *její podgrupa. Pro každé* $a, b \in G$ *platí buď* $aH = bH$*, nebo* $aH \cap bH = \emptyset$*.*

Důkaz. Předpokládejme $aH \cap bH \neq \emptyset$, dokážeme, že $aH = bH$. Uvažujme $c \in aH \cap bH$ a napišme $c = ah_1 = bh_2$ pro nějaká $h_1, h_2 \in H$. Pak pro každé $ah \in aH$ platí

$$ah = ch_1^{-1}h = b\underbrace{h_2h_1^{-1}h}_{\in H} \in bH$$

a podobně pro každé $bh \in bH$ platí

$$bh = ch_2^{-1}h = a\underbrace{h_1h_2^{-1}h}_{\in H} \in aH.$$

Tedy $aH = bH$. $\square$

Lemma 14.8 (velikost rozkladových tříd)**.** *Buď* $\mathbf{G}$ *grupa a* $\mathbf{H}$ *její podgrupa. Pro každé* $a \in G$ *platí* $|aH| = |H|$*.*

Důkaz. Uvažujme zobrazení $f : G \to G$ definované $f(x) = ax$. Toto zobrazení je prosté: kdyby $ax = f(x) = f(y) = ay$, krácením dostaneme $x = y$. Přitom $f(H) = aH$, tedy $f|_H$ je bijekce mezi H a aH, takže jsou tyto množiny stejně velké. $\square$

Lagrangeovu větu lze formulovat i pro nekonečné grupy, s použitím kardinálních čísel pro označení velikostí množin. Čtenáři, který kardinální čísla neviděl, postačí k porozumění tvrzení vlastnost, že součin velikostí množin je roven velikosti kartézského součinu, tj. $|X| \cdot |Y| = |X \times Y|$. Důkaz věty je pro konečné i nekonečné množiny stejný.

Věta 14.9 (Lagrangeova věta)**.** *Buď* $\mathbf{G}$ *grupa a* $\mathbf{H}$ *její podgrupa. Pak*

$$|\mathbf{G}| = |\mathbf{H}| \cdot [\mathbf{G} : \mathbf{H}].$$

Důkaz. Zvolme nějakou transverzálu T a napišme

$$G = \bigcup_{a\in T} aH.$$

Podle Lemmatu 14.7 jde o disjunktní sjednocení, takže počet prvků lze spočítat jako součet velikostí jednotlivých podmnožin:

$$|\mathbf{G}| = \sum_{a\in T} |aH| = \sum_{a\in T} |H| = |T| \cdot |H| = [\mathbf{G} : \mathbf{H}] \cdot |\mathbf{H}|.$$

V druhé rovnosti jsme použili Lemma 14.8 a ve čtvrté rovnosti jsme použili vztah $|T| = [\mathbf{G} : \mathbf{H}]$, který plyne z Lemmatu 14.7. □

Příklad. Speciálním případem Lagrangeovy věty je *Eulerova věta* (Věta 4.10), jejíž elementární důkaz jsme předvedli v sekci 4.4. Zvolme $\mathbf{G} = \mathbb{Z}_n^*$ a $a \in \mathbb{Z}_n^*$, tedy a je celé číslo nesoudělné s n. Pak $\operatorname{ord}(a)$ dělí $|\mathbb{Z}_n^*| = \varphi(n)$, čili $\varphi(n) = k \cdot \operatorname{ord}(a)$ pro nějaké k. V grupě $\mathbb{Z}_n^*$ tedy platí

$$a^{\varphi(n)} = (a^{\operatorname{ord}(a)})^k = 1^k = 1,$$

čili v jazyce teorie čísel $a^{\varphi(n)} \equiv 1 \pmod{n}$.

Na závěr podsekce ukážeme důležité kritérium, podle kterého se snadno pozná, zda jsou dvě rozkladové třídy stejné.

Tvrzení 14.10 (rovnost rozkladových tříd)**.** *Buď* $\mathbf{G}$ *grupa a* $\mathbf{H}$ *její podgrupa. Pro každé* $a, b \in G$ *platí*

(1) $aH = bH$ *právě tehdy, když* $a^{-1}b \in H$;
(2) $Ha = Hb$ *právě tehdy, když* $ab^{-1} \in H$.

Důkaz. (1) ($\Rightarrow$) Protože $aH = bH$, máme $b \in aH$, a tedy $b = ah$ pro nějaké $h \in H$. Tudíž $a^{-1}b = h \in H$. ($\Leftarrow$) Jestliže $a^{-1}b \in H$, pak pro každé $ah \in aH$ platí

$$ah = bb^{-1}ah = b\underbrace{(a^{-1}b)^{-1}h}_{\in H} \in bH$$

a podobně pro každé $bh \in bH$ platí

$$bh = a\underbrace{(a^{-1}b)h}_{\in H} \in aH.$$

Tedy $aH = bH$. (2) se dokáže analogicky. □

Poznámka. Nyní již můžeme dokázat, že levé a pravé rozklady jsou stejně velké. Ukážeme, že zobrazení

$$aH \mapsto Ha^{-1}.$$

je bijektivní. Nejprve musíme dokázat, že jsme korektně definovali zobrazení: mohlo by se stát, že téže rozkladové třídě $aH = bH$ se snažíme přiřadit dvě různé hodnoty $Ha^{-1} \neq Hb^{-1}$. Podle Tvrzení 14.10

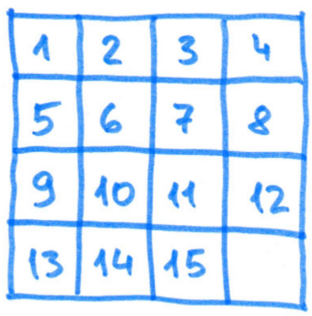
$$aH = bH \;\Leftrightarrow\; a^{-1}b \in H \;\Leftrightarrow\; (a^{-1}b)^{-1} = b^{-1}a \in H \;\Leftrightarrow\; Ha^{-1} = Hb^{-1},$$

a tedy zobrazení je nejen dobře definované, ale také prosté. Evidentně je i na.

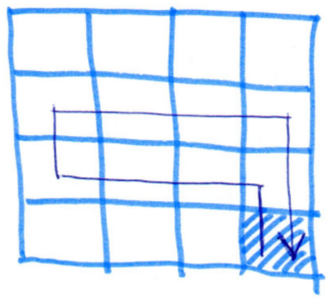
14.3. Loydova patnáctka a generátory alternující grupy.

Loydova patnáctka je známý hlavolam, ve kterém se po hrací ploše ve formě čtvercového pole 4×4 posouvá patnáct čtvercových kostiček s čísly 1, ..., 15. V jednom kroku je možné posunout na prázdné pole jednu ze sousedních kostiček. Cílem je stav, kde jsou kostičky seřazeny vzestupně po řádcích zleva doprava, shora dolů, přičemž prázdné políčko je vpravo dole (viz obrázek). Otázka zní: pro které počáteční stavy lze kostičky přesunout do cílového stavu?

Matematicky lze hlavolam popsat následujícím způsobem. Místo prázdného pole budeme uvažovat kostičku s číslem 16. Označme pole hrací plochy jako v cílovém stavu (pole vpravo dole bude mít číslo 16). *Stav hry* lze popsat jako permutaci $\pi \in S_{16}$, kde na poli číslo i je



Obrázek 15. Loydova patnáctka a číslování polí



Obrázek 16. Loydova patnáctka: průchod ze základního stavu

kostička s číslem $\pi(i)$. *Cílový stav* je popsán identickou permutací. V jednom *kroku* je možné prohodit kostičku číslo 16 se sousední kostičkou, čili ze stavu daného permutací π se dostaneme do stavu daného permutací $\pi \circ (i\ j)$, kde i, j jsou sousední pole a $\pi(i) = 16$. Všimněte si, že v jednom kroku se vždy změní znaménko stavové permutace.

Nejprve ukážeme, které stavy nemají řešení. Popíšeme jistou vlastnost stavu, tzv. *invariant*, kterou zachovává každý krok hry. Stavy, které mají jinou hodnotu invariantu než cílový stav, nemohou být řešitelné. Označme $\mathrm{ny}(\pi)$ tzv. *newyorskou vzdálenost* prázdného pole od pravého dolního pole ve stavu daném permutací π (tj. nejmenší počet tahů, který je potřeba na přesunutí prázdného pole na pozici 16). Označme

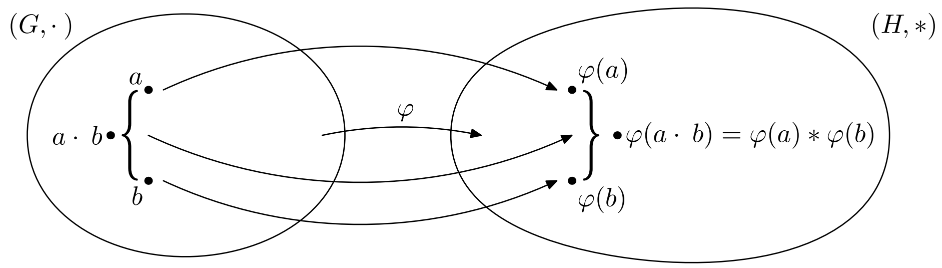
$$I(\pi) = \mathrm{sgn}(\pi) \cdot (-1)^{\mathrm{ny}(\pi)}$$

součin znaménka permutace a (multiplikativní) parity newyorské vzdálenosti prázdného pole. Vidíme, že $I(\pi)$ se v žádném kroku nemění: jeden krok změní znaménko permutace, ale také paritu newyorské vzdálenosti, čili I je invariantem. Vzhledem k tomu, že $I(id) = 1$, stavy dané permutací π s $I(\pi) = -1$ určitě řešitelné nejsou.

Těžší je dokázat, že všechny stavy s $I(\pi) = 1$ jsou řešitelné. Stav nazveme *základní*, pokud je prázdné pole vpravo dole, tj. pro jeho stavovou permutaci platí $\pi(16) = 16$. Bez újmy na obecnosti se lze soustředit na řešitelnost základních stavů: libovolný jiný stav, daný permutací σ, lze několika kroky převést na základní stav daný permutací σ', přičemž $I(\sigma) = I(\sigma') = \mathrm{sgn}(\sigma')$. Otázka tedy zní, zda jsou všechny základní stavy dané sudou permutací řešitelné.

Všimněte si, že ze základního stavu daného permutací π se lze dostat několika kroky do základních stavů daných permutacemi

- $\pi \circ (9\ 10\ 11\ 12\ 15\ 14\ 13)$ — prázdným polem objedeme dolní obdélník 2×4 (po směru hodinových ručiček).
- $\pi \circ (5\ 6\ 7\ 8\ 11\ 10\ 9)$ — prázdné pole posuneme o 1 nahoru, objedeme s ním prostřední obdélník 2×4 a vrátíme jej dolů (viz obrázek).



Obrázek 17. Homomorfismus $\varphi : \mathbf{G} \to \mathbf{H}$.

- $\pi \circ (1\ 2\ 3\ 4\ 7\ 6\ 5)$ — prázdné pole posuneme o 2 nahoru, objedeme s ním horní obdélník 2×4 a vrátíme jej dolů.

Které základní stavy lze převést na cílový, daný permutací *id*? Nebo raději obráceně, na které stavy se lze dostat z identity? Jistě na ty, které jsou dané permutacemi, které poskládáme z tří výše uvedených permutací. Problém řešitelnosti Loydovy patnáctky se tedy redukuje na úlohu, zda

$$\mathbf{A}_{15} = \langle (1\ 2\ 3\ 4\ 7\ 6\ 5), (5\ 6\ 7\ 8\ 11\ 10\ 9), (9\ 10\ 11\ 12\ 15\ 14\ 13) \rangle$$

(vzhledem k tomu, že 16 je pevná, můžeme permutace určující základní stavy považovat za prvky $\mathbf{A}_{15}$). Tuto úlohu necháváme jako cvičení ve stylu sekce 14.1.

15. Grupové homomorfismy

15.1. Základní vlastnosti.

Podobně jako pro okruhy či vektorové prostory, grupové homomorfismy jsou zobrazení, která zachovávají základní grupové operace.

V celé sekci budeme uvažovat dvě grupy $\mathbf{G} = (G, \cdot, ^{-1}, 1)$ a $\mathbf{H} = (H, *, ', e)$.

Definice. Buď $\mathbf{G}, \mathbf{H}$ grupy. Zobrazení $\varphi : G \to H$ je *homomorfismem* těchto grup, pokud pro každé $a, b \in G$ platí

$$\varphi(a \cdot b) = \varphi(a) * \varphi(b), \qquad \varphi(a^{-1}) = \varphi(a)', \qquad \varphi(1) = e.$$

Fakt, že je zobrazení mezi grupami homomorfismem, budeme zapisovat $\varphi : \mathbf{G} \to \mathbf{H}$.

Hned na začátku je dobré si všimnout, že druhá a třetí rovnost plynou z té první, což znatelně zjednodušuje ověřování, zda je dané zobrazení homomorfismem.

Lemma 15.1. *Buď $\mathbf{G}, \mathbf{H}$ grupy a $\varphi : G \to H$ zobrazení. Pak φ je homomorfismem těchto grup právě tehdy, když $\varphi(a \cdot b) = \varphi(a) * \varphi(b)$ pro všechna $a, b \in G$.*

Důkaz. Nejprve dokážeme, že $\varphi(1) = e$. Protože $e * \varphi(1) = \varphi(1) = \varphi(1 \cdot 1) = \varphi(1) * \varphi(1)$, krácením dostaneme $\varphi(1) = e$. Dále dokážeme $\varphi(a^{-1}) = \varphi(a)'$ pro každé $a \in G$. Protože $e = \varphi(1) = \varphi(a \cdot a^{-1}) = \varphi(a) * \varphi(a^{-1})$, z jednoznačnosti inverzních prvků v grupě $\mathbf{H}$ plyne $\varphi(a^{-1}) = \varphi(a)'$. □

Buď $\varphi : \mathbf{G} \to \mathbf{H}$ homomorfismus. Jeho *obrazem* nazýváme jeho obor hodnot, tj. množinu

$$\mathrm{Im}(\varphi) = \{\varphi(a) : \ a \in G\}.$$

Jeho *jádro* definujeme jako množinu

$$\mathrm{Ker}(\varphi) = \{a \in G : \ \varphi(a) = e\}.$$

Tvrzení 15.2 (jádro a obraz jsou podgrupy)**.** *Buď $\mathbf{G}, \mathbf{H}$ grupy a $\varphi : \mathbf{G} \to \mathbf{H}$ homomorfismus. Pak*

1. $\mathrm{Im}(\varphi)$ *tvoří podgrupu grupy* $\mathbf{H}$*;*
2. $\mathrm{Ker}(\varphi)$ *tvoří podrupu grupy* $\mathbf{G}$.

Důkaz. (1) $e \in \mathrm{Im}(\varphi)$, protože $e = \varphi(1)$. Pokud $\varphi(a), \varphi(b) \in \mathrm{Im}(\varphi)$, pak $\varphi(a)' = \varphi(a^{-1}) \in \mathrm{Im}(\varphi)$ a $\varphi(a) * \varphi(b) = \varphi(a \cdot b) \in \mathrm{Im}(\varphi)$.

(2) $1 \in \mathrm{Ker}(\varphi)$, protože $\varphi(1) = e$. Pokud $a, b \in \mathrm{Ker}(\varphi)$, pak a^{-1} a $a \cdot b$ také, protože $\varphi(a^{-1}) = \varphi(a)' = e' = e$ a $\varphi(a \cdot b) = \varphi(a) * \varphi(b) = e * e = e$. □

Tvrzení 15.3. *Buď* $\mathbf{G}, \mathbf{H}$ *grupy a* $\varphi : \mathbf{G} \to \mathbf{H}$ *homomorfismus. Pak* φ *je prostý právě tehdy, když* $\mathrm{Ker}(\varphi) = \{1\}$.

Důkaz. Je-li φ prosté, dva různé prvky se nemohou zobrazovat na e, takže $\mathrm{Ker}(\varphi)$ musí obsahovat jen jeden prvek, a tím je 1. Naopak, $\varphi(a) = \varphi(b)$ právě tehdy, když $e = \varphi(a) * \varphi(b)' = \varphi(a \cdot b^{-1})$, takže neprostá zobrazení obsahují nejednotkový prvek v jádru. □

Příklad. Řada známých zobrazení v matematice je homomorfismem jistých grup.

- Uvažujme zobrazení $z \mapsto |z|$ na komplexních číslech. Toto zobrazení je homomorfismem grup $\mathbb{C}^* \to \mathbb{R}^*$, protože $|a \cdot b| = |a| \cdot |b|$. Jeho jádrem je podgrupa komplexních jednotek, jeho obrazem podgrupa kladných čísel. Naopak, toto zobrazení není homomorfismem grup $\mathbb{C} \to \mathbb{R}$, protože obecně $|a + b| \neq |a| + |b|$.
- Uvažujme zobrazení $z \mapsto e^z$ na komplexních číslech. Toto zobrazení je homomorfismem grup $\mathbb{C} \to \mathbb{C}^*$, protože $e^{a+b} = e^a \cdot e^b$. Jeho jádrem je podgrupa $\langle 2\pi i \rangle = \{k \cdot 2\pi i : \ k \in \mathbb{Z}\}$, jeho obrazem celé $\mathbb{C}^*$.
- Uvažujme zobrazení $A \mapsto \det(A)$ na maticích. Toto zobrazení je homomorfismem grup $\mathbf{GL}_n(\mathbf{T}) \to \mathbf{T}^*$, protože $\det(A \cdot B) = \det(A) \cdot \det(B)$. Jeho jádrem je podgrupa $\mathbf{SL}_n(\mathbf{T})$, jeho obrazem celé $\mathbf{T}^*$.
- Uvažujme zobrazení $\pi \mapsto \mathrm{sgn}(\pi)$ na permutacích. Toto zobrazení je homomorfismem grup $\mathbf{S}_n \to \mathbb{Z}^*$, protože $\mathrm{sgn}(\pi \circ \sigma) = \mathrm{sgn}(\pi) \cdot \mathrm{sgn}(\sigma)$. Jeho jádrem je podgrupa $\mathbf{A}_n$, jeho obrazem celé $\mathbb{Z}^*$.

Homomorfismy jsou určeny svými hodnotami na generátorech: uvažujme homomorfismus $\varphi : \mathbf{G} \to \mathbf{H}$, nechť $\mathbf{G} = \langle X \rangle$ a označme hodnoty $\varphi(a) = h_a$ pro všechna $a \in X$. Obecný prvek grupy $\mathbf{G}$ lze podle Tvrzení 14.2 napsat ve tvaru $g = a_1^{k_1} \cdot \ldots \cdot a_n^{k_n}$, kde $a_1, \ldots, a_n \in X$ a $k_1, \ldots, k_n \in \mathbb{Z}$. Hodnota zobrazení pak bude

$$\varphi(g) = \varphi(a_1)^{k_1} \cdot \ldots \cdot \varphi(a_n)^{k_n} = h_{a_1}^{k_1} \cdot \ldots \cdot h_{a_n}^{k_n}.$$

Avšak pozor, na rozdíl od vektorových prostorů není možné volit obrazy generátorů libovolně, jak ukazuje například následující tvrzení.

Tvrzení 15.4 (řád prvku a jeho obrazu)**.** *Buď* $\varphi : \mathbf{G} \to \mathbf{H}$ *homomorfismus grup. Pak, pro každé* $a \in G$,

$$\mathrm{ord}(\varphi(a)) \mid \mathrm{ord}(a).$$

Je-li navíc φ *prostý, pak*

$$\mathrm{ord}(\varphi(a)) = \mathrm{ord}(a).$$

Důkaz. Označme $\mathrm{ord}(a) = n$. Pak $\varphi(a)^n = \varphi(a^n) = \varphi(1) = e$, čili nutně $\varphi(a)^k = e$ pro nějaké $k \mid n$. Je-li navíc φ prostý, pro všechna $k < n$ musí platit $\varphi(a)^k = \varphi(a^k) \neq e$, protože $a^k \neq 1$ □

Úloha. Popište všechny homomorfismy $\mathbb{Z}_{10} \to \mathbf{S}_3$.

Řešení. Grupa $\mathbb{Z}_{10}$ je cyklická, $\mathbb{Z}_{10} = \langle 1 \rangle$, čili stačí určit přípustné hodnoty $\varphi(1)$: potom $\varphi(k) = \varphi(1 + \ldots + 1) = \varphi(1) \circ \ldots \circ (1) = \varphi(1)^k$. Řád prvku 1 v $\mathbb{Z}_{10}$ je 10, čili řád prvku $\varphi(1)$ v $\mathbf{S}_3$ musí číslo 10 dělit. Avšak v $\mathbf{S}_3$ jsou pouze prvky řádu 1, 2, 3, čili máme nejvýše čtyři možnosti: $\varphi(1) \in \{id, (1\ 2), (1\ 3), (2\ 3)\}$. Je snadné ověřit, že všechna čtyři zobrazení $k \mapsto id$ a $k \mapsto (i\ j)^k$ jsou homomorfismy $\mathbb{Z}_{10} \to \mathbf{S}_3$. (Rozmyslete si, proč zobrazení $k \mapsto (1\ 2\ 3)^k$ nesplňuje definici homomorfismu, bez odkazu na Tvrzení 15.4.) □

Tvrzení 15.5. *Buď* $\mathbf{G}, \mathbf{H}, \mathbf{K}$ *grupy a* $\varphi : \mathbf{G} \to \mathbf{H}$, $\psi : \mathbf{H} \to \mathbf{K}$ *homomorfismy. Pak*

1. $\psi \circ \varphi$ *je homomorfismus* $\mathbf{G} \to \mathbf{K}$,
2. *je-li* φ *bijektivní, pak* φ^{-1} *je homomorfismus* $\mathbf{H} \to \mathbf{G}$.

Důkaz. (1) Označme $\mathbf{K} = (K, +, -, 0)$. Pro $a, b \in G$ platí
$$(\psi \circ \varphi)(a \cdot b) = \psi(\varphi(a \cdot b)) = \psi(\varphi(a) * \varphi(b)) = \psi(\varphi(a)) + \psi(\varphi(b)) = (\psi \circ \varphi)(a) + (\psi \circ \varphi)(b)$$
postupným použitím faktu, že φ a ψ jsou homomorfismy.

(2) Napišme $u, v \in H$ jako $u = \varphi(a)$ a $v = \varphi(b)$ pro jistá $a, b \in G$. Pak
$$\varphi^{-1}(u * v) = \varphi^{-1}(\varphi(a) * \varphi(b)) = \varphi^{-1}(\varphi(a \cdot b)) = a \cdot b = \varphi^{-1}(u) \cdot \varphi^{-1}(v)$$
použitím faktu, že φ je homomorfismus a $\varphi^{-1} \circ \varphi = id$. □

15.2. **Izomorfismus.**

Definice. Bijektivní homomorfismy nazýváme *izomorfismy.*

Z Tvrzení 15.5 ihned plyne, že složení izomorfismů je izomorfismus a inverzní zobrazení k izomorfismu je také izomorfismus.

Na izomorfismus je možné pohlížet jako na „kopírování algebraické struktury“: máme-li grupu $\mathbf{G}$ a bijektivní zobrazení $\varphi : G \to H$, můžeme na množinu H „překopírovat“ grupové operace předpisem
$$e = \varphi(1), \qquad a' = \varphi((\varphi^{-1}(a))^{-1}), \qquad a * b = \varphi(\varphi^{-1}(a) \cdot \varphi^{-1}(b)).$$
Vidíme, že zobrazení φ^{-1} bude izomorfismem mezi novou grupou $\mathbf{H} = (H, *, ', e)$ a starou grupou $\mathbf{G}$. Jedna grupa je kopií druhé, došlo pouze k „přejmenování prvků“ kopírovacím zobrazením φ. Na každý izomorfismus lze pohlížet tímto způsobem.

Dvě grupy $\mathbf{G}, \mathbf{H}$ nazveme *izomorfní*, pokud existuje izomorfismus $\mathbf{G} \to \mathbf{H}$, tento fakt značíme $\mathbf{G} \simeq \mathbf{H}$. Neformálně, jedna grupa je „kopií“ druhé. Tvrzení 15.5 implikuje, že izomorfismus dává ekvivalenci na třídě všech grup:

- reflexivita: $\mathbf{G} \simeq \mathbf{G}$ je zaručeno izomorfismem $id : \mathbf{G} \to \mathbf{G}$;
- symetrie: je-li $\mathbf{G} \simeq \mathbf{H}$ pomocí izomorfismu φ, pak $\mathbf{H} \simeq \mathbf{G}$ pomocí izomorfismu φ^{-1};
- tranzitivita: je-li $\mathbf{G} \simeq \mathbf{H}$ pomocí izomorfismu φ a $\mathbf{H} \simeq \mathbf{K}$ pomocí izomorfismu ψ, pak $\mathbf{G} \simeq \mathbf{K}$ pomocí izomorfismu $\psi \circ \varphi$.

Na prosté homomorfismy lze nahlížet jako na izomorfismy mezi výchozí grupou a obrazem, tj. prostý homomorfismus $\varphi : \mathbf{G} \to \mathbf{H}$ je izomorfismem $\mathbf{G} \simeq \mathbf{Im}(\varphi)$. Takovým homomorfismům se říká *vnoření* grupy $\mathbf{G}$ do grupy $\mathbf{H}$, tj. grupa $\mathbf{H}$ obsahuje izomorfní kopii $\mathbf{G}$ jako podgrupu.

Příklad. Grupy $\mathbb{Z}_2$ a $\mathbb{Z}^*$ jsou izomorfní. Podívejme se na tabulky jejich operací:

$+$	0	1
0	0	1
1	1	0

$\cdot$	1	-1
1	1	-1
-1	-1	1

Tyto tabulky vypadají podobně: jedna je kopií druhé, pokud přepíšeme $0 \mapsto 1$, $1 \mapsto -1$. Toto zobrazení, které můžeme také zapsat $a \mapsto (-1)^a$, je grupový izomorfismus.

Příklad. Grupy $\mathbb{C}$ a $\mathbb{R} \times \mathbb{R}$ jsou izomorfní. Intuitivně, komplexní čísla odpovídají dvojicím reálných čísel, v obou interpretacích se sčítají jednotlivé složky. Není těžké ověřit, že zobrazení $a + bi \mapsto (a, b)$ je grupový izomorfismus $\mathbb{C} \simeq \mathbb{R} \times \mathbb{R}$.

Příklad. Grupy $\mathbb{Z}_n$ a $\mathbb{C}_n = \langle \zeta_n \rangle_{\mathbb{C}^*}$, kde $\zeta_n = e^{2\pi i/n}$, jsou izomorfní. Intuitivně, komplexní čísla tvaru ζ_n^k se násobí tak, že se exponenty sčítají modulo n. Není těžké ověřit, že zobrazení $k \mapsto \zeta_n^k$ je grupový izomorfismus $\mathbb{Z}_n \simeq \mathbb{C}_n$.

Příklad. Všechny tři grupy $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbb{Z}_8^*$ a $\{id, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\} \leq \mathbf{S}_4$ jsou navzájem izomorfní. Není to vidět na první pohled, ale intuice se dá vybudovat přes generátory: všechny tři grupy lze napsat jako $\mathbf{G} = \langle a, b \rangle$, kde $a^2 = 1$, $b^2 = 1$ a $ab = ba$ je ten třetí prvek různý od jednotky. Formální důkaz si udělejte jako cvičení.

Důležitým příkladem izomorfismu je modulární zobrazení z důkazu čínské věty o zbytcích (Věta 4.13). Dokonce, je to nejen grupový, ale také okruhový izomorfismus, tj. zachovává obě základní operace $+, \cdot$.

Tvrzení 15.6 (algebraická verze čínské věty o zbytcích). *Buďte $m_1, \ldots, m_n$ po dvou nesoudělná přirozená čísla a označme $M = m_1 \cdot \ldots \cdot m_n$. Zobrazení*

$$\begin{aligned} \varphi : \mathbb{Z}_M &\to \mathbb{Z}_{m_1} \times \ldots \times \mathbb{Z}_{m_n} \\ a &\mapsto (a \bmod m_1, \ldots, a \bmod m_n). \end{aligned}$$

je izomorfismem těchto okruhů. Restrikce $\varphi|_{\mathbb{Z}_M^*}$ je grupovým izomorfismem

$$\mathbb{Z}_M^* \simeq \mathbb{Z}_{m_1}^* \times \ldots \times \mathbb{Z}_{m_n}^*.$$

Důkaz. Pohledem do důkazu čínské věty o zbytcích zjistíme, že je zobrazení φ bijektivní. Ověříme, že to je homomorfismus: pro obě operace $* \in \{+, \cdot\}$ platí

$$\begin{aligned} \varphi(a) * \varphi(b) &= (a \bmod m_1, \ldots, a \bmod m_n) * (b \bmod m_1, \ldots, b \bmod m_n) \\ &= ((a * b) \bmod m_1, \ldots, (a * b) \bmod m_n) = \varphi(a * b \bmod M), \end{aligned}$$

přičemž v poslední rovnosti využíváme faktu, že všechna m_i dělí M.

Pohledem do druhé části důkazu Tvrzení 4.9 zjistíme, že invertibilní prvky modulo M jsou zobrazeny na invertibilní prvky modulo jednotlivá m_i, čili $\varphi|_{\mathbb{Z}_M^*}$ je skutečně bijekcí na množinu $\mathbb{Z}_{m_1}^* \times \ldots \times \mathbb{Z}_{m_n}^*$. □

15.3. **Neizomorfismus.**

Viděli jsme, že zobrazení $a + bi \mapsto (a, b)$ je izomorfismem grup $\mathbb{C} \simeq \mathbb{R} \times \mathbb{R}$, ale není izomorfismem grup $\mathbb{C}^*$ a $\mathbb{R}^* \times \mathbb{R}^*$. Nemohly by tyto grupy být izomorfní použitím nějakého jiného izomorfismu?

Podobně, čínská věta o zbytcích tvrdí, že pro m, n nesoudělná je $\mathbb{Z}_{mn} \simeq \mathbb{Z}_m \times \mathbb{Z}_n$. Jak je tomu pro m, n soudělná? Zobrazení $x \mapsto (x \bmod m, x \bmod n)$ není ani prosté, ani na: čísla 0 i $\mathrm{NSN}(m, n)$ se zobrazí na dvojici $(0, 0)$. Nemohly by ale tyto grupy být izomorfní použitím nějakého jiného izomorfismu?

Obecným principem, který umožňuje řešit takové úlohy, jsou *invarianty.* Vlastnost V nazveme invariantem, pokud pro každou dvojici izomorfních grup $\mathbf{G} \simeq \mathbf{H}$ platí, že pokud má grupa $\mathbf{G}$ vlastnost V, pak má i grupa $\mathbf{H}$ vlastnost V.

Příkladem invariantu je počet prvků daného řádu: je-li φ izomorfismus, podle Tvrzení 15.4 je řád a a $\varphi(a)$ vždy stejný.

Příklad.

- Grupa $\mathbb{Z}_{mn}$ obsahuje prvek řádu mn. Avšak v grupě $\mathbb{Z}_m \times \mathbb{Z}_n$ mají všechny prvky řád nejvýše $\mathrm{NSN}(m, n)$. Čili pokud jsou m, n soudělné, tyto grupy nemohou být izomorfní.
- Grupa $\mathbb{C}^*$ obsahuje prvky libovolného řádu, avšak grupa $\mathbb{R}^* \times \mathbb{R}^*$ obsahuje pouze prvky řádu $1, 2, \infty$. Čili tyto grupy nemohou být izomorfní.
- Kvaternionová grupa $\mathbf{Q}_8$ i dihedrální grupa $\mathbf{D}_8$ obsahují prvky řádů $1, 2, 4$. Avšak $\mathbf{Q}_8$ obsahuje šest prvků řádu 4, zatímco $\mathbf{D}_8$ pouze dva, takže nemohou být izomorfní.

Jiným příkladem invariantu je minimální počet generátorů.

Tvrzení 15.7. *Buď $\varphi : \mathbf{G} \to \mathbf{H}$ homomorfismus grup, který je* na. *Je-li $\mathbf{G} = \langle X \rangle$, pak $\mathbf{H} = \langle \varphi(X) \rangle$.*

Důkaz. Prvek $b \in H$ napíšeme jako $b = \varphi(a)$ pro nějaký $a \in G$, prvek a vyjádříme v generátorech jako $a = u_1^{k_1} \cdot \ldots \cdot u_n^{k_n}$, kde $u_i \in X$, a prvek b dostaneme jako $b = \varphi(a) = \varphi(u_1)^{k_1} * \ldots * \varphi(u_n)^{k_n} \in \langle \varphi(X) \rangle$. □

Na rozdíl od vektorových prostorů, v grupách mohou být minimální generující množiny různě velké, např. $\mathbb{Z} = \langle 1 \rangle = \langle 2, 3 \rangle$. Invariantem je *nejmenší* počet prvků, který je potřeba k nagenerování dané grupy.

Příklad. Grupy $\mathbb{Z}$ a $\mathbb{Z} \times \mathbb{Z}$ nejsou izomorfní, protože grupu $\mathbb{Z} \times \mathbb{Z}$ nelze nagenerovat jedním prvkem: podgrupa $\langle (a, b) \rangle = \{(ka, kb) : k \in \mathbb{Z}\}$ obsahuje dvojici $(1, 1)$ pouze pro $(a, b) = \pm(1, 1)$, ale ani jedna z těchto dvojic $\mathbb{Z} \times \mathbb{Z}$ negeneruje. O něco složitější argument by prošel i pro úlohu $\mathbb{Z}_{mn} \not\simeq \mathbb{Z}_m \times \mathbb{Z}_n$ pro soudělná m, n.

Uvedené dva invarianty umožňují prokázat neizomorfismus ve spoustě případů, ale ne ve všech. Příkladem je dvojice grup $\mathbb{Q}$ a $\mathbb{Q}^+ = \{a \in \mathbb{Q} : a > 0\} \leq \mathbb{Q}^*$, které nejsou konečně generované a kromě jednotky obsahují pouze prvky nekonečných řádů.

Příklad. Existence odmocnin, tj. vlastnost „pro každé a existuje b takové, že $a = b^2$", je invariantem. Mějme izomorfismus $\varphi : \mathbf{G} \to \mathbf{H}$ a předpokládejme, že tato vlastnost platí v grupě $\mathbf{G}$. Prvek $u \in H$ napíšeme jako $u = \varphi(a)$, vezmeme $b \in G$ takové, že $a = b \cdot b$ a položíme $v = \varphi(b)$. Vidíme, že $u = \varphi(a) = \varphi(b^2) = \varphi(b)^2 = v^2$.

Tento invariant je splněn v grupě $\mathbb{Q}$, kde jde o vlastnost „pro každé $a \in \mathbb{Q}$ existuje $b \in \mathbb{Q}$ takové, že $a = 2b$". Ale není splněn v grupě $\mathbb{Q}^+$, kde jde o vlastnost „pro každé $0 \neq a \in \mathbb{Q}$ existuje $0 \neq b \in \mathbb{Q}$ takové, že $a = b^2$".

Obecně lze říci, že invariantem je každá vlastnost, kterou lze zformulovat pomocí operací dané struktury, rovnosti, logických spojek a kvantifikátorů (tzv. formule prvního řádu). Ani tyto invarianty však nemusí pomoci. Příkladem jsou grupy $\mathbb{Q}$ a $\mathbb{Q} \times \mathbb{Q}$, které nelze odlišit žádnou formulí prvního řádu, ale přesto nejsou izomorfní (viz cvičení).

15.4. **Klasifikační věty.**

Jedním ze základních cílů každé algebraické teorie je tzv. *klasifikace* objektů, tj. *úplný seznam* všech příkladů *až na izomorfismus.* Obvykle není možné provést takovou klasifikaci kompletně, ale často je možné klasifikovat objekty s nějakou speciální, nicméně důležitou vlastností.

Asi nejjednodušším příkladem je *klasifikace cyklických grup.* Grupa se nazývá *cyklická*, pokud má jeden generátor. Každá taková grupa je izomorfní právě jedné z grup $\mathbb{Z}$ nebo $\mathbb{Z}_n$. Jinými slovy, $\mathbb{Z}$ a $\mathbb{Z}_n$ jsou, až na izomorfismus, všechny příklady cyklických grup.

Věta 15.8 (klasifikace cyklických grup)**.** *Buď* $\mathbf{G}$ *cyklická grupa.*

(1) *Je-li* $\mathbf{G}$ *nekonečná, pak je izomorfní grupě* $\mathbb{Z}$.
(2) *Je-li* $\mathbf{G}$ *konečná řádu* n*, pak je izomorfní grupě* $\mathbb{Z}_n$.

Důkaz. Buď $\mathbf{G} = \langle a \rangle$ cyklická grupa.

(1) Předpokládejme, že je $\mathbf{G}$ nekonečná, tedy $\operatorname{ord}(a) = \infty$, a uvažujme zobrazení

$$\mathbb{Z} \to \mathbf{G}, \qquad k \mapsto a^k.$$

Toto zobrazení je homomorfismus, neboť $a^k \cdot a^l = a^{k+l}$. Přitom jádro je triviální, protože $a^k \neq 1$ pro všechna $k \neq 0$, takže podle Tvrzení 15.3 jde o prosté zobrazení. Podle Důsledku 14.3 je toto zobrazení na $\mathbf{G}$.

(2) Předpokládejme, že je $\mathbf{G}$ řádu n, tedy $\operatorname{ord}(a) = n$, a uvažujme zobrazení

$$\mathbb{Z}_n \to \mathbf{G}, \qquad k \mapsto a^k.$$

Toto zobrazení je homomorfismus, neboť $a^k \cdot a^l = a^{k+l} = a^{k+l \bmod n}$, přičemž druhá rovnost plyne z následující úvahy: pokud $k + l < n$, tvrzení je triviální; pokud $k + l \geq n$, pak $k + l \bmod n = k + l - n$, a tedy $a^{k+l \bmod n} = a^{k+l} \cdot a^{-n} = a^{k+l} \cdot 1^{-1} = a^{k+l}$. Podobně jako pro nekonečnou grupu dostáváme, že jádro je triviální a že jde o zobrazení na $\mathbf{G}$. □

Mnohem komplikovanější je *klasifikace konečně generovaných abelovských grup*, která říká, že každá abelovská grupa s konečnou množinou generátorů je izomorfní direktnímu součinu konečně mnoha cyklických grup. Navíc, použitím čínské věty o zbytcích ve formě Tvrzení 15.6, konečné cyklické grupy stačí uvažovat pouze řádu mocniny prvočísla. Tyto komponenty jsou navíc jednoznačně určené (až na pořadí), tj. volbou neizomorfních cyklických grup dostaneme neizomorfní direktní součiny.

Věta 15.9 (klasifikace konečných abelovských grup)**.** *Buď* $\mathbf{G}$ *konečně generovaná abelovská grupa,* $|\mathbf{G}| > 1$*. Pak existují* $m, n \geq 0$*, prvočísla* $p_1, \dots, p_m$ *(ne nutně po dvou různá) a přirozená čísla* $k_1, \dots, k_m$ *taková, že*

$$\mathbf{G} \simeq \mathbb{Z}^n \times \mathbb{Z}_{p_1^{k_1}} \times \mathbb{Z}_{p_2^{k_2}} \times \cdots \times \mathbb{Z}_{p_m^{k_m}}.$$

Čísla m, n *jsou určena jednoznačně a čísla* $p_1^{k_1}, \dots, p_m^{k_m}$ *jednoznačně až na pořadí.*

Důkaz této věty je poměrně zdlouhavý, proto jej přenecháme do některého navazujícího kurzu. Historie této věty je příznačná v kontexu vzniku moderní algebry. Gauss sice neznal pojem grupy, ale na speciální případ této věty přišel při studiu teorie čísel. Důkaz pro konečné abelovské grupy lze vystopovat v práci Leopolda Kroneckera z roku 1870 a obecnější verzi pro konečně generované grupy u Henri Poincarého z roku 1900, ale ani jeden z nich větu neformuloval v abstraktním jazyce teorie grup. Až pozdější matematici si všimli, že jejich důkaz projde zcela obecně.

Příklad. Podle Věty 15.9 je každá čtyřprvková abelovská grupa izomorfní buď grupě $\mathbb{Z}_4$, nebo grupě $\mathbb{Z}_2 \times \mathbb{Z}_2$.

- Grupa $\mathbb{Z}_5^*$ je také čtyřprvková. Vidíme, že $\mathrm{ord}(2) = 4$, takže $\mathbb{Z}_5^* \simeq \mathbb{Z}_4$.
- Grupa $\mathbb{Z}_8^*$ je také čtyřprvková. Vidíme, že $\mathrm{ord}(3) = \mathrm{ord}(5) = \mathrm{ord}(7) = 2$, takže $\mathbb{Z}_8^* \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$.

Oblíbenou kratochvílí grupařů je hledání malých grup, což je svým způsobem také klasifikační věta. V současné době je znám seznam všech grup až do velikosti $2047 = 2^{11} - 1$. Následující tabulka obsahuje klasifikaci všech grup řádu n pro $n \leq 15$ a pro $n = p$, $2p$, p^2, kde p je prvočíslo.

n	grupy řádu n
1	$\mathbb{Z}_1$
2	$\mathbb{Z}_2$
3	$\mathbb{Z}_3$
4	$\mathbb{Z}_4,\ \mathbb{Z}_2 \times \mathbb{Z}_2$
5	$\mathbb{Z}_5$
6	$\mathbb{Z}_6,\ \mathbf{S}_3 = \mathbf{D}_6$
7	$\mathbb{Z}_7$
8	$\mathbb{Z}_8,\ \mathbb{Z}_2 \times \mathbb{Z}_4,\ \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2,\ \mathbf{D}_8,\ \mathbf{Q}_8$
p	$\mathbb{Z}_p$
p^2	$\mathbb{Z}_{p^2},\ \mathbb{Z}_p \times \mathbb{Z}_p$
$2p$	$\mathbb{Z}_{2p},\ \mathbf{D}_{2p}$
12	$\mathbb{Z}_4 \times \mathbb{Z}_3,\ \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3, \mathbf{A}_4, \mathbf{D}_{12}, \mathbf{X}$
15	$\mathbb{Z}_3 \times \mathbb{Z}_5$

Případ $n = p$ je důsledkem Lagrangeovy věty: grupa prvočíselné velikosti nemůže mít vlastní podgrupy, takže musí být generovaná libovolným svým prvkem (kromě jednotky) a podle klasifikace cyklických grup musí být izomorfní $\mathbb{Z}_p$.

Případy $n = p^2$ a $n = 2p$ svojí obtížností mírně přesahují možnosti této učebnice. Pro čtverec prvočísla se použije podobný trik, jako v důkazu Věty 18.6, přičemž se nechá působit grupa na svých prvcích vnitřními automorfismy. K případu $n = 2p$ i ke klasifikaci řádů 8, 12, 15 se hodí pojem semidirektního součinu a pro $n = 12, 15$ také Sylowovy věty. Pomocí semidirektního součinu se také zkonstruuje tajemná grupa $\mathbf{X}$ uvedená v tabulce.

15.5. **Reprezentace grup.**

Slabším typem věty popisující objekty v dané třídě jsou tzv. *reprezentační věty*. Každý objekt je popsán, až na izomorfismus, jako podobjekt objektu nějakého konkrétního typu. Formálně, existuje *vnoření* (prostý homomorfismus) do tohoto objektu. Reprezentační věty ukazují, že zkoumání obecné teorie je stejně těžké, jako zkoumání podobjektů těchto konkrétních objektů.

V úvodu kapitoly o grupách jsme zmínili, že dvěma základními příklady grup jsou grupy permutací a grupy regulárních matic. To proto, že každou grupu lze reprezentovat jako grupu permutací a každou konečnou grupu jako grupu regulárních matic. Tyto dvě reprezentační věty si nyní ukážeme.

Věta 15.10 (Cayleyova reprezentace)**.** *Každou grupu lze vnořit do nějaké symetrické grupy.*

Formálně, pro každou grupu $\mathbf{G}$ existuje množina X a prostý homomorfismus $\varphi : \mathbf{G} \to \mathbf{S}_X$. Čili $\mathbf{G}$ je izomorfní s permutační grupou $\mathbf{Im}(\varphi) \leq \mathbf{S}_X$.

Důkaz. Buď $\mathbf{G}$ grupa a uvažujme pro každé $a \in G$ zobrazení

$$L_a : G \to G, \quad x \mapsto a \cdot x$$

(těmto zobrazením se říká *levé translace*). Tato zobrazení jsou permutace: jediné řešení rovnice $L_a(x) = a \cdot x = y$ je prvek $x = a^{-1} \cdot y$, tedy zobrazení L_a jsou bijektivní. Uvažujme nyní zobrazení

$$\lambda : \mathbf{G} \to \mathbf{S}_G, \quad a \mapsto L_a.$$

Dokážeme, že λ je homomorfismus. Podle Lemmatu 15.1 stačí ověřit, že $\lambda(a \cdot b) = \lambda(a) \circ \lambda(b)$, tj. že zobrazení $L_{a \cdot b}$ je totožné se složením zobrazení $L_a \circ L_b$. Dosazením $u \in G$ vidíme, že

$$L_{a \cdot b}(u) = (a \cdot b) \cdot u = a \cdot (b \cdot u) = L_a(b \cdot u) = L_a(L_b(u)) = (L_a \circ L_b)(u).$$

K ověření prostosti stačí spočítat jádro: kdyby $L_a = id$, pak $a = L_a(1) = id(1) = 1$. Tedy $\mathbf{Ker}(\lambda) = \{1\}$ a λ je prostý homomorfismus. □

Věta 15.11 (lineární reprezentace)**.** *Každou konečnou grupu lze vnořit do nějaké obecné lineární grupy, nad libovolným tělesem.*

Důkaz. Buď $\mathbf{T}$ libovolné těleso, $\mathbf{G}$ daná konečná grupa a označme $n = |G|$. Buď φ bijekce $G \to H = \{1, \ldots, n\}$. Nejprve vytvoříme kopii grupy $\mathbf{G}$ na množině H, tj. vytvoříme grupu $\mathbf{H}$ tak, že φ je izomorfismus $\mathbf{G} \simeq \mathbf{H}$. Dále vezmeme Cayleyovu reprezentaci $\lambda : \mathbf{H} \to \mathbf{S}_n$. Pokud najdeme vnoření ψ grupy $\mathbf{S}_n$ do $\mathbf{GL}_n(\mathbf{T})$, hledaným vnořením $\mathbf{G}$ do $\mathbf{GL}_n(\mathbf{T})$ bude složení $\psi \circ \lambda \circ \varphi$. Uvažujme

$$\psi : \mathbf{S}_n \to \mathbf{GL}_n(\mathbf{T}), \quad \sigma \mapsto \big(\delta_{i,\sigma(j)}\big)_{i,j=1}^{n},$$

kde $\delta_{u,v} = 1$ pokud $u = v$ a $\delta_{u,v} = 0$ v opačném případě. Tedy $\psi(\sigma)$ je matice, ve které je v každém řádku a každém sloupci právě jedna jednička a jinak samé nuly, přičemž ta jednička na i-tém řádku je v $\sigma^{-1}(i)$-tém sloupci. Evidentně jde o prosté zobrazení, zbývá tedy dokázat, že to je homomorfismus, tedy že platí

$$\psi(\pi \circ \sigma) = \psi(\pi) \cdot \psi(\sigma)$$

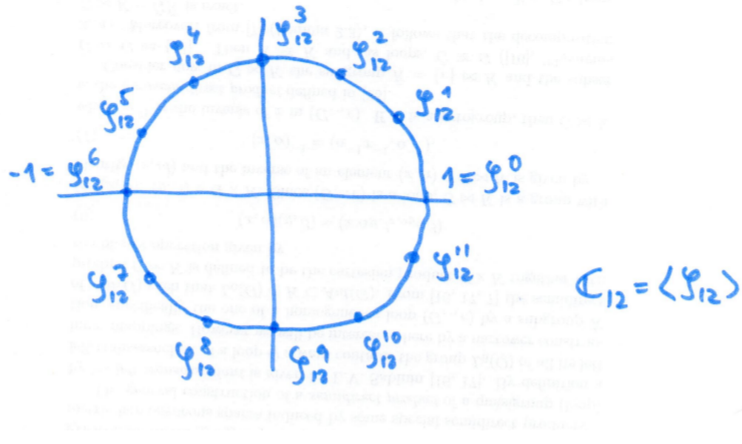
pro všechny permutace $\pi, \sigma \in S_n$. Pravá strana je rovna

$$\big(\delta_{i,\pi(j)}\big)_1^n \cdot \big(\delta_{i,\sigma(j)}\big)_1^n = \Big(\sum_k \delta_{i,\pi(k)} \cdot \delta_{k,\sigma(j)}\Big)_1^n.$$

Přitom $\delta_{i,\pi(k)} \cdot \delta_{k,\sigma(j)} = 1$ právě tehdy, když $i = \pi(k)$ a $k = \sigma(j)$, což je právě tehdy, když $i = \pi(\sigma(j))$ a $k = \sigma(j)$. Tedy celá suma je rovna jedné pro $i = \pi(\sigma(j))$ a nule v opačném případě. Tím pádem je to přesně matice $\psi(\pi \circ \sigma)$. □

Příklad. Rozebereme si reprezentaci grupy $\mathbf{S}_3$ v grupě $\mathbf{GL}_3(\mathbf{T})$. Podle návodu uvedeného v důkazu Věty 15.11 zkonstruujeme

$$\psi(id) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \psi((1\ 2\ 3)) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \psi((1\ 3\ 2)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$
$$\psi((1\ 2)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \psi((2\ 3)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \psi((1\ 3)) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$



Obrázek 18. Ilustrace faktu, proč se cyklické grupy nazývají cyklické.

16. Cyklické grupy

16.1. Podgrupy, generátory, řády prvků.

Grupa $\mathbf{G}$ se nazývá *cyklická*, pokud je generovaná jedním prvkem, tj.

$$\mathbf{G} = \langle a \rangle_{\mathbf{G}}$$

pro nějaké $a \in G$. Její prvky lze díky Důsledku 14.3 vyjádřit jako mocniny generátoru,

$$G = \{a^k : \ k \in \mathbb{Z}\},$$

z čehož je vidět, že je to nutně grupa abelovská. Z Tvrzení 14.6 plyne, že je-li řád a nekonečný, pak jsou tyto mocniny po dvou různé, a je-li $\operatorname{ord}(a) = n$ konečný, pak $G = \{a^0, a^1, \dots, a^{n-1}\}$. Odsud pochází název pro cyklické grupy: při násobení daným prvkem a cyklicky procházíme přes všechny prvky grupy $\mathbf{G}$.

Klasifikace cyklických grup (Věta 15.8) říká, že každá cyklická grupa je izomorfní jedné z grup $\mathbb{Z}$, $\mathbb{Z}_n$. Přirozených příkladů je však více.

Příklady (cyklické grupy)**.**

- Grupy $\mathbb{Z}$ a $\mathbb{Z}_n$, $n \in \mathbb{N}$, jsou cyklické, generované prvkem 1.
- Grupy $\mathbb{C}_n \leq \mathbb{C}^*$ sestávající ze všech komplexních kořenů polynomu $x^n - 1$ jsou cyklické, $\mathbb{C}_n = \langle e^{2\pi i/n} \rangle$.
- V této sekci si dokážeme, že grupy $\mathbb{Z}_p^*$ jsou cyklické pro každé prvočíslo p (Věta 16.7). Například $\mathbb{Z}_5^* = \langle 2 \rangle$, $\mathbb{Z}_7^* = \langle 3 \rangle$, $\mathbb{Z}_{11}^* = \langle 2 \rangle$.
- Některé grupy $\mathbb{Z}_n^*$, n složené, jsou cyklické, např. $\mathbb{Z}_6^* = \{1, 5\} = \langle 5 \rangle$, ale některé ne, např. grupa $\mathbb{Z}_8^*$ cyklická není.

Příklad. Každá grupa $\mathbf{G}$ prvočíselného řádu je cyklická. Uvažujme podgrupu $\langle a \rangle$, $a \neq 1$. Podle Lagrangeovy věty je $|\langle a \rangle|$ dělí $|\mathbf{G}|$, přitom $|\langle a \rangle| > 1$, tedy $|\langle a \rangle| = |\mathbf{G}|$ a prvek a tuto grupu generuje.

Nejprve se podíváme, jak vypadají podgrupy cyklických grup.

Tvrzení 16.1. *Každá podgrupa cyklické grupy je cyklická.*

Důkaz. Buď $\mathbf{H}$ podgrupa cyklické grupy $\mathbf{G} = \langle a \rangle$. Je-li $H = \{1\}$, pak $\mathbf{H} = \langle 1 \rangle$. V opačném případě označme k nejmenší kladné číslo takové, že $a^k \in H$ (takové jistě existuje: je-li $1 \neq b \in H$, pak $b = a^l$ pro nějaké l a buď b nebo b^{-1} má exponent kladný). Dokážeme, že $\mathbf{H} = \langle a^k \rangle$. Inkluze $\langle a^k \rangle \subseteq H$ je zřejmá. Pro spor tedy předpokládejme, že existuje nějaký prvek $a^n \in H \setminus \langle a^k \rangle$. Nutně $k \nmid n$, jinak bychom měli $a^n = (a^k)^{n/k} \in \langle a^k \rangle$. Napišme $n = kq + r$, kde $0 < r < k$. Pak

$$a^r = a^{n-kq} = a^n \cdot (a^k)^{-q} \in H,$$

protože a^n i a^k leží v H, což je spor s volbou k jako nejmenšího kladného čísla s vlastností $a^k \in H$. $\square$

Tvrzení nelze obrátit: Prüferova grupa $\mathbb{C}_{p^\infty}$ je příkladem grupy, která není cyklická (ani konečně generovaná), přestože všechny její vlastní podgrupy cyklické jsou.

Příklad (podgrupy grupy $\mathbb{Z}$)**.** Grupa $\mathbb{Z}$ je cyklická, čili její podgrupy jsou cyklické, tedy tvaru

$$\langle k\rangle = k\mathbb{Z} = \{a \in \mathbb{Z} : \ k \mid a\}$$

pro nějaké $k \in \mathbb{Z}$. Přitom $k\mathbb{Z} = l\mathbb{Z}$ právě tehdy, když $k = \pm l$. Podgrupy jsou tedy ve vzájemně jednoznačné korepondenci s nezápornými čísly a $k\mathbb{Z} \subseteq l\mathbb{Z}$ právě tehdy, když $l \mid k$. Čili podgrupy jsou uspořádány vzhledem k inkluzi opačně než množina $\mathbb{N} \cup \{0\}$ dělitelností.

Pro konečné cyklické grupy je situace složitější, mnoho různých prvků může generovat stejné podgrupy.

Lemma 16.2 (podgrupy cyklických grup)**.** *Buď* $\mathbf{G} = \langle a\rangle$ *cyklická grupa. Pak*

(1) $\langle a^k, a^l\rangle = \langle a^{\mathrm{NSD}(k,l)}\rangle$,
(2) *je-li* $|\mathbf{G}| = n$, *pak* $\langle a^k\rangle = \langle a^{\mathrm{NSD}(k,n)}\rangle$.

Důkaz. (1) Protože $\mathrm{NSD}(k,l)$ dělí k i l, platí $a^k, a^l \in \langle a^{\mathrm{NSD}(k,l)}\rangle$, čímž máme prokázánu inkluzi $\subseteq$. Naopak, podle Bézoutovy rovnosti je $\mathrm{NSD}(k,l) = uk + vl$ pro nějaká $u, v \in \mathbb{Z}$, a tedy

$$a^{\mathrm{NSD}(k,l)} = a^{uk+vl} = (a^k)^u \cdot (a^l)^v \in \langle a^k, a^l\rangle,$$

čímž máme prokázánu inkluzi $\supseteq$.

(2) Dosadíme $l = n$: pak $\langle a^{\mathrm{NSD}(k,n)}\rangle = \langle a^k, a^n\rangle = \langle a^k\rangle$, protože $a^n = 1$. $\square$

Tvrzení 16.3 (generátory cyklických grup)**.** *Buď* $\mathbf{G} = \langle a\rangle$ *cyklická grupa.*

(1) *Pokud je* $\mathbf{G}$ *nekonečná, generátorem jsou pouze prvky* a, a^{-1}.
(2) *Pokud je* $\mathbf{G}$ *konečná řádu* n*, generátorem jsou právě prvky* a^k*, kde* $k \in \{1, \dots, n-1\}$ *je nesoudělné s* n.

Důkaz. (1) Oba prvky a, a^{-1} grupu $\mathbf{G}$ generují, protože $\{a^k : \ k \in \mathbb{Z}\} = \{a^{-k} : \ k \in \mathbb{Z}\}$. Žádný jiný generátor grupa $\mathbf{G}$ nemá: kdyby $\mathbf{G} = \langle a^n\rangle$ pro nějaké n, pak by existovalo $m \in \mathbb{Z}$ takové, že $a = (a^n)^m$, a dostali bychom $1 = (a^n)^m \cdot a^{-1} = a^{mn-1}$; řád a je ovšem nekonečný, a tedy $mn = 1$, čili $n = \pm 1$.

(2) Podle Lemmatu 16.2 je $\langle a^k\rangle = \langle a^{\mathrm{NSD}(k,n)}\rangle$. Pokud $\mathrm{NSD}(k,n) = 1$, pak $\langle a^k\rangle = \langle a\rangle = \mathbf{G}$. Pokud $\mathrm{NSD}(k,n) = d \neq 1$, pak $\langle a^k\rangle = \langle a^d\rangle = \{a^d, a^{2d}, \dots, a^{\frac{n}{d}d}\}$ je vlastní podgrupa. $\square$

Příklad (podgrupy grupy $\mathbb{Z}_n$)**.** Grupa $\mathbb{Z}_n$ je cyklická, čili její podgrupy jsou cyklické, tedy tvaru

$$\langle k\rangle = k\mathbb{Z}_n = \{ku \bmod n : \ u = 0, \dots, n-1\},$$

pro nějaké $k \in \{0, \dots, n-1\}$. Z Lemmatu 16.2(2) s volbou $a = 1$ plyne, že $k\mathbb{Z}_n = \mathrm{NSD}(k,n)\mathbb{Z}_n$, tedy $k\mathbb{Z}_n = l\mathbb{Z}_n$ právě tehdy, když $\mathrm{NSD}(k,n) = \mathrm{NSD}(l,n)$. Podgrupy jsou tedy ve vzájemně jednoznačné korepondenci s děliteli čísla n. Pro $k, l \mid n$ pak platí, že $k\mathbb{Z}_n \subseteq l\mathbb{Z}_n$ právě tehdy, když $l \mid k$. Čili podgrupy jsou uspořádány vzhledem k inkluzi opačně než množina všech dělitelů čísla n dělitelností. Podle Tvrzení 16.3 je $\mathbb{Z}_n = \langle k\rangle$ právě tehdy, když jsou k, n nesoudělná.

Příklad (podgrupy grupy $\mathbb{Z}_p^*$)**.** Grupa $\mathbb{Z}_{11}^* = \langle 2\rangle$ je cyklická řádu 10, čili její podgrupy jsou cyklické, tedy tvaru

$$\langle 2^k\rangle = \{2^{uk} \bmod 11 : \ u = 0, \dots, 10\},$$

pro nějaké $k \in \{0, \dots, 9\}$. Z Lemmatu 16.2(2) plyne, že $\langle 2^k\rangle = \langle 2^l\rangle$ právě tehdy, když $\mathrm{NSD}(k,n) = \mathrm{NSD}(l,n)$. Podgrupy jsou tedy ve vzájemně jednoznačné korepondenci s děliteli čísla 10, máme tedy čtyři podgrupy,

$$\langle 2^1\rangle = \mathbb{Z}_{11}^*, \ \langle 2^2\rangle = \{1, 4, 5, 9, 3\}, \ \langle 2^5\rangle = \{1, 10\}, \ \langle 2^{10}\rangle = \{1\}.$$

Generátory jsou prvky 2^k takové, že k je nesoudělné s 10, tedy prvky $2^1 = 2$, $2^3 = 8$, $2^6 = 7$ a $2^9 = 6$. Všimněte si, že to jsou právě čísla, která nepatří do žádné z vlastních podgrup vypsaných výše.

Úloha se přímočaře zobecní na libovolnou grupu $\mathbb{Z}_p^*$, p prvočíslo, o které si ukážeme, je vždy cyklická, byť není zřejmé, který prvek a je generátorem. Podgrupy pak budou právě $\langle a^k \rangle$, kde $k \mid p-1$, generátory budou právě prvky a^k, kde $\mathrm{NSD}(k, p-1) = 1$.

Z Tvrzení 16.3 plyne, že cyklická grupa řádu n má právě $\varphi(n)$ generátorů, kde φ značí Eulerovu funkci. Tohoto faktu využijeme k řešení obecnější úlohy: spočítáme počet prvků každého řádu. V nekonečných cyklických grupách mají všechny prvky kromě jednotky řád nekonečný. V konečných grupách dává Lagrangeova věta omezení na přípustné řády. Ukážeme si, že v cyklických grupách prvky všech přípustných řádů existují a jejich počet je dán Eulerovou funkcí.

Tvrzení 16.4 (řády prvků cyklických grup)**.** *Cyklická grupa konečného řádu n obsahuje právě $\varphi(d)$ prvků řádu d pro každé $d \mid n$.*

Důkaz. Buď $\mathbf{G}$ cyklická grupa konečného řádu n. Každý prvek řádu $d \mid n$ je generátorem nějaké cyklické podgrupy řádu d. Taková podgrupa však v $\mathbf{G}$ existuje pouze jedna: podle Lemmatu 16.2 jsou všechny podgrupy v $\mathbf{G}$ tvaru $\langle a^k \rangle$, $k \mid n$. Přitom $|\langle a^k \rangle| = \frac{n}{k}$, čili $\langle a^{\frac{n}{d}} \rangle$ je jediná podgrupa řádu d. Ta má podle Tvrzení 16.3 právě $\varphi(d)$ generátorů. □

Tvrzení o počtu prvků daného řádu lze použít k důkazu následující kombinatorické identity.

Tvrzení 16.5. *Pro každé $n \in \mathbb{N}$ platí $\sum_{d \mid n} \varphi(d) = n$.*

Důkaz. Budeme počítat počet prvků grupy $\mathbb{Z}_n$ dvěma způsoby. Jeden způsob je triviální: grupa obsahuje čísla $0, \dots, n-1$, tedy $|\mathbb{Z}_n| = n$. Podruhé spočítáme prvky podle řádů: podle Lagrangeovy věty jsou přípustné řády $d \mid n$, tedy $|\mathbb{Z}_n| = \sum_{d \mid n} u_d$, kde u_d značí počet prvků řádu d. Tvrzení 16.4 říká, že $u_d = \varphi(d)$. □

16.2. **Multiplikativní grupy konečných těles jsou cyklické.**

Tvrzení uvedené v názvu podsekce má dalekosáhlé důsledky v teorii konečných těles. K jeho důkazu použijeme následující kritérium cykličnosti.

Lemma 16.6. *Buď $\mathbf{G}$ konečná grupa a předpokládejme, že pro každé k grupa $\mathbf{G}$ obsahuje nejvýše k prvků a splňujících $a^k = 1$. Pak je grupa $\mathbf{G}$ cyklická.*

Důkaz. Označme $n = |\mathbf{G}|$ a u_k počet prvků řádu k v grupě $\mathbf{G}$. Podle Lagrangeovy věty je $u_k = 0$ pro všechna $k \nmid n$, a tedy $n = \sum_{d \mid n} u_d$ (počítáme prvky $\mathbf{G}$ podle jejich řádu jako v Tvrzení 16.5).

Uvažujme nějaký prvek a řádu k v $\mathbf{G}$. Podgrupa $\langle a \rangle$ je cyklická řádu k a všechny prvky $b \in \langle a \rangle$ splňují $b^k = 1$. Podle předpokladu v $\mathbf{G}$ žádné jiné prvky s touto vlastností nejsou, takže $\langle a \rangle$ je jediná cyklická podgrupa řádu k v $\mathbf{G}$. Podle Tvrzení 16.3 má $\varphi(k)$ generátorů, a tedy $u_k = \varphi(k)$.

Čili pro každé $d \mid n$ platí $u_d = 0$ nebo $u_d = \varphi(d)$. Dokážeme, že vždy nastane druhá možnost: podle Tvrzení 16.5 je $\sum_{d \mid n} \varphi(d) = n = \sum_{d \mid n} u_d$, takže $u_d = \varphi(d)$ pro všechna $d \mid n$. Speciálně $u_n \neq 0$, a tedy v $\mathbf{G}$ existuje prvek řádu n, neboli generátor. □

Věta 16.7. *Buď $\mathbf{T}$ těleso a $\mathbf{G}$ konečná podgrupa grupy $\mathbf{T}^*$. Pak $\mathbf{G}$ je cyklická.*

Důkaz. Podle Věty 2.4 má polynom $x^k - 1$ nejvýše k kořenů v tělese $\mathbf{T}$. Tedy grupa $\mathbf{G} \leq \mathbf{T}^*$ může obsahovat nejvýše k prvků a splňujících $a^k = 1$ a můžeme aplikovat předchozí kritérium. □

Speciálně, multiplikativní grupy konečných těles jsou cyklické. Jejich generátorům se říká *primitivní prvky.* Každé konečné těleso $\mathbf{T} \geq \mathbb{Z}_p$ lze napsat jako $\mathbf{T} = \mathbb{Z}_p(a)$, kde a je libovolný primitivní prvek. Avšak pozor, v tělese $\mathbb{Z}_p[\alpha]/(m)$ není prvek α nutně primitivní: například v $\mathbb{F}_9 = \mathbb{Z}_3[\alpha]/(\alpha^2 + 1)$ generuje α pouze čtyřprvkovou podgrupu.

Není bez zajímavosti, že pro grupy $\mathbb{Z}_p^*$ lze znění Věty 16.7 interpretovat čistě v jazyce elementární teorie čísel: pro každé prvočíslo p existuje číslo a (generátor té grupy) takové, že každé $b \in \{1, \ldots, p-1\}$ lze vyjádřit právě jedním způsobem jako $b = a^k \bmod p$ pro nějaké $k \in \{0, \ldots, p-2\}$.

16.3. **Diskrétní logaritmus a kryptografie.**

Podívejme se znovu na klasifikaci cyklických grup (Věta 15.8). Cyklická grupa $\mathbf{G} = \langle a \rangle$ řádu n je izomorfní grupě $\mathbb{Z}_n$, izomorfismem je zobrazení

$$exp : \mathbb{Z}_n \to \mathbf{G}, \quad k \mapsto a^k.$$

Zobrazení se nazývá *diskrétní exponenciála* a inverznímu zobrazení se říká *diskrétní logaritmus.* Jinými slovy, diskrétní logaritmus prvku $b \in G$ přiřadí to jediné $k \in \{0, \ldots, n-1\}$, pro které $b = a^k$; budeme jej značit $k = \log_a b$. (Pro nekonečné grupy se používá analogická terminologie, ale z výpočetního hlediska nejsou tak zajímavé.)

Počítat diskrétní exponenciálu je zpravidla snadné. Přesněji řečeno, kdykoliv lze v dané grupě efektivně násobit, pak lze také efektivně mocnit. Určitě ne tak, že bychom počítali a^k jako k součinů. Stačí jich méně než $2\lfloor \log_2 k \rfloor$: napíšeme si k ve dvojkové soutavě, $k = \sum_{i=0}^{\lfloor \log_2 k \rfloor} u_i 2^i$, kde $u_i \in \{0, 1\}$, a vyjádříme mocninu jako

$$a^k = a^{\sum_{i=0}^{\lfloor \log_2 k \rfloor} u_i 2^i} = \prod_{i:\, u_i = 1} a^{2^i}.$$

Přitom prvků tvaru a^{2^i} se ve vzorci vyskytuje nejvýše $\lfloor \log_2 k \rfloor + 1$ a spočteme je postupným mocněním

$$a,\ a^2,\ (a^2)^2,\ \ldots,\ a^{2^i} = (a^{2^{i-1}})^2,\ \ldots,$$

čili pomocí $\lfloor \log_2 k \rfloor$ součinů. Vidíme, že spočítat libovolnou mocninu v n-prvkové grupě vyžaduje méně než $2 \log_2 n$ násobení.

Empirická zkušenost ukazuje, že pro spoustu grup je výpočet diskrétního logaritmu obtížný. Hledání logaritmu hrubou silou, procházením všech n možných exponentů, je exponenciálně pomalejší než výpočet mocniny. Pro některé grupy je výpočet snadný (viz příklad níže), ale například pro grupy $\mathbb{Z}_p^*$, p prvočíslo, nebo grupy odvozené z eliptických křivek nad konečnými tělesy, není v současnosti znám výrazně lepší postup než hrubá síla.

Příklad. Uvažujme cyklickou grupu $\mathbb{Z}_n = \langle a \rangle$. Logaritmus $\log_a b$ je roven tomu (jedinému) $k \in \{0, \ldots, n-1\}$, pro které

$$ka \equiv b \pmod{n}.$$

Takové k najdeme snadno Eukleidovým algoritmem: podle Tvrzení 16.3 je generátor nesoudělný s n, spočteme Bézoutovy koeficienty $1 = \mathrm{NSD}(a, n) = ua + vn$ a vidíme, že $b = uab + vnb \equiv uba \pmod{n}$, čili $\log_a b = ub \bmod n$.

Nadále uvažujme libovolnou cyklickou grupu $\mathbf{G}$, pro kterou je diskrétní exponenciála výpočetně zvladatelná, ale logaritmus nikoliv (používají se např. grupy $\mathbb{Z}_p^*$ pro prvočísla $p \geq 2^{1000}$). Ukážeme si dva kryptografické algoritmy založené na diskrétním logaritmu: *Diffie-Hellmanův protokol* pro výměnu klíče (jde o nejpoužívanější algoritmus svého druhu) a *El Gamalův algoritmus* pro kryptografii s veřejným klíčem (v praxi se používá, i když algoritmus RSA ze sekce 4.4, který řeší stejnou úlohu, je výrazně populárnější).

Diffie-Hellmanův protokol. Alice a Bob se potřebují dohodnout na nějakém společném hesle (odborně *klíči*), přičemž k dispozici mají pouze veřejný kanál (např. odposlouchávaný telefon). Jak to provést?

Nejprve se Alice a Bob dohodnou na nějaké cyklické grupě a generátoru, $\mathbf{G} = \langle a \rangle$. Dále si Alice zvolí číslo m a Bob číslo n z intervalu $2, \ldots, |G| - 1$, přičemž každý bude svoje číslo držet v tajnosti. Pak provedou následující úkony: Alice spočte $u = a^m$ a pošle u Bobovi, Bob spočte $v = a^n$ a pošle v Alici. Poté Alice spočte $v^m = (a^n)^m = a^{mn}$ a Bob spočte $u^n = (a^m)^n = a^{mn}$. Oba získali stejný prvek a^{mn} a ten prohlásí za společný klíč.

Kdyby nepřítel poslouchal jejich komunikaci, co zjistí? Bude znát grupu $\mathbf{G}$, generátor a a hodnoty $u = a^m$ a $v = a^n$. Chtěl by spočítat prvek a^{mn}. Této úloze se říká *Diffie-Hellmanův problém.* Očividným řešením je provést diskrétní logaritmus, získat čísla m, n, vynásobit je a dopočítat a^{mn}. Toto řešení však není výpočetně zvladatelné a žádný efektivní postup není v současné době znám.

<u>*El Gamalův protokol*</u>. Příjemce zvolí cyklickou grupu $\mathbf{G} = \langle a \rangle$, náhodné číslo k z intervalu $2, \ldots, |G| - 1$ a spočte $b = a^k$. *Veřejným klíčem* bude $\mathbf{G}, a, b$, jeho *soukromým klíčem* bude k.

Odesílatel zprávy zvolí náhodné číslo l z intervalu $2, \ldots, |G| - 1$, které po odeslání zprávy zničí, a zprávu $x \in G$ zašifruje jako dvojici

$$y = (a^l, x \cdot b^l).$$

Příjemce obdrží dvojici $y = (u, v)$ a dešifruje ji pomocí k takto:

$$v \cdot u^{-k} = x \cdot b^l \cdot (a^l)^{-k} = x \cdot (a^k)^l \cdot (a^l)^{-k} = x.$$

Kdybychom uměli rychle počítat diskrétní logaritmus, okamžitě získáme soukromý klíč. Jsou známy i jiné způsoby útoku na El Gamalův algoritmus, díky nimž například grupy $\mathbb{Z}_p^*$ nejsou považovány za bezpečné. Obecný postup však znám není a El Gamal je považován za bezpečný například na dostatečně velkých grupách odvozených z eliptických křivek.

Jedním ze základních konceptů v kryptografii je pojem *jednosměrné funkce*. Velmi zjednodušeně řečeno, je to bijektivní zobrazení f takové, že hodnoty $f(x)$ se dají počítat rychle, ale není znám postup, kterým by bylo možné získat nějakou statisticky významnou informaci o hodnotách inverzního zobrazení $f^{-1}(y)$. Příkladem je

- diskrétní exponenciála v některých grupách, např. zobrazení $\mathbb{Z}_{p-1} \to \mathbb{Z}_p^*$, $k \mapsto a^k \bmod p$ pro dostatečně velká prvočísla p;
- zobrazení $\mathbb{Z}_N \to \mathbb{Z}_N$, $a \mapsto a^k \bmod N$ pro vhodná k, N, např. je-li N součinem dvou dostatečně velkých velkých prvočísel (na tomto příkladu je založena šifra RSA, viz sekce 4.4).

K čemu může být dobrá funkce, kterou lze zprávu důkladně zašifrovat, ale za žádných okolností dešifrovat? Alice a Bob si chtějí na dálku zahrát hru „*panna nebo orel*“. Alice bude házet mincí, Bob hádat. Jak to ale udělat, aby Alice Boba nepodvedla, když se Bob nemůže na minci podívat? Zvolme nějakou jednosměrnou funkci f na množině $\{1, \ldots, n\}$. Pokud Alice hodí orla, zvolí náhodné liché číslo x, v opačném případě zvolí sudé číslo. Bobovi pošle hodnotu $f(x)$. Protože je f jednosměrná, Bob neumí spočítat, co padlo, zvolí tedy odpověď náhodně a sdělí ji Alici. Nyní Alice zveřejní číslo x a Bob ihned vidí, zda vyhrál. Pro kontrolu si spočte $f(x)$ a porovná ho s hodnotou, kterou dostal na začátku. Pokud se hodnoty neshodují, Alice podvádí. Může Alice Boba podvést tak, aby na to nepřišel? Dejme tomu, že padl orel a to samé si tipnul Bob. Aby Alice Boba podvedla, musela by Bobovi ukázat sudé x' takové, že $f(x') = f(x)$. Jenže takové x' neexistuje, když je f bijekce.

Pro kryptografii jsou zvlášť cenné jednosměrné funkce, ke kterým existují tzv. *zadní vrátka*, dodatečná informace, pomocí které lze inverzní zobrazení počítat efektivně. Příkladem je zobrazení z šifry RSA: počítat odmocniny modulo N je obecně obtížné, ale známe-li prvočíselný rozklad čísla N, je to snadné.

V poslední době je populárním tématem tzv. *homomorfní šifrování*. Cílem je najít jednosměrné funkce (se zadními vrátky), které by byly homomorfismem vůči nějaké zajímavé operaci. Motivace vychází z praxe vzdáleného počítání (*cloud computing*): rádi bychom, aby pro nás někdo spočítal časově náročné úlohy na našich datech, ale zároveň bychom nechtěli tato data prozradit. Přeloženo do matematického jazyka, máme nějakou operaci $*$ na datech (typickým příkladem z praxe je třeba součin velkých matic) a chceme jednosměrnou funkci $X \to X$ takovou, že když pošleme poskytovateli služeb hodnoty $f(x), f(y)$, on spočte výsledek $f(x) * f(y)$ a my jej dešifrujeme zobrazením f^{-1}, dostaneme správný výsledek $x * y$. To jest, chceme, aby platilo $f(x) * f(y) = f(x * y)$, jinými slovy, aby f byl homomorfismus vzhledem k operaci $*$ (která je často grupová, ale jsou i obecnější koncepty). Tento příklad slouží i jako dobrá motivace pojmu homomorfismus.

17. Grupy symetrií

17.1. Symetrie geometrických objektů.

Jednou z původních motivací pro rozvoj teorie grup bylo studium symetrií geometrických objektů. Pro jednoduchost se soustředíme na eukleidovskou geometrii, ačkoliv podobné úvahy lze provádět i v jiných geometriích.

Z kurzu lineární algebry si připomeňme větu, která říká, že izometrie na eukleidovském prostoru $\mathbb{R}^n$ jsou právě ortogonální afinní zobrazení, tj. zobrazení

$$\mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto Ax + b,$$

kde A je ortogonální matice a $b \in \mathbb{R}^n$. Izometrie roviny jsou rotace (otočení) se středem v počátku složené s posunutím a reflexe (osové symetrie) podle osy procházející počátkem složené z posunutím. Izometrie třidimenzionálního prostoru jsou rotace kolem osy procházející počátkem složené z posunutím a dále tato zobrazení složená s reflexí podle roviny procházející počátkem.

Je-li dána podmnožina X eukleidovského prostoru $\mathbb{R}^n$, uvažujme prvky eukleidovské grupy $\mathbf{E}_n$ zachovávající tuto množinu, tj.

$$\mathrm{Sym}(X) = \{\varphi \in E_n : \varphi(X) = X\}.$$

Je snadné ověřit, že $\mathrm{Sym}(X)$ tvoří podgrupu v $\mathbf{E}_n$. Podgrupa $\mathbf{Sym}(X)$ se nazývá *grupa symetrií objektu X*.

Příklad. Uvažujme n-dimenzionální kouli $X = \{(x_1, \dots, x_n) : \sum x_i \leq 1\} \subset \mathbb{R}^n$. Grupa $\mathbf{Sym}(X) = \mathbf{Sym}(\{0\})$ sestává z izometrií, které zachovávají nulu, tj. ze zobrazení $x \mapsto Ax$, kde A je ortogonální matice. Vidíme, že $\mathbf{Sym}(X) \simeq \mathbf{O}_n(\mathbb{R})$.

Příklad. Uvažujme pravidelný n-úhelník $X \subset \mathbb{R}^2$ se středem v nule. Grupa $\mathbf{Sym}(X)$ sestává z rotací se středem v nule o úhly $k \cdot 2\pi/n$, $k = 0, \dots, n-1$, a z reflexí, jejichž osy procházejí

- pro liché n, vrcholem a středem protilehlé strany,
- pro sudé n, protilehlými vrcholy a protilehlými středy stran.

Vidíme, že $\mathbf{Sym}(X) \simeq \mathbf{D}_{2n}$, kde dané izometrii odpovídá příslušná permutace na očíslovaných vrcholech.

Z druhého příkladu si lze vzít obecné ponaučení: jsou-li izometrie zachovávající daný objekt X určeny hodnotami v n bodech, lze se na grupu $\mathbf{Sym}(X)$ dívat jako na podgrupu grupy $\mathbf{S}_n$. Tento duální pohled budeme používat často, nejen pro dihedrální grupy.

Příklad. Uvažujme krychli $X \subset \mathbb{R}^3$ se středem v nule. Zřejmě $\mathbf{Sym}(X) \leq \mathbf{Sym}(\{0\})$. Jak vypadají rotace, které zachovávají krychli?

Nejprve ukážeme, že existuje nejvýše 24 rotací zachovávajících danou krychli. Zvolme vrchol. Ten lze otočit na libovolný z osmi vrcholů krychle. Jeho tři sousedi se ovšem musí zobrazit na sousedy obrazu, a to ve stejném pořadí, čili jsou nejvýše tři možnosti, jak to udělat. Obrazy protilehlých vrcholů jsou těmito čtyřmi jednoznačně určeny, nemůže tedy být více než $8 \cdot 3 = 24$ rotací.

Je snadné nahlédnout, že následujících 24 rotací krychli zachovává:

- identita,
- rotace s osou přes středy protilehlých stěn o úhly 90, 180 a 270 stupňů,
- rotace s osou přes středy protilehlých hran o úhel 180 stupňů,
- rotace s osou přes protilehlé vrcholy o úhly 120 a 240 stupňů.

Čili podgrupa rotací $\mathbf{R} \leq \mathbf{Sym}(X)$ má 24 prvků. Všimněte si, že různé rotace permutují různým způsobem tělesové úhlopříčky, čili grupa $\mathbf{R}$ je izomorfní nějaké podgrupě grupy $\mathbf{S}_4$ (rotaci přiřaďte permutaci na očíslovaných úhlopříčkách). Protože mají obě grupy 24 prvků, nutně musí platit $\mathbf{R} \simeq \mathbf{S}_4$.

Zvolme pevně jednu reflexi $\sigma_0 \in \mathrm{Sym}(X)$. Pro každou další reflexi $\sigma \in \mathrm{Sym}(X)$ platí $\sigma \circ \sigma_0^{-1} \in R$, protože složením reflexí (determinant příslušné matice je -1) dostaneme rotaci (determinant 1), čili podle Tvrzení 14.10 patří σ do rozkladové třídy $\sigma_0 R$. Dokázali jsme, že všechny reflexe tvoří jednu rozkladovou třídu, čili $[\mathbf{Sym}(X):\mathbf{R}]=2$, a tedy $\mathbf{Sym}(X)$ má 48 prvků. V pokročilejším kurzu teorie grup se dozvíte, jak tuto grupu reprezentovat pomocí konstrukce zvané semidirektní součin.

17.2. **Automorfismy matematických struktur.**

Motivací pro rozvoj teorie grup bylo studium symetrií nejen geometrických, ale také algebraických a kombinatorických objektů. Standardní definicí symetrie je pojem *automorfismu.* Obecně lze říci, že automorfismem matematického objektu $\mathbf{X}=(X,\dots)$ (například okruhu, tělesa, grupy, vektorového prostoru, grafu, metrického prostoru, topologického prostoru atd.) se rozumí permutace množiny X, která *zachovává strukturu tohoto objektu.* Pro algebraické struktury jde o bijektivní homomorfismy (tj. izomorfismy na sebe sama) ve smyslu sekcí 1.3, 15.1. Pro grafy jde o bijektivní zobrazení, která zachovávají hrany i nehrany. Pro spojité struktury jde o tzv. homeomorfismy, tedy bijektivní zobrazení, která jsou spojitá v obou směrech.

Automorfismy daného objektu vždy tvoří podgrupu grupy $\mathbf{S}_X$, nazývanou *grupa automorfismů* a označovanou $\mathbf{Aut}(\mathbf{X})$.

Příklad. Automorfismem grafu (V,E) se rozumí permutace φ na množině vrcholů V taková, že $(x,y)\in E$ právě tehdy, když $(\varphi(x),\varphi(y))\in E$. Je snadné nahlédnout, že automorfismy daného grafu skutečně tvoří podgrupu grupy $\mathbf{S}_V$. Například,

- grupa automorfismů úplného grafu na n vrcholech je grupa $\mathbf{S}_n$,
- grupa automorfismů n-prvkové kružnice je dihedrální grupa $\mathbf{D}_{2n}$,
- grupa automorfismů cesty délky $n\geq 2$ je dvouprvková.

Příklad. Automorfismem vektorového prostoru $\mathbf{V}$ nad tělesem $\mathbf{T}$ se rozumí bijektivní lineární zobrazení $\mathbf{V}\to\mathbf{V}$. Z lineární algebry víte, že pro prostory konečné dimenze automorfismy vzájemně jednoznačně odpovídají regulárním maticím, přičemž složení automorfismů odpovídá součinu příslušných matic. Čili $\mathbf{Aut}(\mathbf{V})\simeq\mathbf{GL}_n(\mathbf{T})$, kde $n=\dim\mathbf{V}$.

Příklad. Jak uvidíme v sekci 25, celá Galoisova teorie a důkaz neřešitelnosti polynomiálních rovnic v radikálech jsou založeny na studiu grup automorfismů těles.

Sekci zakončíme stručnou informací o grupových automorfismech. V neabelovských grupách pochází spousta symetrie z konjugace: je-li $\mathbf{G}$ grupa a $a\in G$, pak zobrazení

$$\varphi_a: G\to G,\qquad x\mapsto axa^{-1}$$

je automorfismem grupy $\mathbf{G}$. Tyto automorfismy se nazývají *vnitřní* a je snadné ověřit, že tvoří podgrupu grupy $\mathbf{Aut}(\mathbf{G})$, značenou $\mathbf{Inn}(\mathbf{G})$.

Příklad. Dokážeme že $\mathbf{Aut}(\mathbf{S}_3)=\mathbf{Inn}(\mathbf{S}_3)\simeq\mathbf{S}_3$. Protože $\mathbf{S}_3=\langle(1\ 2),(2\ 3)\rangle$, každý její automorfismus je určený hodnotami na těchto dvou transpozicích. Ty se mohou zobrazit pouze na prvky řádu 2 (Tvrzení 15.4), tedy v tomto případě transpozice, čili máme nejvýše $3\cdot 2=6$ možností, jak to udělat. Na druhou stranu, není těžké nahlédnout, že vnitřní automorfismy jsou po dvou různé, čili je jich šest a $\mathbf{Aut}(\mathbf{S}_3)=\mathbf{Inn}(\mathbf{S}_3)$. Jako cvičení si dokažte, že pro všechna n platí $\mathbf{Inn}(\mathbf{S}_n)\simeq\mathbf{S}_n$.

Podobné tvrzení platí pro všechny grupy $\mathbf{S}_n$ kromě $n=6$, ale je to o dost těžší dokázat. Grupa $\mathbf{S}_6$ je výjimečná: její grupa automorfismů je dvakrát větší než podgrupa vnitřních automorfismů.

Operace, kterou provádějí vnitřní automorfismy, se nazývá *konjugace*. Formálně, buď $\mathbf{G}$ grupa a $a,b\in G$. Prvky a,b nazýváme *konjugované* v $\mathbf{G}$, pokud existuje $c\in G$ takové, že $a=cbc^{-1}$. Je vidět, že relace konjugace je ekvivalencí. Její bloky se nazývají *třídy konjugace.*

Příklad. Dvě permutace jsou konjugované v grupě $\mathbf{S}_n$ právě tehdy, když mají stejnou strukturu cyklů, viz Tvrzení 13.1.

Příklad. V lineární algebře se konjugovaným maticím se říká *podobné.* Konjugace odpovídá změně báze lineárního zobrazení, tj. dvě matice jsou konjugované v grupě $\mathbf{GL}_n(\mathbf{T})$ právě tehdy, když jsou maticí téhož lineárního zobrazení vzhledem k jistým bázím.

Příklad. Permutace (1 2 3) a (2 3 4) jsou konjugované v grupě $\mathbf{S}_4$, protože obě mají jeden cyklus délky 1 a jeden cyklus délky 3. Tyto permutace ovšem nejsou konjugované v grupě $\mathbf{A}_4$: jak plyne z důkazu Tvrzení 13.1, jediné permutace ρ splňující $(2\ 3\ 4) = \rho \circ (1\ 2\ 3) \circ \rho^{-1}$ jsou (1 4), (1 2 3 4) a (1 3 2 4). Žádná z nich ovšem není sudá.

18. Působení grupy na množině

18.1. **Abstraktní grupa jako grupa permutací.**

V mnoha situacích se hodí interpretovat danou abstraktní grupu jako grupu permutací na jisté množině. Například číselnou grupu $\mathbb{Z}_n$ lze interpretovat jako grupu permutací roviny, kde číslu k odpovídá rotace o úhel $k \cdot 2\pi/n$. Součet dvou čísel modulo n odpovídá složení příslušných rotací, opačné číslo odpovídá opačné rotaci a nula identické permutaci, čili jde o homomorfismus. Toto pozorování motivuje následující definici.

Definice. *Působením grupy* $\mathbf{G}$ *na množině* X rozumíme libovolný homomorfismus $\pi : G \to S_X$. Hodnotu permutace $\pi(g)$ na prvku $x \in X$ budeme značit krátce $g(x)$.

Z definice homomorfismu plyne, že jednotka v $\mathbf{G}$ působí jako identita, g^{-1} působí jako inverzní permutace k $\pi(g)$ a platí vztah $(g \cdot h)(x) = g(h(x))$. Můžeme si představovat, že prvky grupy $\mathbf{G}$ „hýbou“ s prvky množiny X, přičemž jak se prvky v $\mathbf{G}$ násobí, tak se příslušné „pohyby“ skládají.

Příklad. Triviálním případem je přirozené působení permutační grupy $\mathbf{G} \leq \mathbf{S}_X$ na množinu X, kde $\pi(g) = g$.

Příklad. Působení z úvodního odstavce odpovídá následující konfiguraci: $\mathbf{G} = \mathbb{Z}_n$, $X = \mathbb{R}^2$ a $\pi(k)$ je permutace na X daná předpisem

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} \cos(k \cdot 2\pi/n) & -\sin(k \cdot 2\pi/n) \\ \sin(k \cdot 2\pi/n) & \cos(k \cdot 2\pi/n) \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

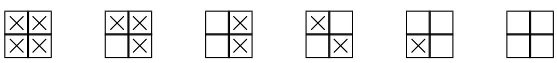
Příklad. Maticové grupy lze interpretovat jako permutace příslušného vektorového prostoru dané příslušným lineárním zobrazením: zde $\mathbf{G} \leq \mathbf{GL}_n(\mathbf{T})$, $X = T^n$ a $\pi(A)$ je permutace množiny T^n, která vektor v zobrazí na součin Av.

Příklad. Cayleovu reprezentaci (Věta 15.10) lze interpretovat tak, že daná grupa $\mathbf{G}$ působí na množině svých prvků G translací: $g(x) = L_g(x) = gx$. Jsou i jiná přirozená působení grupy na množině svých prvků, například konjugací, kde $g(x) = \varphi_g(x) = gxg^{-1}$ (viz cvičení). Tato a podobná působení umožňují dokázat zajímavé vlstnosti konečných grup.

Jako motivaci, proč uvažovat abstraktní koncept působení grupy na množině, si ukážeme jednu pěknou aplikaci v kombinatorice. Jak spočítat jistý typ objektů až na dané symetrie? Například, kolika způsoby je možné obarvit stěny krychle dvěma barvami až na otočení, tj. když dvě obarvení považujeme za totožná, pokud jedno z druhého dostaneme rotací krychle? Pro jednoduchost budeme metodu ilustrovat v dvojrozměrném případě.

Úloha. Kolika způsoby je možné obarvit políčka čtvercové tabulky 2×2 dvěma barvami až na otočení? Tj. dvě obarvení považujeme za totožná, pokud jedno z druhého dostaneme otočením tabulky.

Tuto úlohu je snadné řešit prostým výčtem všech možných obarvení, vyjde nám následujících šest:



Ale při větším počtu barev nebo větším počtu políček bychom se nedopočítali. Nejprve si ujasněme, co přesně znamená počítání „až na danou symetrii“. Dva objekty považujeme za totožné, pokud jeden z druhého dostaneme pomocí nějakého povoleného zobrazení. V naší úloze jsou to rotace, které zachovávají daný čtverec, tj. rotace roviny o 0, 90, 180 a 270 stupňů kolem středu čtverce. Uvažujme tedy působení grupy $\mathbf{G}$ sestávající z těchto čtyřech rotací na množině X sestávající ze včech možných obarvení čtverce 2×2 dvěma barvami (čili $|X| = 2^4 = 16$), kde $\pi(g)$ je permutace, která otočí tabulku o příslušný úhel i s daným obarvením.

Nyní zpět k teorii. V celém zbytku sekce budeme uvažovat libovolné působení grupy $\mathbf{G}$ na množinu X. Budeme potřebovat několik užitečných definic a vlastností.

Definice. Zavedeme tzv. *relaci tranzitivity* $\sim$ na množině X: definujeme $x \sim y$, pokud existuje $g \in G$ takové, že $g(x) = y$.

Volně řečeno, $x \sim y$, pokud nějaká permutace přesouvá prvek x na prvek y.

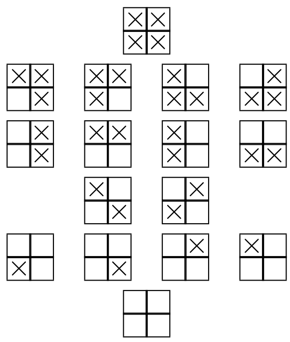
Lemma 18.1. *Relace $\sim$ je ekvivalence na množině X.*

Důkaz. Reflexivita: jednotka působí jako identita, tj. $1(x) = id(x) = x$. Symetrie: inverz prvku působí jako inverzní permutace, tj. $g(x) = y$ implikuje $g^{-1}(y) = x$. Tranzitivita: násobení odpovídá skládání permutací, čili pokud $x \sim y \sim z$, tj. $g(x) = y$ a $h(y) = z$ pro nějaká $g, h \in G$, pak $(h \cdot g)(x) = h(g(x)) = h(y) = z$, a tedy $x \sim z$. $\square$

Bloky ekvivalence $\sim$ nazýváme *orbity*. Orbitu obsahující prvek x budeme značit

$$[x] = \{y \in X : \ x \sim y\} = \{g(x) : \ g \in G\}.$$

Příklad. V motivační úloze jsou v relaci $\sim$ právě ty dvojice obravení, kde lze jedno z druhého dostat otočením. Množina všech obarvení se tedy rozpadne na šest orbit následujícím způsobem:



Řešením motivační úlohy je *počet orbit* v tomto působení.

Definice. Bod $x \in X$ se nazývá *pevným bodem* prvku $g \in G$, pokud $g(x) = x$. Množinu všech pevných bodů prvku $g \in G$ budeme značit

$$X_g = \{x \in X : \ g(x) = x\}$$

a *stabilizátorem prvku $x \in X$* nazveme množinu

$$G_x = \{g \in G : \ g(x) = x\}.$$

Příklad. Stabilizátorem obou jednobarevných obarvení je celá grupa $\mathbf{G}$. Stabilizátor obarvení obsahuje pouze identitu. Stabilizátor obarvení obsahuje identitu a rotaci o 180 stupňů.

Lemma 18.2. *Stabilizátor G_x tvoří podgrupu grupy $\mathbf{G}$.*

Důkaz. Jednotka náleží G_x, neboť $1(x) = id(x) = x$. Pokud $g(x) = x$, pak také $g^{-1}(x) = x$. A pokud $g, h \in G_x$, tj. platí $g(x) = h(x) = x$, pak $(gh)(x) = g(h(x)) = g(x) = x$, čili $gh \in G_x$. $\square$

Zásadní význam má následující tvrzení, které dává do souvislosti velikost stabilizátoru a velikost orbity.

Tvrzení 18.3 (velikost orbity vs. index stabilizátoru)**.** *Nechť grupa $\mathbf{G}$ působí na množině X. Pak pro každé $x \in X$ platí*

$$\big|[x]\big| = [\mathbf{G} : \mathbf{G}_x].$$

Důkaz. Index $[\mathbf{G} : \mathbf{G}_x]$ značí počet rozkladových tříd podgrupy $\mathbf{G}_x$, stačí tedy najít bijekci mezi prvky orbity a množinou rozkladových tříd. Uvažujme zobrazení

$$\varphi : \{gG_x :\ g \in G\} \to [x], \qquad gG_x \mapsto g(x).$$

Dokážeme, že to je bijekce. Předně je třeba ověřit, že jsme dobře definovali zobrazení: mohlo by se stát, že tutéž rozkladovou třídu máme označenu dvěma různými způsoby, tj. že $gG_x = hG_x$, a přitom se jí snažíme přiřadit různé hodnoty $g(x) \neq h(x)$. Ovšem podle Tvrzení 14.10 platí

$$gG_x = hG_x \ \Leftrightarrow\ h^{-1}g \in G_x \ \Leftrightarrow\ h^{-1}g(x) = x \ \Leftrightarrow\ g(x) = h(x),$$

a tedy φ je nejen dobře definované, ale také prosté. Navíc pro každý prvek $y \in [x]$ existuje $g \in G$ splňující $g(x) = y$, tedy φ je bijekce. $\square$

Je-li grupa $\mathbf{G}$ konečná, kombinací Tvrzení 18.3 a Lagrangeovy věty dostáváme vztah

$$|\mathbf{G}| = |\mathbf{G}_x| \cdot |[x]|.$$

Speciálně, velikost každé orbity dělí řád grupy $\mathbf{G}$ (všimněte si, že to je splněno v naší motivační úloze).

18.2. **Burnsideova věta a počítání orbit.**
Označme $X/\!\sim$ množinu všech bloků ekvivalence $\sim$ na množině X. V našem kontextu bude $|X/\!\sim|$ značit počet orbit daného působení.

Věta 18.4 (Burnsideova věta)**.** *Nechť konečná grupa* $\mathbf{G}$ *působí na konečnou množinu* X*. Pak*

$$\big|X/\!\sim\big| = \frac{1}{|G|} \cdot \sum_{g \in G} \big|X_g\big|.$$

Vzorec lze interpretovat tak, že *počet orbit je roven průměrnému počtu pevných bodů*, kde průměr počítáme přes všechny prvky grupy $\mathbf{G}$.

Důkaz. Označme

$$M = \big\{(g, x) \in G \times X :\ g(x) = x\big\}.$$

Prvky této množiny můžeme spočítat dvěma způsoby: buď ke každému g spočítáme počet x takových, že $(g, x) \in M$, nebo naopak, ke každému x spočítáme počet g takových, že $(g, x) \in M$. Dostaneme tak následující rovnost:

$$|M| = \sum_{g \in G} |X_g| = \sum_{x \in X} |G_x|.$$

Použitím této rovnosti dopočítáme uvedený vzorec:

$$\frac{1}{|G|} \cdot \sum_{g \in G} |X_g| \ = \ \frac{1}{|G|} \cdot \sum_{x \in X} |G_x| \ \stackrel{18.3}{=} \ \frac{1}{|G|} \cdot \sum_{x \in X} \frac{|G|}{\big|[x]\big|} \ = \ \sum_{x \in X} \frac{1}{\big|[x]\big|} \ =$$

$$\sum_{O \in (X/\sim)} \sum_{x \in O} \frac{1}{\big|[x]\big|} \ = \ \sum_{O \in (X/\sim)} \sum_{x \in O} \frac{1}{|O|} \ = \ \sum_{O \in (X/\sim)} |O| \cdot \frac{1}{|O|} \ = \ \sum_{O \in (X/\sim)} 1.$$

Výsledek je tedy roven velikosti množiny $X/\!\sim$. $\square$

Příklad. Vraťme se k motivační úloze. Rotace o 0 stupňů (tj. identita) zachovává všechna obarvení, tedy $|X_0| = |X| = 16$. Rotace o 90 stupňů zobrazuje levé dolní políčko na levé horní, levé horní na pravé horní, atd., čili abychom dostali stejné obarvení, musí mít všechna čtyři políčka stejnou barvu. Tedy $|X_{90}| = 2$. Podobně $|X_{270}| = 2$. Rotace o 180 stupňů zaměňuje levé dolní políčko za pravé horní a levé horní za pravé dolní. Tyto dvě dvojice tedy musí být stejnobarevné a to lze provést čtyřmi způsoby. Tedy $|X_{180}| = 4$. Podle Burnsideovy věty je počet obarvení až na otočení $\frac{1}{4} \cdot (16 + 2 + 4 + 2) = 6$.

Metodu ilustrujeme na několika dalších úlohách.

Úloha. (a) Dětská stavebnice obsahuje tři červené, tři zelené a tři modré čtvercové destičky. Kolika způsoby je lze sestavit do velkého čtverce 3×3? Dvě sestavy považujeme za totožné, pokud jednu z druhé dostaneme otočením. (b) Jak se výsledek změní, pokud je možné dílky pevně spojovat? Tedy pokud dvě sestavy považujeme za totožné, dostaneme-li jednu z druhé otočením a převrácením.

Řešení. Místo sestav budeme uvažovat barvení jednotlivých políček čtverce. Čili X bude množina všech obarvení čtverce 3×3 daným počtem barev a $\mathbf{G}$ bude (a) grupa $\mathbb{Z}_4$ interpretovaná jako rotace čtverce, (b) grupa $\mathbf{D}_8$ všech symetrií čtverce. Grupa $\mathbf{G}$ působí na X tak, že příslušná permutace otočí/převrátí čtverec i s jeho obarvením. Řešením úlohy je počet orbit tohoto působení (dvě obarvení jsou v jedné orbitě právě tehdy, když jedno z druhého dostaneme otočením, resp. převrácením).

Napíšeme tabulku, v jejímž prvním sloupci bude seznam prvků grupy $\mathbf{G}$, přičemž zobrazení „podobného typu“ budeme sdružovat (rozumí se, že prvky „podobného typu“ mají stejně velké množiny pevných bodů), v druhém sloupci bude počet prvků daného typu a ve třetím počet pevných bodů těchto prvků. Pevným bodem se rozumí takové obarvení, které po daném otočení/převrácení vypadá stejně.

g	#	$\lvert X_g\rvert$
id	1	1680
rotace o $\pm\, 90°$	2	0
rotace o $180°$	1	0
osa přes vrcholy	2	36
osa středem hran	2	36

Podle Burnsideovy věty je počet obarvení

(a) $\frac{1}{4} \cdot (1680 + 2 \cdot 0 + 1 \cdot 0) = 420$,
(b) $\frac{1}{8} \cdot (1680 + 2 \cdot 0 + 1 \cdot 0 + 2 \cdot 36 + 2 \cdot 36) = 228$.

□

Úloha. Kolik náhrdelníků lze sestavit (a:Sparta) ze tří červených, tří žlutých a tří modrých kuliček, (b:Bohemians) z šesti zelených a tří bílých kuliček? (Nezáleží na poloze náhrdelníku, je možno jej převracet či otáčet.)

Řešení. Náhrdelník reprezentujeme jako obarvení vrcholů pravidelného devítiúhelníka. Čili X, resp. Y, bude množina všech obarvení vrcholů pravidelného devítiúhelníka danými barvami a $\mathbf{G} = \mathbf{D}_{18}$ bude grupa všech symetrií pravidelného devítiúhelníka, která působí na X, resp. Y, tak, že příslušná permutace otočí/převrátí devítiúhelník i s jeho obarvením. Každé orbitě tohoto působení odpovídá právě jeden náhrdelník (jehož kuličky jsou uspořádány podle toho obarvení). Napíšeme tabulku podobně jako v předchozí úloze.

g	#	$\lvert X_g\rvert$	$\lvert Y_g\rvert$
id	1	1680	84
rotace o $\pm\, 1$ vrchol	2	0	0
rotace o $\pm\, 2$ vrcholy	2	0	0
rotace o $\pm\, 3$ vrcholy	2	6	3
rotace o $\pm\, 4$ vrcholy	2	0	0
reflexe	9	0	4

Podle Burnsideovy věty je počet náhrdelníků (a) $\frac{1}{18} \cdot (1680 + 2 \cdot 6) = 94$, resp. (b) $\frac{1}{18} \cdot (84 + 2 \cdot 3 + 9 \cdot 4) = 7$. □

Úloha. Kolika způsoby je možné obarvit stěny krychle dvěma barvami? Kolika způsoby lze přiřadit stěnám čísla 1 až 6? A kolik existuje hracích kostek, tj. kolika způsoby lze přiřadit čísla 1 až 6 tak, že součet protilehlých stěn je sedm? Dvě obarvení/přiřazení považujeme za totožná, pokud lze jedno z druhého dostat otočením krychle.

Řešení. Buď X množina všech obarvení stěn krychle dvěma barvami, Y množina všech přiřazení čísel 1 až 6 stěnám a Z množina těch přiřazení z Y, jejichž protilehlé stěny dávají součet sedm. $\mathbf{G}$ bude grupa všech rotací krychle působící na X, Y i Z tak, že příslušná permutace otočí krychli i s jejím obarvením/přiřazením. Napíšeme tabulku podobně jako v předchozí úloze.

g	#	$\|X_g\|$	$\|Y_g\|$	$\|Z_g\|$
identita	1	2^6	$6!$	48
osa přes středy protilehlých stěn, $\pm 90°$	6	2^3	0	0
osa přes středy protilehlých stěn, $+180°$	3	2^4	0	0
osa přes středy protilehlých hran, $+180°$	6	2^3	0	0
osa přes protilehlé vrcholy, $\pm 120°$	8	2^2	0	0

Tedy počty orbit jsou

- $|X/\sim| = \frac{1}{24} \cdot (2^6 + 3 \cdot 2^4 + 12 \cdot 2^3 + 8 \cdot 2^2) = 10$,
- $|Y/\sim| = \frac{1}{24} \cdot 6! = 30$,
- $|Z/\sim| = \frac{1}{24} \cdot 48 = 2$.

Jak známo, hrací kostky jsou dvě, pravotočivá a levotočivá, podle pořadí stěn 1, 2, 3 při pohledu na roh kostky. který tato čísla sdílí. □

Burnsideovu větu lze použít v řadě dalších aplikací, např. pokud chceme zjistit *počet nějakých struktur dané velikosti až na izomorfismus.* Metodu ilustrujeme na grafech se čtyřmi vrcholy.

Příklad. Buď X množina všech grafů s vrcholy $1, 2, 3, 4$. Dva grafy jsou izomorfní, pokud existuje permutace z $\mathbf{S}_4$, která převádí hrany na hrany a mezery na mezery. Uvažujme tedy působení grupy $\mathbf{S}_4$ na X tak, že daná permutace přehází vrcholy i s hranami, které do nich vedou. Orbity tohoto působení budou obsahovat právě všechny navzájem izomorfní grafy, počet neizomorfních grafů je tedy roven počtu orbit. Řešením je tabulka

g	#	$\|X_g\|$
id	1	2^6
(..)	6	2^4
(..)(..)	3	2^4
(...)	8	2^2
(....)	6	2^2

Vidíme, že čtyřprvkových grafů je 11.

Na závěr ukážeme jednu zajímavost s elegantním důkazem. Permutační grupa se nazývá *tranzitivní*, má-li jen jednu orbitu (ve svém přirozeném působení). Např. grupy $\mathbf{S}_n$, $\mathbf{A}_n$, $\mathbf{D}_{2n}$ jsou tranzitivní, grupa $\langle (1\ 2)(3\ 4) \rangle_{\mathbf{S}_4}$ není.

Věta 18.5 (Jordanova věta)**.** *Každá alespoň dvouprvková konečná tranzitivní grupa obsahuje alespoň jednu permutaci bez pevného bodu.*

Důkaz. Podle Burnsideovy věty je počet orbit roven průměrnému počtu pevných bodů. Tranzitivita říká, že počet orbit je 1. Přitom identita má alespoň dva pevné body, tedy *nadprůměrné* množství, takže musí existovat permutace, která má *podprůměrné* množství pevných bodů. Protože je počet pevných bodů nezáporné celé číslo, jediná podprůměrná hodnota je 0. Tedy existuje permutace bez pevného bodu. □

18.3. **Cauchyova věta.**

Tvrzení 18.3 lze použít k důkazu spousty užitečných tvrzení o konečných grupách. Jedno takové si nyní ukážeme.

Lagrangeova věta říká, že řád každého prvku dělí řád celé grupy. Naopak, pokud k dělí $|G|$, prvek řádu k v grupě $\mathbf{G}$ existovat nemusí. Leda by ovšem k bylo prvočíslo, pak musí. Následující věta má řadu důsledků v teorii konečných grup a i my ji budeme potřebovat v kapitole o Galoisově teorii k důkazu, že jisté polynomy stupně 5 nemají vzorec na vyjádření kořenů.

Věta 18.6 (Cauchyova věta)**.** *Buď* **G** *konečná grupa a p prvočíslo, které dělí $|G|$. Pak v* **G** *existuje prvek řádu p.*

Důkaz. Označme
$$X = \{(a_1, \ldots, a_p) \in G^p : \; a_1 \cdot \ldots \cdot a_p = 1\}.$$
Prvních $p-1$ prvků v každé p-tici můžeme zvolit libovolně a poslední je jimi jednoznačně určen, čili $|X| = |G|^{p-1}$ a vidíme, že $p \mid |X|$. Grupa $\mathbb{Z}_p$ působí na X rotací složek. Podle Tvrzení 18.3 velikost každé orbity dělí řád působící grupy, čili možné velikosti orbit jsou pouze 1 a p. Přitom aspoň jedna orbita velikosti 1 existuje, konkrétně $\{(1, \ldots, 1)\}$. Protože je velikost $|X|$ dělitelná p, musí existovat ještě aspoň $p-1$ jednoprvkových orbit. Přitom v jednoprvkové orbitě může být pouze konstantní p-tice, nějaká $(a, \ldots, a)$, která splňuje $a^p = a \cdot \ldots \cdot a = 1$, čímž jsme objevili hledaný prvek řádu p. □

19. Faktorgrupy

19.1. Normální podgrupy.

Předmětem této sekce je velmi důležitá konstrukce *faktorgrup*. Jako parametr slouží jistý typ podgrup, zvaných normální. S tímto pojmem se nyní seznámíme.

Tvrzení 19.1 (ekvivalentní definice normální podgrupy)**.** *Buď* **G** *grupa a* **H** *její podgrupa. Následující tvrzení jsou ekvivalentní:*

1. *$aH = Ha$ pro každé $a \in G$ (tj. levé a pravé rozkladové třídy daného prvku jsou stejné),*
2. *$aha^{-1} \in H$ pro každé $h \in H$ a každé $a \in G$ (tj. je uzavřena na konjugaci libovolným prvkem).*

Důkaz. (1) ⇒ (2). Buď $h \in H$ a $a \in G$. Pak $ah \in aH = Ha$, a tedy existuje $k \in H$ takové, že $ah = ka$. Dostáváme $aha^{-1} = k \in H$.

(2) ⇒ (1). Dokážeme obě inkluze v rovnosti $aH = Ha$. Nejprve uvažujme $ah \in aH$. Pak $k = aha^{-1} \in H$, a tedy $ah = ka \in Ha$. Nyní uvažujme $ha \in Ha$. Pak $l = a^{-1}ha \in H$, tedy $ha = al \in aH$. □

Definice. Podgrupa **H** se nazývá *normální* v grupě **G**, pokud splňuje ekvivalentní podmínky formulované v Tvrzení 19.1. Tento fakt značíme $\mathbf{H} \trianglelefteq \mathbf{G}$.

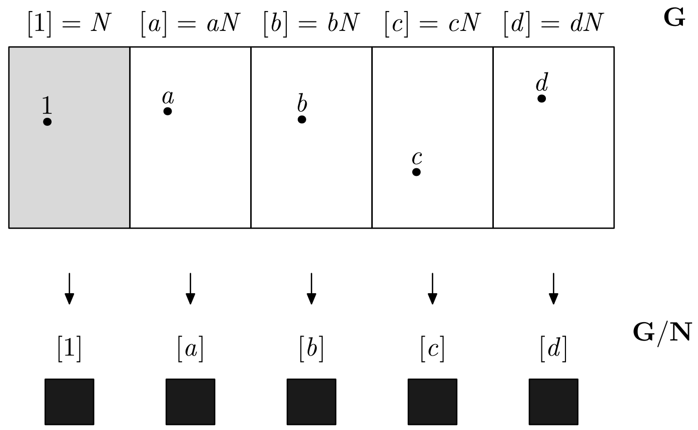
V abelovských grupách je každá podgrupa normální, obě ekvivalentní podmínky jsou triviálně splněny. Z triviálních důvodů platí také $\{1\} \trianglelefteq \mathbf{G}$ a $\mathbf{G} \trianglelefteq \mathbf{G}$.

Příklad.

- Podgrupa $\mathbf{SL}_n(\mathbf{T})$ matic s determinantem 1 je normální v grupě $\mathbf{GL}_n(\mathbf{T})$, jak plyne z podmínky (2) užitím součinového vzorce pro determinanty: $\det(AHA^{-1}) = (\det A)(\det H)(\det A)^{-1} = \det H$.
- Podgrupa $\mathbf{A}_n$ sudých permutací je normální v grupě $\mathbf{S}_n$, jak plyne ze součinového vzorce pro znaménko: $\operatorname{sgn}(aha^{-1}) = (\operatorname{sgn} a)(\operatorname{sgn} h)(\operatorname{sgn} a)^{-1} = \operatorname{sgn} h$.
- Podgrupa $\mathbf{D}_{2n}$ není normální v grupě $\mathbf{S}_n$, což je vidět z Tvrzení 13.1.

Možné jsou oba extrémy: jsou grupy, kde je každá podgrupa normální (kromě abelovských například kvaternionová grupa $\mathbf{Q}_8$), ale také grupy, které obsahují pouze nevlastní normální podgrupy (například grupy $\mathbf{A}_n$, $n \neq 4$). Grupy, které nemají žádné vlastní normální podgrupy, se nazývají *jednoduché*. Jedním z největších algebraických výsledků 20. století je klasifikace (tzn. úplný seznam až na izomorfismus) všech konečných jednoduchých grup.

Příklad. Jediné normální podgrupy v grupě $\mathbf{S}_n$, $n \neq 4$, jsou $\{1\}$, $\mathbf{A}_n$, $\mathbf{S}_n$. Grupa $\mathbf{S}_4$ navíc obsahuje čtyřprvkovou normální podgrupu
$$\mathbf{K}_4 = \{id, (12)(34), (13)(24), (14)(23)\},$$
které se říká *Kleinova*. Důkaz pro $n = 3, 4, 5$ si proveďte jako cvičení za pomoci Tvrzení 13.1 a 14.5.



Obrázek 19. Konstrukce faktorgrupy

Na závěr uvedeme jedno důležité pozorování doplňující Tvrzení 15.2.

Tvrzení 19.2. *Jádro homomorfismu je normální podgrupa.*

Důkaz. Uvažujme $\varphi : \mathbf{G} \to \mathbf{H}$. V Tvrzení 15.2 jsme dokázali, že $\mathbf{Ker}(\varphi)$ je podgrupa. Normalita plyne z toho, že pro libovolné $a \in \mathrm{Ker}(\varphi)$ a $u \in G$ platí $\varphi(uau^{-1}) = \varphi(u) * \varphi(a) * \varphi(u)' = \varphi(u) * \varphi(u)' = e$. □

19.2. **Konstrukce faktorgrupy.**

V různých odvětvích matematiky se opakuje myšlenka konstrukce faktorobjektu. Neformálně řečeno, je dán objekt s velmi jemnou strukturou (hvězdy na obloze). Pokud od objektu poodstoupíme, některé prvky splynou (při pohledu pouhým okem například hvězdy v jedné galaxii). To, co vidíme, je faktorobjekt původního objektu (světlé body na obloze). O trochu formálněji, ztotožníme *podobné* objekty (na obloze ty, které jsou příliš blízko). Co se přesně myslí relací podobnosti už závisi na konkrétním typu objektu.

V sekci 9.2 jsme se seznámili se speciálním případem konstrukce faktorokruhu: dva polynomy jsme prohlásili za podobné, pokud dávají stejný zbytek modulo daný polynom m, a s reprezentanty zbytků jsme počítali modulo m. Tato konstrukce funguje obecněji: polynom m můžeme zaměnit za libovolný ideál I a dva prvky prohlásíme za podobné, pokud je jejich rozdíl v I. Analogickou myšlenku lze použít i pro grupy, kde místo ideálu budeme uvažovat normální podgrupu $\mathbf{N}$ a dva prvky prohlásíme za podobné, pokud je jejich podíl v $\mathbf{N}$.

Definice. Buď $\mathbf{G}$ grupa a $\mathbf{N}$ její normální podgrupa. Definujeme relaci na množině G předpisem

$$a \sim b \quad \Leftrightarrow \quad a \cdot b^{-1} \in N.$$

Podle Tvrzení 14.10 je $a \sim b$ právě tehdy, když $Na = Nb$, čili relace $\sim$ je ekvivalencí na množině G. Její bloky jsou rozkladové třídy grupy $\mathbf{G}$ podle podgrupy $\mathbf{N}$, a protože je $\mathbf{N}$ normální, levé i pravé rozkladové třídy jsou totéž (Tvrzení 19.1), čili

$$[a] = aN = Na.$$

Na těchto blocích definujeme operace předpisy

$$[a] \cdot [b] = [a \cdot b] \qquad \text{a} \qquad [a]^{-1} = [a^{-1}]$$

(v následujícím lemmatu ověříme, že je tato definice korektní, tj. že výsledek operace nezávisí na tom, kterým prvkem si daný blok označíme), za jednotkový prvek vezmeme blok $[1] = N$. Množina bloků s výše uvedenými operacemi se nazývá *faktorgrupa grupy* $\mathbf{G}$ *podle podgrupy* $\mathbf{N}$,

$$\mathbf{G}/\mathbf{N} = (\{[a] : a \in G\}, \cdot, ^{-1}, [1]).$$

Lemma 19.3. *Buď* $\mathbf{G}$ *grupa a* $\mathbf{N}$ *její normální podgrupa.*

(1) *Výše uvedené operace na blocích jsou dobře definovány.*

(2) *Faktorgrupa* $\mathbf{G}/\mathbf{N}$ *je skutečně grupa.*

Důkaz. (1) Uvažujme dva bloky označené dvěma způsoby, $[a] = [c]$ a $[b] = [d]$. Ověříme, že $[a \cdot b] = [c \cdot d]$ a $[a^{-1}] = [c^{-1}]$. Protože $a \sim c$ a $b \sim d$, tj. $a \cdot c^{-1} \in N$ a $b \cdot d^{-1} \in N$, z uzavřenosti množiny N na násobení i konjugaci libovolným prvkem dostáváme

$$(ab) \cdot (cd)^{-1} = abd^{-1}c^{-1} = ac^{-1}cbd^{-1}c^{-1} = \underbrace{(ac^{-1})}_{\in N} \cdot \underbrace{c(bd^{-1})c^{-1}}_{\in N} \in N,$$

čili $a \cdot b \sim c \cdot d$, tj. $[a \cdot b] = [c \cdot d]$. Pro inverz stačí využít faktu, že $ac^{-1} \in N \;\Leftrightarrow\; a^{-1}c \in N$, protože levé i pravé rozkladové třídy jsou stejné, a tedy $Na = Nc \;\Leftrightarrow\; aN = cN$.

(2) Ověříme, že $\mathbf{G}/\mathbf{N}$ splňuje axiomy grup. Operace $\cdot$ je asociativní, neboť $[a] \cdot ([b] \cdot [c]) = [a \cdot (b \cdot c)] = [(a \cdot b) \cdot c] = ([a] \cdot [b]) \cdot [c]$, a podobně se ověří i $[a] \cdot [1] = [a \cdot 1] = [a] = [1 \cdot a] = [1] \cdot [a]$ a $[a] \cdot [a]^{-1} = [a \cdot a^{-1}] = [1] = [a]^{-1} \cdot [a]$. □

Příklad. Uvažujme grupu $\mathbf{G} = \mathbb{Z}$ a normální podgrupu $\mathbf{H} = n\mathbb{Z}$. Platí

$$a \sim b \;\Leftrightarrow\; n \mid a - b \;\Leftrightarrow\; a \equiv b \pmod n,$$

bloky této ekvivalence jsou rozkladové třídy

$$[a] = \{k \in \mathbb{Z} : \; k \equiv a \pmod n\} = a + n\mathbb{Z}, \quad a = 0, \ldots, n-1.$$

Přitom $[a] + [b] = [a+b] = [a+b \bmod n]$ a $-[a] = = [-a] = [n-a]$, čili operace na prvcích $\mathbb{Z}/n\mathbb{Z}$ jsou jako operace na číslech $0, \ldots, n-1$ modulo n. Není těžké ověřit, že $[a] \mapsto a$ je izomorfismus $\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}_n$.

Příklad. Uvažujme grupu $\mathbf{G} = \mathbf{S}_n$ a normální podgrupu $\mathbf{H} = \mathbf{A}_n$. Platí

$$\pi \sim \sigma \;\Leftrightarrow\; \pi \circ \sigma^{-1} \in A_n \;\Leftrightarrow\; \mathrm{sgn}(\pi) = \mathrm{sgn}(\sigma),$$

čili tato ekvivalence má právě dva bloky: množinu S sudých permutací a množinu L lichých permutací. Operace na těchto třídách je $S \circ S = L \circ L = S$ a $S \circ L = L \circ S = L$, jde o dvouprvkovou grupu izomorfní grupě $\mathbb{Z}^*$.

Jak jednoduše, ale přitom formálně určit, jak vypadá faktorgupa dané grupy? Pomůže nám následující věta, která dává do souvislosti faktorgrupy a homomorfní obrazy grup.

Věta 19.4 (věta o homomorfismu). *Buď* $\varphi : \mathbf{G} \to \mathbf{H}$ *homomorfismus grup.*

(1) *Je-li* $\mathbf{N} \leq \mathbf{Ker}(\varphi)$ *normální podgrupa grupy* $\mathbf{G}$*, pak je zobrazení*

$$\psi : \mathbf{G}/\mathbf{N} \to \mathbf{H}, \qquad [a] \mapsto \varphi(a)$$

dobře definované a je to grupový homomorfismus.

(2) (1. věta o izomorfismu) $\mathbf{G}/\mathbf{Ker}(\varphi) \simeq \mathbf{Im}(\varphi)$.

Důkaz. (1) Předně je třeba ověřit, že je zobrazení ψ dobře definované: mohlo by se stát, že máme tentýž blok označen dvěma různými způsoby, tj. že $[a] = [b]$ pro nějaká $a \neq b$, a přitom se těmto blokům snažíme přiřadit dvě různé hodnoty $\varphi(a) \neq \varphi(b)$. Ovšem

$$[a] = [b] \;\Leftrightarrow\; a \cdot b^{-1} \in N \;\Rightarrow\; a \cdot b^{-1} \in \mathrm{Ker}(\varphi) \;\Leftrightarrow\; \varphi(a \cdot b^{-1}) = 1 \;\Leftrightarrow\; \varphi(a) = \varphi(b),$$

tedy ψ je dobře definované zobrazení. Protože $\psi([a \cdot b]) = \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b) = = \psi([a]) \cdot \psi([b])$, je to homomorfismus.

(2) Použijeme (1) pro $\mathbf{N} = \mathbf{Ker}(\varphi)$. Výsledný homomorfismus je prostý, neboť

$$[a] = [b] \;\Leftrightarrow\; a \cdot b^{-1} \in \mathrm{Ker}(\varphi) \;\Leftrightarrow\; \varphi(a \cdot b^{-1}) = 1 \;\Leftrightarrow\; \varphi(a) = \varphi(b),$$

a uvažujeme-li jej jako zobrazení $\mathbf{G}/\mathbf{Ker}(\varphi) \to \mathbf{Im}(\psi) = \mathbf{Im}(\varphi)$, pak je také na. □

1. věta o izomorfismu je dobrým nástrojem, pokud chceme určit, jak vypadá daná faktorgrupa. Chceme-li dokázat, že $\mathbf{G}/\mathbf{N} \simeq \mathbf{H}$, stačí najít homomorfismus z $\mathbf{G}$ na $\mathbf{H}$, jehož jádrem je $\mathbf{N}$. Metodu ilustrujeme na několika příkladech.

Příklad. Jak vypadá faktorgrupa $\mathbb{Z}/n\mathbb{Z}$? Analýzu situace jsme provedli výše a vidíme, že bychom měli hledat homomorfismus $\mathbb{Z} \to \mathbb{Z}_n$, jehož jádrem je podgrupa $n\mathbb{Z}$. Situaci řeší zobrazení $a \mapsto a \bmod n$, které je očividně homomorfismem na $\mathbb{Z}_n$, jehož jádrem je $\{a \in \mathbb{Z} : a \bmod n = 0\} = n\mathbb{Z}$. Podle 1. věty o izomorfismu

$$\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}_n.$$

Příklad. Jak vypadá faktorgrupa $\mathbf{S}_n/\mathbf{A}_n$? Analýzu situace jsme provedli výše a vidíme, že bychom měli hledat homomorfismus $\mathbf{S}_n \to \mathbb{Z}^*$, jehož jádrem je podgrupa A_n. Situaci řeší zobrazení $\pi \mapsto \mathrm{sgn}(\pi)$, které je očividně homomorfismem na $\mathbb{Z}^*$, jehož jádro tvoří sudé permutace. Podle 1. věty o izomorfismu

$$\mathbf{S}_n/\mathbf{A}_n \simeq \mathbb{Z}^*.$$

Příklad. Jak vypadá faktorgrupa $\mathbf{GL}_n(\mathbf{T})/\mathbf{SL}_n(\mathbf{T})$? Platí

$$A \sim B \;\Leftrightarrow\; AB^{-1} \in \mathrm{SL}_n(\mathbf{T}) \;\Leftrightarrow\; \det AB^{-1} = \det A(\det B)^{-1} = 1 \;\Leftrightarrow\; \det A = \det B.$$

Bloky této ekvivalence jsou tedy určeny hodnotou determinantu, kterou může být libovolný nenulový prvek tělesa. Přitom determinant součinu je součin determinantů, tedy bloky se násobí tak, jak se násobí příslušné prvky tělesa, čili faktorgrupa $\mathbf{GL}_n(\mathbf{T})/\mathbf{SL}_n(\mathbf{T})$ by měla být izomorfní grupě $\mathbf{T}^*$. Skutečně, zobrazení $\det : \mathbf{GL}_n(\mathbf{T}) \to \mathbf{T}^*$ je homomorfismem na grupu $\mathbf{T}^*$, jehož jádro tvoří matice s determinantem 1, čili podgrupa $\mathbf{SL}_n(\mathbf{T})$. Podle 1. věty o izomorfismu je

$$\mathbf{GL}_n(\mathbf{T})/\mathbf{SL}_n(\mathbf{T}) \simeq \mathbf{T}^*.$$

Jsou případy, kdy podobná analýza nedává žádný dobrý náhled. Leckdy je možné použít dalších triků, například úvah o počtu prvků a znalosti malých grup.

Příklad. Jak vypadá faktorgrupa $\mathbf{S}_4/\mathbf{K}_4$, kde $\mathbf{K}_4$ je Kleinova podgrupa? Podle Lagrangeovy věty je $|\mathbf{S}_4/\mathbf{K}_4| = [\mathbf{S}_4 : \mathbf{K}_4] = 24/4 = 6$, čili faktorgrupa $\mathbf{S}_4/\mathbf{K}_4$ je izomorfní buď grupě $\mathbf{S}_3$, nebo cyklické grupě $\mathbb{Z}_6$. Dokážeme, že není abelovská, což potvrdí první variantu:

$$[(1\ 2\ 3)] \circ [(1\ 2\ 3\ 4)] = [(1\ 2\ 3) \circ (1\ 2\ 3\ 4)] = [(1\ 3\ 4\ 2)],$$
$$[(1\ 2\ 3\ 4)] \circ [(1\ 2\ 3)] = [(1\ 2\ 3\ 4) \circ (1\ 2\ 3)] = [(1\ 3\ 2\ 4)],$$

ovšem $[(1\ 3\ 4\ 2)] \neq [(1\ 3\ 2\ 4)]$, neboť $(1\ 3\ 4\ 2) \circ (1\ 3\ 2\ 4)^{-1} = (1\ 2\ 4) \notin K_4$.

Jiným příkladem použití 1. věty o izomorfismu je elegantnější důkaz klasifikace cyklických grup.

Alternativní důkaz Věty 15.8. Buď $\mathbf{G} = \langle a \rangle$ cyklická grupa a uvažujme zobrazení

$$\varphi : \mathbb{Z} \to \mathbf{G}, \qquad k \mapsto a^k.$$

Podle Důsledku 14.3 je toto zobrazení na $\mathbf{G}$. Je-li φ také prosté, pak je izomorfismem $\mathbf{G} \simeq \mathbb{Z}$. V opačném případě je $\mathbf{Ker}(\varphi) = n\mathbb{Z}$, kde $n = \mathrm{ord}(a)$, a podle 1. věty o izomorfismu je $\mathbf{G} \simeq \mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}_n$. □

V sekci o řešitelných grupách bude potřeba porozumět tomu, jak vypadají podgrupy faktorgrup. O tom hovoří 2. věta o izomorfismu.

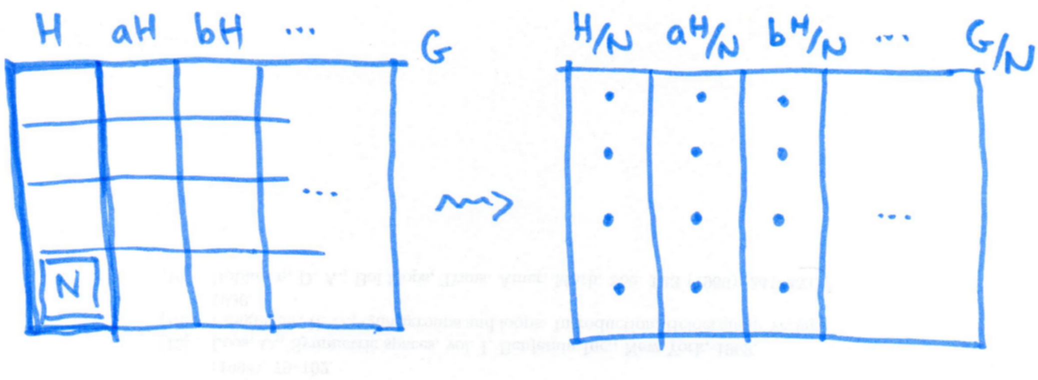
Tvrzení 19.5 (2. věta o izomorfismu)**.** *Buď* $\mathbf{G}$ *grupa a* $\mathbf{N}$ *její normální podgrupa.*

(1) *Je-li* $\mathbf{N} \trianglelefteq \mathbf{H} \trianglelefteq \mathbf{G}$*, pak* $\mathbf{H}/\mathbf{N}$ *je normální podgrupa v* $\mathbf{G}/\mathbf{N}$.
(2) *Je-li* $\mathbf{K} \trianglelefteq \mathbf{G}/\mathbf{N}$*, pak existuje normální podgrupa* $\mathbf{H} \trianglelefteq \mathbf{G}$ *taková, že* $\mathbf{K} = \mathbf{H}/\mathbf{N}$.
(3) *Pro* $\mathbf{N} \trianglelefteq \mathbf{H} \trianglelefteq \mathbf{G}$ *platí*

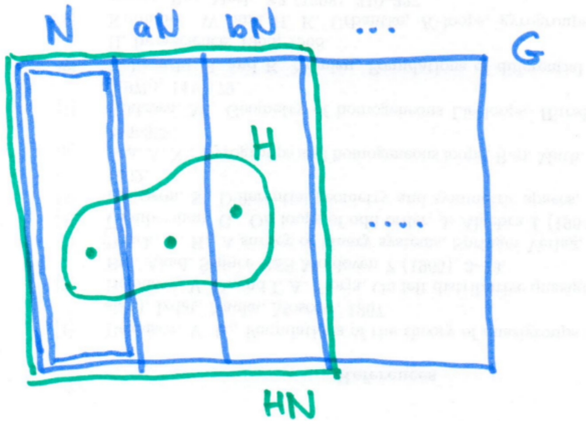
$$(\mathbf{G}/\mathbf{N})/(\mathbf{H}/\mathbf{N}) \simeq \mathbf{G}/\mathbf{H}.$$

Důkaz. (1) Buď $[a], [b]$ prvky $\mathbf{H}/\mathbf{N}$, čili $a, b \in H$, a buď $[g]$ prvek $\mathbf{G}/\mathbf{N}$. Pak $[a][b] = [ab]$ je prvek $\mathbf{H}/\mathbf{N}$, protože $ab \in H$, a ze stejného důvodu jsou $\mathbf{H}/\mathbf{N}$ také $[1]$, $[a]^{-1} = [a^{-1}]$ a $[g][a][g]^{-1} = [gag^{-1}]$.

(2) Buď $H = \{a \in G : [a] \in K\}$. Pro $a, b \in H$ a $g \in G$ platí $ab \in H$, protože $[ab] = [a][b] \in K$, a ze stejného důvodu jsou prvky $\mathbf{H}$ také 1, a^{-1} a gag^{-1}. Zjevně $\mathbf{K} = \mathbf{H}/\mathbf{N}$.



OBRÁZEK 20. Ilustrace 2. věty o izomorfismu. Větší podgrupa **H** určuje hrubší ekvivalenci (s většími bloky).



OBRÁZEK 21. Ilustrace 3. věty o izomorfismu. Podgrupa HN je sjednocení rozkladových tříd xN, které H protne.

(3) Uvažujme homomorfismus $\varphi : \mathbf{G}/\mathbf{N} \to \mathbf{G}/\mathbf{H}$, $[a]_\mathbf{N} \mapsto [a]_\mathbf{H}$. Je dobře definovaný, protože $\mathbf{N} \leq \mathbf{H}$, a tedy $[a]_\mathbf{N} = [b]_\mathbf{N}$ implikuje $[a]_\mathbf{H} = [b]_\mathbf{H}$. Je to homomorfismus, $\varphi([a]_\mathbf{N}[b]_\mathbf{N}) = \varphi([ab]_\mathbf{N}) = [ab]_\mathbf{H} = [a]_\mathbf{H}[b]_\mathbf{H} = \varphi([a]_\mathbf{N})\varphi([b]_\mathbf{N})$. Jeho obraz je celé $\mathbf{G}/\mathbf{H}$ a jeho jádro sestává z těch $[a]_\mathbf{N}$, pro které je $a \in H$, tedy $\mathbf{Ker}(\varphi) = \mathbf{H}/\mathbf{N}$. Aplikací 1. věty o izomorfismu dostaneme uvedený vztah. □

Je-li dána podgrupa **H** a normální podgrupa **N**, máme dvě přirozeně definované dvojice podgrup: jednou je $\mathbf{H} \cap \mathbf{N} \trianglelefteq \mathbf{H}$, druhou je $\mathbf{N} \trianglelefteq \mathbf{HN}$, kde $HN = \{hn : h \in H, n \in N\}$. Oba faktory jsou izomorfní.

Tvrzení 19.6 (3. věta o izomorfismu)**.** *Buď* **G** *grupa,* **N** *její normální podgrupa a* **H** *její libovolná podgrupa. Pak HN tvoří podgrupu grupy* **G***, $H \cap N$ tvoří normální podgrupu grupy* **H** *a*

$$\mathbf{HN}/\mathbf{N} \simeq \mathbf{H}/(\mathbf{H} \cap \mathbf{N}).$$

Důkaz si proveďte jako cvičení v podobném stylu, jako jsme dokazovali 2. větu o izomorfismu.

19.3. **Řešitelné grupy.**

Pojem řešitelnosti vychází z Galoisovy věty (Věta 25.1), která říká, že danou polynomiální rovnici lze řešit v radikálech právě tehdy, když je Galoisova grupa tohoto polynomu řešitelná. Pojmy týkající se řešení polynomiálních rovnic si vysvětlíme později, teď se podíváme na pojem grupové řešitelnosti.

Definice. Grupa $\mathbf{G}$ se nazývá *řešitelná*, pokud existuje číslo k a normální podgrupy $\mathbf{N}_0, \ldots, \mathbf{N}_k \trianglelefteq \mathbf{G}$ takové, že $\{1\} = \mathbf{N}_0 \leq \mathbf{N}_1 \leq \ldots \leq \mathbf{N}_k = \mathbf{G}$ a každá faktorgrupa $\mathbf{N}_i/\mathbf{N}_{i-1}$, $i = 1, \ldots, k$, je abelovská. Nejmenšímu k, pro které taková řada podgrup existuje, se říká *stupeň řešitelnosti* grupy $\mathbf{G}$.

Vidíme, že grupa je řešitelná stupně 1 právě tehdy, když je abelovská. Řešitelné grupy stupně ≤ 2 se nazývají *metabelovské*.

Příklady.

- Grupa $\mathbf{S}_3$ je metabelovská, jak prokazuje řada podgrup
$$\{1\} \leq \mathbf{A}_3 \leq \mathbf{S}_3.$$
Obě faktorgrupy $\mathbf{A}_3/\{1\} \simeq \mathbf{A}_3 \simeq \mathbb{Z}_3$ a $\mathbf{S}_3/\mathbf{A}_3 \simeq \mathbb{Z}_2$ jsou abelovské.
- Obecněji, dihedrální grupy $\mathbf{D}_{2n}$ jsou metabelovské, jak prokazuje řada podgrup
$$\{1\} \leq \mathbf{R} \leq \mathbf{D}_{2n},$$
kde $\mathbf{R}$ sestává ze všech rotací. Obě faktorgrupy $\mathbf{R}/\{1\} \simeq \mathbf{R} \simeq \mathbb{Z}_n$ a $\mathbf{D}_{2n}/\mathbf{R} \simeq \mathbb{Z}_2$ jsou abelovské.
- Grupa $\mathbf{S}_4$ je řešitelná stupně 3, jak prokazuje řada podgrup
$$\{1\} \leq \mathbf{K}_4 \leq \mathbf{A}_4 \leq \mathbf{S}_4,$$
kde $\mathbf{K}_4$ je Kleinova podgrupa. Všechny faktorgrupy $\mathbf{K}_4/\{1\} \simeq \mathbf{K}_4 \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbf{A}_4/\mathbf{K}_4 \simeq \mathbb{Z}_3$ a $\mathbf{S}_4/\mathbf{A}_4 \simeq \mathbb{Z}_2$ jsou abelovské.
- Všechny grupy řádu p^k, p prvočíslo, jsou řešitelné, ale důkaz není úplně jednoduchý. Slavná a velmi obtížná *Feit-Thompsonova věta* říká, že všechny grupy lichého řádu jsou řešeitelné.

Příklad. Grupy $\mathbf{S}_n$, $n \geq 5$, nejsou řešitelné, protože $\mathbf{S}_n$ má jedinou vlastní normální podgrupu $\mathbf{A}_n$, tedy jediná možná řada je $\{1\} \leq \mathbf{A}_n \leq \mathbf{S}_n$, avšak podgrupa $\mathbf{A}_n$ není abelovská. Ani grupy $\mathbf{A}_n$ nejsou řešitelné, protože nemají vůbec žádné vlastní normální podgrupy. (Pro $n = 5$ jste tento fakt měli ověřit ve cvičeních k sekci 19.1.) Grupa $\mathbf{A}_5$ řádu 30 je nejmenší grupa, která není řešitelná.

S abelovskostí faktorgrupy úzce souvisí následující pojem: pro danou grupu $\mathbf{G}$ definujeme *derivovanou podgrupu*
$$\mathbf{G}' = \langle aba^{-1}b^{-1} : a, b \in G.$$
Tato podgrupa je vždy normální v $\mathbf{G}$, jak si plyne z následujícího tvrzení.

Lemma 19.7. *Buď $\mathbf{G}$ grupa a $\mathbf{N}$ její normální podgrupa.*

(1) *$\mathbf{N}'$ je normální podgrupa v $\mathbf{G}$.*
(2) *$\mathbf{G}/\mathbf{N}$ je abelovská právě tehdy, když $\mathbf{G}' \leq \mathbf{N}$.*

Důkaz. (1) Generátory grupy $\mathbf{N}'$ jsou uzavřeny na konjugaci: pro $a, b \in N$ a $g \in G$ platí
$$g(aba^{-1}b^{-1})g^{-1} = (gag^{-1})(gbg^{-1})(ga^{-1}g^{-1})(gb^{-1}g^{-1}),$$
přičemž všechny čtyři prvky v závorkách jsou v $\mathbf{N}$, protože je tato podgrupa normální. Z Tvrzení 14.2 pak stejným argumentem plyne, že pokud je množina generátorů X uzavřena na konjugaci, podgrupa jimi generovaná je normální: pro prvek $a_1^{k_1} \ldots a_n^{k_n}$, kde $a_i \in X$, platí
$$g(a_1^{k_1} \ldots a_n^{k_n})g^{-1} = (ga_1g^{-1})^{k_1} \ldots (ga_ng^{-1})^{k_n},$$
přičemž všechny prvky v závorkách jsou v X.

(2) Faktorgrupa $\mathbf{G}/\mathbf{N}$ je abelovská právě tehdy, když pro všechna $a, b \in G$ platí $[a][b] = [b][a]$, neboli $[1] = [a][b][a]^{-1}[b]^{-1} = [aba^{-1}b^{-1}]$, tj. $aba^{-1}b^{-1} \in N$. Nejmenší taková podgrupa je z definice $\mathbf{G}'$, všechny ostatní $\mathbf{N}$ ji musí obsahovat. □

Následující tvrzení dává do souvislosti řešitelnost dané grupy a jejích podgrup a faktorgrup. Několikrát jej využijeme v důkazu Galoisovy věty.

Tvrzení 19.8 (řešitelnost podgrup a faktorgrup)**.** *Buď* $\mathbf{G}$ *grupa.*

(1) *Je-li* $\mathbf{G}$ *řešitelná a* $\mathbf{H}$ *její podgrupa, pak je* $\mathbf{H}$ *řešitelná.*
(2) *Je-li* $\mathbf{G}$ *řešitelná a* $\mathbf{K}$ *její normální podgrupa, pak je* $\mathbf{G}/\mathbf{K}$ *řešitelná.*
(3) *Pokud* $\mathbf{G}$ *obsahuje normální podgrupu* $\mathbf{N}$ *takovou, že jsou obě grupy* $\mathbf{N}$ *i* $\mathbf{G}/\mathbf{N}$ *řešitelné, pak je* $\mathbf{G}$ *řešitelná.*

Důkaz. (1) Uvažujme řadu normálních podgrup $\{1\} = \mathbf{N}_0 \leq \mathbf{N}_1 \leq \ldots \leq \mathbf{N}_k = \mathbf{G}$, kde je každá faktorgrupa $\mathbf{N}_i/\mathbf{N}_{i-1}$ abelovská. Dokážeme, že

$$\{1\} = \mathbf{N}_0 \cap \mathbf{H} \leq \mathbf{N}_1 \cap \mathbf{H} \leq \ldots \leq \mathbf{N}_k \cap \mathbf{H} = \mathbf{H}$$

je řada prokazující řešitelnost grupy $\mathbf{H}$. Pomocí 3. věty o izomorfismu upravíme

$$\begin{aligned}(\mathbf{N}_i \cap \mathbf{H})/(\mathbf{N}_{i-1} \cap \mathbf{H}) &= (\mathbf{N}_i \cap \mathbf{H})/((\mathbf{N}_i \cap \mathbf{H}) \cap \mathbf{N}_{i-1}) \\ &\simeq (\mathbf{N}_{i-1}(\mathbf{N}_i \cap \mathbf{H}))/\mathbf{N}_{i-1} \leq \mathbf{N}_i/\mathbf{N}_{i-1}.\end{aligned}$$

Vidíme, že původní faktorgrupa je podgrupou abelovské grupy $\mathbf{N}_i/\mathbf{N}_{i-1}$, čili je abelovská.

(2) Vyjdeme ze stejné řady a dokážeme, že

$$\mathbf{K}/\mathbf{K} = \mathbf{N}_0\mathbf{K}/\mathbf{K} \leq \mathbf{N}_1\mathbf{K}/\mathbf{K} \leq \ldots \leq \mathbf{N}_k\mathbf{K}/\mathbf{K} = \mathbf{G}/\mathbf{K}$$

je řada prokazující řešitelnost grupy $\mathbf{G}/\mathbf{K}$. Použijeme postupně 2., 3. a 2. větu o izomorfismu a dostaneme

$$\begin{aligned}(\mathbf{N}_i\mathbf{K}/\mathbf{N})/(\mathbf{N}_{i-1}\mathbf{K}/\mathbf{K}) &\simeq \mathbf{N}_i\mathbf{K}/\mathbf{N}_{i-1}\mathbf{K} = \mathbf{N}_i(\mathbf{N}_{i-1}\mathbf{K})/\mathbf{N}_{i-1}\mathbf{K} \\ &\simeq \mathbf{N}_i/(\mathbf{N}_{i-1}\mathbf{K}) \cap \mathbf{N}_i \simeq (\mathbf{N}_i/\mathbf{N}_{i-1})/\left((\mathbf{N}_{i-1}\mathbf{K}) \cap \mathbf{N}_i/\mathbf{N}_{i-1}\right)\end{aligned}$$

(poslední krok využívá pozorování, že $\mathbf{N}_{i-1} \leq (\mathbf{N}_{i-1}\mathbf{K}) \cap \mathbf{N}_i$). Vidíme, že původní faktorgrupa je faktorgrupou abelovské grupy $\mathbf{N}_i/\mathbf{N}_{i-1}$, čili je abelovská.

(3) Uvažujme řadu $\mathbf{N}/\mathbf{N} = \mathbf{L}_0/\mathbf{N} \leq \mathbf{L}_1/\mathbf{N} \leq \ldots \leq \mathbf{L}_l/\mathbf{N} = \mathbf{G}/\mathbf{N}$, kde je každá faktorgrupa $(\mathbf{L}_i/\mathbf{N})/(\mathbf{L}_{i-1}/\mathbf{N}) \simeq \mathbf{L}_i/\mathbf{L}_{i-1}$ abelovská (takový zápis podgrup je možný díky 2. větě o izomorfismu, bod (2)). Tvrzení dokážeme indukcí podle stupně řešitelnosti grupy $\mathbf{N}$.

Je-li $\mathbf{N}$ řešitelná stupně 1, tedy abelovská, pak

$$\{1\} \leq \mathbf{N} = \mathbf{L}_0 \leq \mathbf{L}_1 \leq \ldots \leq \mathbf{L}_l = \mathbf{G},$$

je řada prokazující řešitelnost grupy $\mathbf{G}$.

Nyní předpokládejme, že tvrzení platí, kdykoliv je $\mathbf{N}$ řešitelná stupně $< k$, a uvažujme situaci, kdy je stupeň řešitelnosti roven k. Uvažujme řadu $\{1\} = \mathbf{K}_0 \leq \mathbf{K}_1 \leq \ldots \leq \mathbf{K}_k = \mathbf{N}$, kde je každá faktorgrupa $\mathbf{K}_i/\mathbf{K}_{i-1}$ abelovská. Předposlední člen řady, grupa $\mathbf{K}_{k-1}$, je řešitelná stupně $< k$ a $\mathbf{N}/\mathbf{K}_{k-1}$ je abelovská. Z Lemmatu 19.7 plyne, že $\mathbf{N}' \leq \mathbf{K}_{k-1}$ a že $\mathbf{N}'$ je normální v $\mathbf{G}$. Část (1) pak říká, že grupa $\mathbf{N}'$ je řešitelná stupně $< k$, díky řadě

$$\{1\} = \mathbf{K}_0 \cap \mathbf{N}' \leq \mathbf{K}_1 \cap \mathbf{N}' \leq \ldots \leq \mathbf{K}_{k-1} \cap \mathbf{N}' = \mathbf{N}'.$$

Také $\mathbf{G}/\mathbf{N}'$ je řešitelná grupa, což prokazuje řada

$$\mathbf{N}'/\mathbf{N}' \leq \mathbf{N}/\mathbf{N}' = \mathbf{L}_0/\mathbf{N}' \leq \mathbf{L}_1/\mathbf{N}' \leq \ldots \leq \mathbf{L}_l/\mathbf{N}' = \mathbf{G}/\mathbf{N}'.$$

Z indukčního předpokladu plyne, že $\mathbf{G}$ je řešitelná. □

Důsledek 19.9. *Buď* $\mathbf{G}$ *grupa a* $\mathbf{N}_0, \ldots, \mathbf{N}_k \trianglelefteq \mathbf{G}$ *takové, že* $\{1\} = \mathbf{N}_0 \leq \mathbf{N}_1 \leq \ldots \leq \mathbf{N}_k = \mathbf{G}$ *a každá faktorgrupa* $\mathbf{N}_i/\mathbf{N}_{i-1}$, $i = 1, \ldots, k$, *je řešitelná. Pak je* $\mathbf{G}$ *řešitelná.*

Důkaz. Indukcí podle k: z indukčního předpokladu je řešitelná grupa $\mathbf{N}_{k-1}$, řešitelná je také faktorgrupa $\mathbf{N}_k/\mathbf{N}_{k-1}$, takže díky Tvrzení 19.8(3) je řešitelná i grupa $G = \mathbf{N}_k$. □

Číselná tělesa a kořeny polynomů

20. Okruhové homomorfismy a faktorokruhy

20.1. Homomorfismy.
Výklad teorie tělesových rozšíření začneme krátkým pojednáním o základních vlastnostech okruhových homomorfismů a s nimi související obecnou konstrukcí faktorokruhu podle ideálu. Většina faktů v této sekci je přímou analogií situace, kterou jsme viděli v grupách.

V celé sekci budou $\mathbf{R}$, $\mathbf{S}$ dva okruhy (ne nutně komutativní). Zobrazení $\varphi : R \to S$ je *homomorfismem* těchto okruhů, zapisujeme $\varphi : \mathbf{R} \to \mathbf{S}$, pokud pro každé $a, b \in R$ platí

$$\varphi(a+b) = \varphi(a) + \varphi(b), \quad \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b), \quad \varphi(-a) = -\varphi(a) \quad \text{a} \quad \varphi(0) = 0.$$

Z Lemmatu 15.1 ihned plyne, že první dvě podmínky implikují třetí a čtvrtou, čili jsme v souladu s definicí ze sekce 1.3.

Obrazem homomorfismu nazýváme jeho obor hodnot, tj. množinu

$$\mathrm{Im}(\varphi) = \{b \in H : \ b = \varphi(a) \text{ pro nějaké } a \in G\}.$$

Jádro homomorfismu definujeme jako množinu

$$\mathrm{Ker}(\varphi) = \{a \in G : \ \varphi(a) = 0\}.$$

Připomeňme, že podmnožinu I v okruhu $\mathbf{R}$ nazveme ideálem, pokud je $(I, +, -, 0)$ podgrupou grupy $(R, +, -, 0)$ a pro každé $a \in I$, $r \in R$ platí $r \cdot a \in I$ a $a \cdot r \in I$. Následující tvrzení mimo jiné říká, že ideály hrají vzhledem k homomorfismům roli normálních podgrup.

Tvrzení 20.1 (jádro a obraz jsou podokruhy)**.** *Buď* $\mathbf{R}, \mathbf{S}$ *okruhy a* $\varphi : \mathbf{R} \to \mathbf{S}$ *homomorfismus. Pak*

(1) $\mathrm{Im}(\varphi)$ *tvoří podokruh okruhu* $\mathbf{S}$*;*
(2) $\mathrm{Ker}(\varphi)$ *tvoří ideál okruhu* $\mathbf{R}$

Důkaz. Okruhový homomorfismus je zároveň grupový homomorfismus vzhledem k operacím $+, -, 0$, čili můžeme použít Tvrzení 15.2 a ihned dostaneme uzavřenost jádra a obrazu na operace $+, -, 0$. Uzavřenost obrazu na násobení se dokáže stejně jako pro sčítání. Zbývá dokončit část (1): je-li $\varphi(a) = 0$ a $r \in R$ libovolné, pak $\varphi(r \cdot a) = \varphi(r) \cdot \varphi(a) = \varphi(r) \cdot 0 = 0$ a analogicky pro součin $a \cdot r$, čili $\mathrm{Ker}(\varphi)$ tvoří ideál v $\mathbf{R}$. □

Z Tvrzení 15.3 ihned plyne jeho analogie pro okruhy.

Tvrzení 20.2. *Buď* $\mathbf{R}, \mathbf{S}$ *okruhy a* $\varphi : \mathbf{R} \to \mathbf{S}$ *homomorfismus. Pak* φ *je prostý právě tehdy, když* $\mathrm{Ker}(\varphi) = \{0\}$*.*

Podobně jako pro grupy se dokáže také následující tvrzení (proveďte jako cvičení!).

Tvrzení 20.3. *Buď* $\mathbf{R}, \mathbf{S}, \mathbf{T}$ *okruhy a* $\varphi : \mathbf{R} \to \mathbf{S}$, $\psi : \mathbf{S} \to \mathbf{T}$ *homomorfismy. Pak*

(1) $\psi \circ \varphi$ *je homomorfismus* $\mathbf{R} \to \mathbf{T}$,
(2) *je-li* φ *bijektivní, pak* φ^{-1} *je homomorfismus* $\mathbf{S} \to \mathbf{T}$.

Bijektivní homomorfismy se nazývají *izomorfismy.* Dva okruhy nazveme *izomorfní*, pokud mezi nimi vede izomorfismus (viz sekce 1.3). Vše, co bylo v sekci 15.2 řečeno o izomorfismech grup, platí analogicky i o izomorfismech okruhů.

Příklad. Důležitou rodinou jsou tzv. *modulární homomorfismy.* V číselné variantě zvolme číslo $m > 0$ a definujme zobrazení

$$\varphi_m : \mathbb{Z} \to \mathbb{Z}_m, \qquad a \mapsto a \bmod m.$$

V polynomiální variantě zvolme polynom $0 \neq m \in T[x]$ a definujme zobrazení

$$\varphi_m : \mathbf{T}[x] \to \mathbf{T}[x]/(m), \qquad f \mapsto f \bmod m.$$

Není těžké ověřit, že jde o homomorfismy, jejich jádra jsou $m\mathbb{Z}$, resp. $m\mathbf{T}[x]$. Podobně bychom mohli postupovat v každém oboru, kde je definováno jednoznačné dělení se zbytkem.

Modulární zobrazení z čínské věty o zbytcích je okruhový izomorfismus, viz Tvrzení 15.6. Toto zobrazení lze interpretovat jako simultánní modulární homomorfismus do několika okruhů $\mathbb{Z}_m$ najednou.

Příklad. Důležitou rodinou jsou tzv. *dosazovací homomorfismy.* Uvažujme komutativní okruhy $\mathbf{R} \leq \mathbf{S}$ a prvek $a \in S$ a definujme zobrazení

$$\varphi_a : \mathbf{R}[x] \to \mathbf{S}, \qquad f \mapsto f(a).$$

Není těžké ověřit, že jde o homomorfismus. Je-li $\mathbf{R} = \mathbf{S}$, jeho jádrem je hlavní ideál $(x-a)T[x]$ (viz Tvrzení 2.3) a obrazem celé $\mathbf{R}$ (díky konstantním polynomům). Obecně může jádro a obraz vycházet všelijak, např. pro $\mathbf{R} = \mathbb{Z}$, $\mathbf{S} = \mathbb{C}$, $a = i$ dostaneme jádro $(x^2+1)\mathbb{Z}[x]$ a obraz $\mathbb{Z}[i]$.

Příklad. Oba výše uvedené typy lze kombinovat, například zobrazení

$$\varphi : \mathbb{Z}[x] \to \mathbb{Z}_2, \qquad f \mapsto f(0) \bmod 2$$

je také homomorfismem, jeho jádro sestává z těch polynomů, jejichž absolutní člen je sudý, což není hlavní ideál.

Tuto podsekci ukončíme příkladem homomorfismu, který hraje zásadní roli v okruzích nenulové charakteristiky. Využijeme jej například při klasifikaci konečných těles.

Tvrzení 20.4 (Frobeniův endomorfismus)**.** *Buď* $\mathbf{R}$ *okruh s jednotkou prvočíselné charakteristiky* p *a definujme zobrazení*

$$\varphi_p : \mathbf{R} \to \mathbf{R}, \qquad a \mapsto a^p.$$

(1) *Zobrazení* φ_p *je homomorfismus.*
(2) *Je-li* $\mathbf{R}$ *obor integrity, pak je* φ_p *prosté.*
(3) *Je-li* $\mathbf{R}$ *konečné těleso, pak je* φ_p *automorfismus.*

Zobrazení φ_p se říká *Frobeniův endomorfismus*, resp. *Frobeniův automorfismus* v případě, kdy je bijektivní. Speciálním případem bodu (1) je rovnost

$$(a+b)^p = a^p + b^p,$$

která je snem každého středoškoláka, ale platí pouze za předpokladu prvočíselné charakteristiky p.

Důkaz. (1) Zřejmě $(a \cdot b)^p = a^p \cdot b^p$ a podle binomické věty

$$(a+b)^p = \sum_{i=0}^{p} \binom{p}{i} a^i b^{p-i} = a^p + b^p,$$

neboť p dělí všechny binomické koeficienty $\binom{p}{i}$, $i = 1, \dots, p-1$, protože všechna prvočísla obsažená ve jmenovateli jsou menší.

(2) Z podmínky $\varphi_p(a) = a^p = 0$ plyne, že $a = 0$. Tedy jádro homomorfismu φ_p je triviální, čili je prostý (Tvrzení 20.2).

(3) Prosté zobrazení na konečné množině je bijektivní (Lemma 1.4). $\square$

Důsledek 20.5. *Buď* p *prvočíslo a* f *polynom ze* $\mathbb{Z}_p[x]$. *Pak* $f(x^p) = f^p$.

Důkaz. Označme $f = \sum a_i x^i$. Pak $f^p = (\sum a_i x^i)^p = \sum a_i^p (x^i)^p = \sum a_i x^{ip}$, kde druhá rovnost plyne z Tvrzení 20.4 a třetí z malé Fermatovy věty. $\square$

20.2. **Konstrukce faktorokruhu podle ideálu.**

Vzhledem k tomu, že ideál tvoří normální podgrupu aditivní grupy daného okruhu, můžeme v konstrukci faktorokruhu vyžít vše, co jsme udělali pro faktorgrupy.

Definice. Buď $\mathbf{R}$ okruh a I jeho ideál. Definujeme ekvivalenci na množině R předpisem

$$a \sim b \;\Leftrightarrow\; a - b \in I.$$

Podle Tvrzení 14.10 je $a \sim b$ právě tehdy, když $a + I = b + I$, čili bloky jsou rozkladové třídy $[a] = a + I$. Na těchto blocích definujeme operace předpisy

$$[a] + [b] = [a + b], \qquad -[a] = [-a], \qquad [a] \cdot [b] = [a \cdot b].$$

(v následujícím lemmatu ověříme, že je tato definice korektní). Množina bloků s výše uvedenými operacemi se nazývá *faktorokruh okruhu* $\mathbf{R}$ *podle ideálu* I,

$$\mathbf{R}/I = (\{[a] : a \in R\}, +, -, \cdot, [0]).$$

Lemma 20.6. *Buď* $\mathbf{R}$ *okruh a* I *jeho ideál.*

(1) *Výše uvedené operace na blocích jsou dobře definovány.*
(2) *Faktorokruh* $\mathbf{R}/I$ *je skutečně okruh.*

Důkaz. Z konstrukce faktorgrupy již víme, že je sčítání a odčítání dobře definované. Pro násobení předpokládejme $[a] = [c]$ a $[b] = [d]$, ověříme, že $[ab] = [cd]$. Předpokládaje $a \sim c$ a $b \sim d$, tj. $a - c \in I$ a $b - d \in I$, spočteme

$$ab - cd = a \underbrace{(b - d)}_{\in I} + \underbrace{(a - c)}_{\in I} d \in I,$$

čili $ab \sim cd$, tj. $[ab] = [cd]$. Podobně jako pro grupy se ukáže, že takto definované operace splňují všechny axiomy okruhů, včetně komutativity a existence jednotky, pokud tyto platily v původním okruhu $\mathbf{R}$ (pozor, vlastnost býti oborem integrity zachována být nemusí, viz sekce 20.3). $\square$

Příklad. Uvažujme komutativní okruh $\mathbf{R} = \mathbb{Z}$ a ideál $I = n\mathbb{Z}$. Platí

$$a \sim b \;\Leftrightarrow\; n \mid a - b \;\Leftrightarrow\; a \equiv b \pmod{n}.$$

Stejně jako pro grupy není problém ukázat, že $\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}_n$.

Podobně jako pro grupy platí *věta o homomorfismu* a *1. věta o izomorfismu.*

Věta 20.7 (věta o homomorfismu). *Buď* $\varphi : \mathbf{R} \to \mathbf{S}$ *homomorfismus okruhů.*

(1) *Je-li* $I \subseteq \mathrm{Ker}(\varphi)$ *ideál v* $\mathbf{R}$, *pak je zobrazení*

$$\psi : \mathbf{R}/I \to \mathbf{S}, \qquad [a] \mapsto \varphi(a)$$

dobře definované a je to okruhový homomorfismus.

(2) [1. věta o izomorfismu] $\mathbf{R}/\mathrm{Ker}(\varphi) \simeq \mathbf{Im}(\varphi)$.

Důkaz. Plyne přímo z grupové verze (Věta 19.4), pouze je třeba si uvědomit, že všechna definovaná zobrazení jsou i okruhové homomorfismy. $\square$

Příklad. Modulární homomorfismus $\mathbb{Z} \to \mathbb{Z}_n$, $a \mapsto a \bmod n$, prokazuje, že $\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}_n$.

Je-li $\mathbf{R}$ komutativní okruh a $I = mR$ jeho hlavní ideál, pak

$$a \sim b \;\Leftrightarrow\; m \mid a - b \;\Leftrightarrow\; a \equiv b \pmod{m}.$$

Je-li v okruhu $\mathbf{R}$ definováno jednoznačné dělení se zbytkem (např. $\mathbf{R} = \mathbb{Z}$ nebo $\mathbf{R} = \mathbf{T}[x]$, $\mathbf{T}$ těleso), prvky $\mathbf{R}/mR$ můžeme reprezentovat jako všechny možné zbytky po dělení prvkem m a operace v $\mathbf{R}/mR$ budou jako operace v původním okruhu modulo m, neboť

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

Speciálně, pro $\mathbf{R} = \mathbf{T}[x]$, $\mathbf{T}$ těleso, vidíme, že konstrukce faktorokruhu ve smyslu této sekce a ve smyslu sekce 9.2 jsou „v podstatě totožné". 1. věta o izomorfismu použitá na modulární zobrazení $\mathbf{T}[x] \to \mathbf{T}[x]/(m)$ dává izomorfismus

$$\mathbf{T}[x]/mT[x] \simeq \mathbf{T}[x]/(m),$$

v němž polynom f stupně $< \deg m$ jednoznačně odpovídá příslušnému bloku $[f]$. Značení se zpravidla směšuje, používají se oba zápisy $\mathbf{T}[x]/mT[x]$ i $\mathbf{T}[x]/(m)$ pro obě formálně různé, nicméně izomorfní konstrukce.

Příklad. Jak vypadá faktorokruh $\mathbf{T}[x]/(x-a)$, kde $a \in T$? Uvažujme dosazovací homomorfismus

$$\mathbf{T}[x] \to \mathbf{T}, \qquad f \mapsto f(a).$$

Je to zobrazení na $\mathbf{T}$ a jeho jádro je

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

Podle 1. věty o izomorfismu je $\mathbf{T}[x]/(x-a) \simeq \mathbf{T}$.

Pro polynomy vyššího stupně je situace složitější.

Příklad. Jak vypadá faktorokruh $\mathbb{Q}[x]/(x^2+1)$? Uvažujme homomorfismus

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

Je to zobrazení na $\mathbb{Q}(i)$ a jeho jádro je

$$\begin{aligned}\{f \in \mathbb{Q}[x] : \ f(i) = 0\} &= \{f \in \mathbb{Q}[x] : \ f(i) = f(-i) = 0\}\\ &= \{f \in \mathbb{Q}[x] : \ x - i \mid f, \ x + i \mid f\}\\ &= \{f \in \mathbb{Q}[x] : \ (x-i)(x+i) = x^2+1 \mid f\}\\ &= (x^2+1)\mathbb{Q}[x].\end{aligned}$$

Podle 1. věty o izomorfismu je $\mathbb{Q}[x]/(x^2+1) \simeq \mathbb{Q}(i)$.

Příklad. Jak vypadá faktorokruh $\mathbb{Q}[x]/(x^2-1)$? Uvažujme homomorfismus

$$\mathbb{Q}[x] \to \mathbb{Q} \times \mathbb{Q}, \qquad f \mapsto (f(1), f(-1)).$$

Je to zobrazení na $\mathbb{Q} \times \mathbb{Q}$ a jeho jádro je

$$\begin{aligned}\{f \in \mathbb{Q}[x] : \ f(1) = f(-1) = 0\} &= \{f \in \mathbb{Q}[x] : \ x - 1 \mid f, \ x + 1 \mid f\}\\ &= \{f \in \mathbb{Q}[x] : \ (x-1)(x+1) = x^2-1 \mid f\}\\ &= (x^2-1)\mathbb{Q}[x].\end{aligned}$$

Podle 1. věty o izomorfismu je $\mathbb{Q}[x]/(x^2-1) \simeq \mathbb{Q} \times \mathbb{Q}$.

Na závěr jeden příklad s ideálem, který není hlavní.

Příklad. Jak vypadá faktorokruh $\mathbb{Z}[x]/I$, kde $I = \{f \in \mathbb{Z}[x] : \ m \mid f(0)\}$? Dva polynomy f, g jsou ekvivalentní právě tehdy, když $f - g \in I$, tj. právě tehdy, když $m \mid f(0) - g(0)$, tj. právě tehdy, když $f(0) \equiv g(0) \pmod m$. Existuje tedy přesně m rozkladových tříd. Není těžké nahlédnout, že zobrazení

$$\mathbb{Z}[x] \to \mathbb{Z}_m, \qquad f \mapsto f(0) \bmod m$$

je homomorfismem, jehož jádro je I, a tedy $\mathbb{Z}[x]/I \simeq \mathbb{Z}_m$.

Podobně jako pro grupy se dokáží i 2. a 3. věta o izomorfismu.

Tvrzení 20.8 (2. věta o izomorfismu)**.** *Buď* $\mathbf{R}$ *okruh a* I *jeho ideál.*

(1) *Je-li* $I \subseteq J$ *ideál v* $\mathbf{R}$*, pak* $J/I = \{[a] : a \in J\}$ *je ideál v* $\mathbf{R}/I$*.*
(2) *Je-li* K *ideál v* $\mathbf{R}/I$*, pak existuje ideál* J *v* $\mathbf{R}$ *takový, že* $K = J/I$*.*
(3) *V obou případech platí*

$$(\mathbf{R}/I)/(J/I) \simeq \mathbf{R}/J.$$

Tvrzení 20.9 (3. věta o izomorfismu)**.** *Buď* $\mathbf{R}$ *okruh,* I *jeho ideál a* $\mathbf{S}$ *podokruh. Pak* $S + I$ *tvoří podokruh okruhu* $\mathbf{R}$ *a*

$$(\mathbf{S} + \mathbf{I})/I \simeq \mathbf{S}/(S \cap I).$$

20.3. **Faktorokruhy podle maximálních ideálů a prvoideálů.**

V této části si ukážeme analogii k Tvrzení 9.4, které říká, že faktorokruh $\mathbf{T}[\alpha]/(m)$ je těleso právě tehdy, když to je obor integrity, a to právě tehdy, když m je ireducibilní prvek v $\mathbf{T}[\alpha]$. Obecně je situace složitější: může se stát, že faktorokruh podle ideálu je oborem integrity, ale ne tělesem.

Definice. Ideál I okruhu $\mathbf{R}$ nazveme

- *prvoideálem*, pokud pro každé $a, b \in R$ platí, že kdykoliv $ab \in I$, pak $a \in I$ nebo $b \in I$;
- *maximální*, pokud je I maximální v uspořádané množině vlastních ideálů, tj. pokud neexistuje ideál J splňující $I \subset J \subset R$.

Věta 20.10 (faktor podle prvoideálu a maximálního ideálu)**.** *Buď* $\mathbf{R}$ *komutativní okruh s jednotkou a* I *jeho ideál. Pak*

(1) $\mathbf{R}/I$ *je obor integrity právě tehdy, když* I *je prvoideál;*
(2) $\mathbf{R}/I$ *je těleso právě tehdy, když* I *je maximální ideál.*

Důkaz. (1) Faktorokruh $\mathbf{R}/I$ je oborem integrity právě tehdy, když pro každé $a, b \in R$ platí, že kdykoliv $[ab] = [a] \cdot [b] = [0]$, pak $[a] = [0]$ nebo $[b] = 0$. Přeloženo do řeči ideálů, $ab \in I$ implikuje $a \in I$ nebo $b \in I$, což je definice prvoideálu.

(2) Podle Tvrzení 7.6 je $\mathbf{R}/I$ tělesem právě tehdy, když neobsahuje žádné vlastní ideály. 2. věta o izomorfismu říká, že jakýkoliv vlastní ideál v $\mathbf{R}/I$ je tvaru J/I, kde $I \subset J \subset R$ je ideál v $\mathbf{R}$. Čili neexistence vlastního ideálu v $\mathbf{R}/I$ je ekvivalentní tomu, že je I maximální ideál. $\square$

Všimněte si, že v oborech hlavních ideálů jsou

- prvoideály právě ty ideály, které jsou generované prvočinitelem (obě podmínky požadují, aby $m \mid ab$ implikovalo $m \mid a$ nebo $m \mid b$);
- maximální právě ty ideály, které jsou generované ireducibilním prvkem (protože b je vlastním dělitelem a právě tehdy, když $aR \subset bR$).

Obory hlavních ideálů jsou gaussovské (Věta 7.8) a v nich jsou ireducibilní prvky totéž, co prvočinitelé (Důsledek 6.2(2)). Čili Tvrzení 9.4 je speciálním případem Věty 20.10.

Příklad. Připomeňte si příklady $\mathbb{Q}[x]/(f)$ ze sekce 20.2:

- polynom x^2+1 je ireducibilní, tedy ideál $(x^2+1)\mathbb{Q}[x]$ je maximální, a skutečně, $\mathbb{Q}[x]/(x^2+1) \simeq \mathbb{Q}(i)$ těleso;
- polynom $x^2 - 1$ není ireducibilní, tedy ideál $(x^2 - 1)\mathbb{Q}[x]$ není maximální (např. ideál $(x - 1)\mathbb{Q}[x]$ je větší), a skutečně, $\mathbb{Q}[x]/(x^2 - 1) \simeq \mathbb{Q} \times \mathbb{Q}$ není těleso.

Příklad. Obor $\mathbb{Z}[x]$ není oborem hlavních ideálů. Například ideál $I = x\mathbb{Z}[x]$ je prvoideálem (protože je polynom x prvočinitelem), avšak není to maximální ideál (například ideál $\{f \in \mathbb{Z}[x] : 2 \mid f(0)\}$ je větší). A vskutku, faktorokruh $\mathbb{Z}[x]/I \simeq \mathbb{Z}$ (dosazovací homomorfismus) je oborem integrity, ale ne tělesem.

21. Tělesové rozšíření jako vektorový prostor

Tématem celé kapitoly je studium vlastností tělesových rozšíření. Soustředíme se na dva základní pojmy: *stupeň rozšíření*, tj. dimenze většího tělesa jakožto vektorového prostoru nad menším tělesem, a jeho souvislost s vlastnostmi algebraických čísel, a dále na *Galoisovu grupu rozšíření*, tj. grupu symetrií většího tělesa vůči tomu menšímu.

Pomocí těchto pojmů lze studovat řadu klasických problémů. Ukážeme si dva příklady, které patří mezi stěžejní výsledky matematiky první poloviny 19. století. Pomocí stupně rozšíření se charakterizují body, které lze sestrojit pravítkem a kružítkem, a touto metodou dokážeme nemožnost řešení některých planimetrických úloh. Vyjádřitelnost kořenů daného polynomu vzorcem používajícím jeho koeficienty úzce souvisí se symetriemi jeho rozkladového nadtělesa a touto metodou dokážeme neexistenci vzorců na řešení polynomiálních rovnic stupně 5 a více. Jako vedlejší produkt teorie rozkladových nadtěles dostaneme také klasifikaci konečných těles.

Ve výše uvedených tématech nás budou zajímat především tzv. *číselná tělesa*, tj. reálná či komplexní rozšíření racionálních čísel konečného stupně, čili tělesa tvaru $\mathbb{Q}(a_1,\dots,a_n)$, kde $a_1,\dots,a_n$ jsou algebraická komplexní čísla. Většina teorie nicméně platí obecně. Připomeňme si zde konstrukce těles, které jsme dosud potkali:

- podílová tělesa oborů integrity, např. těleso racionálních funkcí (sekce 1.4),
- faktorokuhy podle ireducibilního prvku (Tvrzení 9.4), resp. podle maximálního ideálu (Věta 20.10),
- speciálně, konečná tělesa $\mathbb{Z}_p$ a $\mathbb{Z}_p[\alpha]/(f)$ (sekce 9.2).

Rozšířením těles budeme rozumět libovolnou dvojici těles $\mathbf{T},\mathbf{S}$ takovou, že $\mathbf{T}\leq\mathbf{S}$. Říkáme, že $\mathbf{T}$ je podtělesem $\mathbf{S}$, nebo že $\mathbf{S}$ je rozšířením $\mathbf{T}$.

Klíčem k pochopení celé kapitoly je myšlenka, že těleso $\mathbf{S}$ lze považovat za vektorový prostor nad tělesem $\mathbf{T}$: sčítání a odčítání ponecháme a místo násobení jakožto operace $S\times S\to S$ uvažujeme pouze restrikci $T\times S\to S$. Neformálně, prvky většího tělesa $\mathbf{S}$ považujeme za vektory, prvky menšího tělesa $\mathbf{T}$ za skaláry a uvažujeme pouze násobení skalár krát vektor. Tento vektorový prostor budeme značit $\mathbf{S_T}$.

Uvědomte si, že jde skutečně o vektorový prostor: aditivní struktura $(S,+,-,0)$ je abelovskou grupu a pro všechna $a,b\in T$ (skaláry), $v,w\in S$ (vektory) platí každý z axiomů vektorových prostorů: $a(bv)=(ab)v$ plyne z asociativity násobení, $1v=v$ z vlastnosti jednotky a $a(v+w)=av+aw$ a $(a+b)v=av+bv$ z distributivity.

Definice. Dimenze vektorového prostoru $\mathbf{S_T}$ se nazývá *stupeň rozšíření* a značí se

$$[\mathbf{S}:\mathbf{T}]=\dim\mathbf{S_T}.$$

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

Příklady.

- $[\mathbb{C}:\mathbb{R}]=2$. Každé komplexní číslo lze zapsat právě jedním způsobem jako $a+bi$, $a,b\in\mathbb{R}$, čili prvky $1,i$ tvoří bázi prostoru $\mathbb{C}_\mathbb{R}$.
- Analogicky, pro s, které není čtvercem, je stupeň $[\mathbb{Q}(\sqrt{s}):\mathbb{Q}]=2$, prvky $1,\sqrt{s}$ tvoří bázi prostoru $\mathbb{Q}(\sqrt{s})_\mathbb{Q}$.
- $[\mathbb{Q}(\sqrt{2},\sqrt{3}):\mathbb{Q}]=4$, bázi prostoru $\mathbb{Q}(\sqrt{2},\sqrt{3})_\mathbb{Q}$ tvoří například prvky $1,\sqrt{2},\sqrt{3},\sqrt{6}$.
- Pozor, pro $\zeta_3=e^{2\pi i/3}$ je stupeň $[\mathbb{Q}(\zeta_3):\mathbb{Q}]=2$ a nikoliv 3: prvky $1,\zeta_3,\zeta_3^2$ jsou lineárně závislé, protože $\zeta_3^2=-1-\zeta_3$.
- Je-li u transcendentní číslo (např. konstanty e nebo π), stupeň $[\mathbb{Q}(u):\mathbb{Q}]$ je nekonečný (spočetný): lineárně nezávislou množinu tvoří třeba prvky $1,u,u^2,\dots$ (viz Věta 22.4).
- Stupeň $[\mathbb{R}:\mathbb{Q}]$ je dokonce nespočetný: prostory spočetné dimenze nad spočetným tělesem jsou spočetné, zatímco reálných čísel je nespočetně.

Připomeňme pojem prvookruhu. Pro libovolný okruh $\mathbf{R}$ s jednotkou uvažujme zobrazení

$$\mathbb{Z}\to\mathbf{R},\qquad n\mapsto\underbrace{1+\ldots+1}_{n}.$$

Je vidět, že jde o homomorfismus, jehož obrazem je tzv. *prvookruh* okruhu $\mathbf{R}$ a jehož jádrem je ideál $n\mathbb{Z}$, kde n je *charakteristika* okruhu $\mathbf{R}$. Použitím 1. věty o izomorfismu dostáváme, že prvookruh libovolného okruhu je izomorfní buď okruhu $\mathbb{Z}$ v případě charakteristiky 0, nebo okruhu $\mathbb{Z}/n\mathbb{Z}\simeq\mathbb{Z}_n$ v případě charakteristiky n.

Nyní uvažujme těleso $\mathbf{T}$. Jeho *prvotělesem* se rozumí nejmenší podtěleso. To v sobě jistě obsahuje prvookruh, ale navíc musí ke každému nenulovému prvku obsahovat jeho inverz. Podle Tvrzení 1.6 je charakteristika tělesa 0 nebo prvočíslo p. V druhém případě je prvookruh již tělesem (izomorfním $\mathbb{Z}_p$), čili pojmy splývají. V případě charakteristiky 0 prvotěleso sestává ze všech zlomků ab^{-1}, kde a,b jsou prvky prvookruhu, čili je izomorfní tělesu $\mathbb{Q}$.

Každé těleso je samozřejmě rozšířením svého prvotělesa. Speciálně, pro konečná tělesa dostáváme velmi zajímavý důsledek vektorového pohledu na tělesová rozšíření.

Tvrzení 21.1. *Počet prvků konečného tělesa je mocnina prvočísla.*

Důkaz. Konečné těleso $\mathbf{T}$ charakteristiky p je rozšířením svého prvotělesa $\mathbf{P} \simeq \mathbb{Z}_p$. Čili vektorový prostor $\mathbf{T_P}$ je izomorfní prostoru $(\mathbb{Z}_p)^k$, kde $k = [\mathbf{T} : \mathbf{P}]$, čili má p^k prvků. □

22. Algebraické prvky a rozšíření konečného stupně

22.1. Minimální polynom a stupeň jednoduchého rozšíření.

V této sekci se vrátíme k pojmu algebraického čísla a uvedeme jej do souvislosti s vlastnostmi tzv. *jednoduchých rozšíření*, tj. rozšíření tvaru $\mathbf{T}(a)$, určených jedním prvkem. Hlavním cílem této sekce je vybudovat teorii, jejímž důsledkem je následující charakterizace: číslo a je algebraické právě tehdy, když je stupeň $[\mathbb{Q}(a) : \mathbb{Q}]$ konečný, přičemž tento stupeň je pak roven stupni libovolného ireducibilního polynomu, jehož je a kořenem.

Začneme obecnější definicí algebraičnosti, nad libovolným tělesem.

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 $\mathbf{T}$.

Uvědomte si, že číslo je algebraické ve smyslu sekce 2.6 právě tehdy, když je algebraickým prvkem nad tělesem $\mathbb{Q}$: dané číslo je kořenem nějakého racionálního polynomu právě tehdy, když je kořenem nějakého celočíselného polynomu, stačí přenásobit koeficienty.

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

Tvrzení 22.1 (vlastnosti minimálních polynomů). *Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles a $a \in S$ algebraický nad $\mathbf{T}$. Pak*

1. *minimální polynom $m_{a,\mathbf{T}}$ existuje a je jednoznačně určený;*
2. *prvek a je kořenem polynomu $f \in T[x]$ právě tehdy, když $m_{a,\mathbf{T}} \mid f$.*

Důkaz. Množina $I = \{f \in T[x] : f(a) = 0\}$ tvoří ideál v oboru $\mathbf{T}[x]$, a protože je $\mathbf{T}[x]$ oborem hlavních ideálů (Věta 7.5), existuje monický polynom $m \in T[x]$ takový, že $I = mT[x]$. Vidíme, že $f(a) = 0$ právě tehdy, když $m \mid f(a)$. Kdyby polynom m nebyl ireducibilní v $\mathbf{T}[x]$, tj. kdyby $m = fg$, kde $f, g \nparallel m$, pak $0 = m(a) = f(a)g(a)$, čili prvek a by byl kořenem alespoň jednoho z polynomů f, g, ale $m \nmid f, g$, spor. Čili m je minimální polynom prvku a nad $\mathbf{T}$. Pro jakýkoliv jiný monický ireducibilní polynom $\tilde{m} \in T[x]$, jehož je a kořenem, platí $m \mid \tilde{m}$ a z ireducibility a moničnosti dostáváme $\tilde{m} = m$. □

Příklad. Je ihned vidět, že

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

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

Příklad. Pozor, pro $\zeta_3 = e^{2\pi i/3}$ minimální polynom $m_{\zeta_3,\mathbb{Q}}$ není $x^3 - 1$, neboť tento polynom není ireducibilní. Platí $x^3 - 1 = (x - 1)(x^2 + x + 1)$, ζ_3 je kořenem druhého činitele, ten je ireducibilní, a tedy $m_{\zeta_3,\mathbb{Q}} = x^2 + x + 1$.

Příklad. Spočteme minimální polynom prvku $a = \sqrt{2} + \sqrt{3}$. Platí

$$a^2 = 5 + 2\sqrt{6}, \quad a^3 = 11\sqrt{2} + 9\sqrt{3}, \quad a^4 = 49 + 20\sqrt{6}$$

a vidíme, že $a^4 = 10a^2 - 1$. Čili a je kořenem polynomu $x^4 - 10x^2 + 1$. Tento polynom je ireducibilní: díky Tvrzení 8.5 nemá racionální kořen a na součin dvou polynomů stupňů 2 se rozkládat nemůže, neboť $\sqrt{2} + \sqrt{3}$ není řešením žádné kvadratické rovnice.

Tvrzení 22.2 (struktura jednoduchých rozšíření)**.** *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 3.1 je

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

Dokážeme, že tyto prvky tvoří podtěleso. Mějme tedy nějaký prvek $0 \neq f(a) \in T[a]$, hledáme jeho inverz, tedy 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 $\mathrm{NSD}(m_{a,\mathbf{T}}, f) = 1$, čili 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 prvek k $f(a)$. $\square$

Alternativní důkaz. Uvažujme homomorfismus $\varphi : \mathbf{T}[x] \to \mathbf{T}[a]$, $f \mapsto f(a)$. Ten je zřejmě na, jeho jádro je ideál $m_{a,\mathbf{T}}T[x]$, a tak podle 1. věty o izomorfismu $\mathbf{T}[x]/(m_{a,\mathbf{T}}) \simeq \mathbf{T}[a]$. Protože je $m_{a,\mathbf{T}}$ ireducibilní, podle Tvrzení 9.4 je $\mathbf{T}[a]$ těleso, a tedy $\mathbf{T}[a] = \mathbf{T}(a)$. $\square$

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

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

Pro rozšíření vyšších stupňů vycházejí vzorce ošklivě (zkuste si to!) a Tvrzení 22.2 má svoji cenu.

Poznámka. Je-li a transcendentní prvek nad $\mathbf{T}$, pak $\mathbf{T}[a] \neq \mathbf{T}(a)$. Kdyby $\frac{1}{a} \in \mathbf{T}[a]$, pak by existoval polynom $f \in \mathbf{T}[x]$ takový, že $f(a) = a^{-1}$, čili $af(a) = 1$, a tedy a by bylo kořenem polynomu $xf - 1 \in T[x]$, spor.

Tvrzení 22.3 (stupeň jednoduchých rozšíření)**.** *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 $\sum_{i=0}^{n-1} t_i a^i = 0$ pro nějaká $t_i \in T$, z nichž by aspoň jedno bylo nenulové. Prvek a by tedy byl kořenem (nenulového) polynomu $\sum_{i=0}^{n-1} t_i x^i \in T[x]$ s menším stupněm než $m_{a,\mathbf{T}}$, což by byl spor s minimalitou.

Nyní dokážeme, že prvky $1, a, a^2, \ldots, a^{n-1}$ 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 $f(a) = r(a) = \sum_{i=0}^{n-1} t_i a^i$, kde $t_i \in T$ jsou koeficienty polynomu r. $\square$

Příklad. Pomocí Tvrzení 22.3 lze určit stupeň jednoduchého rozšíření.

- $[\mathbb{C} : \mathbb{R}] = [\mathbb{R}(i) : \mathbb{R}] = \deg m_{i,\mathbb{R}} = \deg(x^2 + 1) = 2$.
- $[\mathbb{Q}(\sqrt[n]{p}) : \mathbb{Q}] = \deg(x^n - p) = n$ pro libovolné $n \in \mathbb{N}$ a prvočíslo p, protože uvedený polynom je podle Eisensteinova kritéria ireducibilní. (Pokud p není prvočíslo, situace je složitější.)
- $[\mathbb{Q}(e^{2\pi i/n}) : \mathbb{Q}] = \varphi(n)$ (hodnota Eulerovy funkce), což ale není snadné dokázat, používá se k tomu teorie cyklotomických polynomů, viz přednáška Teorie čísel. Je-li n prvočíslo, minimálním polynomem je $x^{n-1} + x^{n-2} + \ldots + 1 = \frac{x^n - 1}{x - 1}$, jehož ireducibilitu lze po substituci ukázat z Eisensteinova kritéria.

Důsledek 22.4. *Buď* $\mathbf{T} \leq \mathbf{S}$ *rozšíření těles a* $a \in S$*. Prvek a je algebraický nad* $\mathbf{T}$ *právě tehdy, když je stupeň* $[\mathbf{T}(a) : \mathbf{T}]$ *konečný.*

Důkaz. Je-li a transcendentní, pak $1, a, a^2, \dots$ tvoří nekonečnou lineárně nezávislou množinu: kdyby $\sum_{i=0}^{n} t_i a^i = 0$ pro nějaké koeficienty $t_i \in T$, aspoň jeden nenulový, bylo by a kořenem nenulového polynomu $\sum_{i=0}^{n} t_i x^i$ z $\mathbf{T}[x]$, spor. Opačná implikace plyne z Tvrzení 22.3. $\square$

Příklad. 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

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

Buď $1, a$ báze prostoru $\mathbf{S_T}$. Pak $\mathbf{S} = \mathbf{T}(a)$ a podle Tvrzení 22.3 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 $\mathbf{S} = \mathbf{T}(u + v\sqrt{s}) = \mathbf{T}(\sqrt{s})$.

22.2. **Vícenásobná rozšíření.**

K výpočtu stupně vícenásobného rozšíření slouží následující obecné pravidlo. (Platí i pro nekonečné stupně, viz poznámka nad Větou 14.9.)

Tvrzení 22.5 (stupeň vícenásobných rozšíření)**.** *Buď $\mathbf{T} \leq \mathbf{S} \leq \mathbf{U}$ rozšíření těles. Pak*

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

Důkaz. Zvolme bázi A vektorového prostoru $\mathbf{S_T}$ a bázi B vektorového prostoru $\mathbf{U_S}$. Dokážeme, že

$$C = \{ab : \ a \in A, \ b \in B\}$$

je bází vektorového prostoru $\mathbf{U_T}$.

Nejprve dokážeme, že C generuje prostor $\mathbf{U_T}$ (jistě $C \subseteq U$, a tedy C generuje podprostor $\mathbf{U_T}$). Je-li $u \in U$, pak $u = \sum_j s_j b_j$ pro nějaká $s_j \in S$ a $b_j \in B$. Každé s_j lze napsat jako $s_j = \sum_i t_{ij} a_i$ pro nějaká $t_{ij} \in T$ a $a_i \in A$, a dosazením druhé rovnosti do první dostaneme

$$u = \sum_j \Big(\sum_i t_{ij} a_i\Big) b_j = \sum_{i,j} t_{ij} \cdot a_i b_j.$$

Tedy u je lineární kombinací prvků C 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 $t_{ij} \in T$ a $a_i b_j \in C$. 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_j nad tělesem $\mathbf{S}$ nám dává $\sum_i t_{ij} a_i = 0$ pro každé j a z lineární nezávislosti prvků a_i nad tělesem $\mathbf{T}$ dostáváme $t_{ij} = 0$ pro všechna i, j.

Z lineární nezávislosti také plyne, že prvky ab, $a \in A$, $b \in B$, jsou po dvou různé, a tedy

$$[\mathbf{U} : \mathbf{T}] = |C| = |A \times B| = |A| \cdot |B| = [\mathbf{S} : \mathbf{T}] \cdot [\mathbf{U} : \mathbf{S}].$$

$\square$

Tvrzení 22.3 a 22.5 můžeme aplikovat na výpočet stupně vícenásobných rozšíření typu $\mathbf{T}(a_1, a_2, \dots)$. Dvojité rozšíření $\mathbf{T} \leq \mathbf{T}(a, b)$ můžeme rozbít na dvě jednoduchá rozšíření $\mathbf{T} \leq \mathbf{T}(a) \leq \mathbf{T}(a, b)$ a spočteme

$$\begin{aligned}[\mathbf{T}(a, b) : \mathbf{T}] = [\mathbf{T}(a, b) : \mathbf{T}(a)] \cdot [\mathbf{T}(a) : \mathbf{T}] &= \deg m_{b,\mathbf{T}(a)} \cdot \deg m_{a,\mathbf{T}} \\ &\leq \deg m_{b,\mathbf{T}} \cdot \deg m_{a,\mathbf{T}}.\end{aligned}$$

Pozor, při vyjádření stupně $[\mathbf{T}(a, b) : \mathbf{T}(a)]$ musíme použít minimální polynom prvku b nad tělesem $\mathbf{T}(a)$, který může být menšího stupně, než minimální polynom nad tělesem $\mathbf{T}$. Vícenásobným použitím popsaného postupu snadno dokážeme následující důsledek.

Důsledek 22.6. *Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles a $a_1, \dots, a_n \in S$ prvky algebraické nad $\mathbf{T}$. Pak $\mathbf{T}(a_1, \dots, a_n)$ je rozšířením konečného stupně nad $\mathbf{T}$.*

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:

- $m_{\sqrt{2}+\sqrt{3},\mathbb{Q}} = x^4 - 10x^2 + 1$;
- $m_{\sqrt{2},\mathbb{Q}} = x^2 - 2$;
- $m_{\sqrt{3},\mathbb{Q}(\sqrt{2})} = x^2 - 3$ (ověřte, že je opravdu ireducibilní v $\mathbb{Q}(\sqrt{2})[x]$!).

Podle Tvrzení 22.3 a 22.5 dostáváme $[\mathbb{Q}(\sqrt{2}+\sqrt{3}) : \mathbb{Q}] = 4$ a $[\mathbb{Q}(\sqrt{2},\sqrt{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{2},\sqrt{3}) : \mathbb{Q}(\sqrt{2})] \cdot [\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = 2 \cdot 2 = 4$.

Je-li $\mathbf{T} \leq \mathbf{S}$ rozšíření těles a každý prvek tělesa $\mathbf{S}$ je algebraický nad $\mathbf{T}$, hovoříme o *algebraickém rozšíření.* Tuto vlastnost mají všechna rozšíření konečného stupně.

Tvrzení 22.7. *Rozšíření konečného stupně jsou algebraická.*

Důkaz. Označme $n = [\mathbf{S} : \mathbf{T}]$. Pro libovolný prvek $a \in S$ dokážeme, že 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 $t_i \in T$, aspoň jeden z nich nenulový, kterými lze lineárně nakombinovat nulu, tj. $\sum_{i=0}^n t_i a^i = 0$. Čili prvek a je kořenem nenulového polynomu $\sum_{i=0}^n t_i x^i \in T[x]$. $\square$

Opačná implikace neplatí: příkladem je algebraický uzávěr tělesa $\mathbb{Q}$, který má nekonečný stupeň nad $\mathbb{Q}$.

Tvrzení 22.7 je principem nekonstruktivních důkazů algebraičnosti: k důkazu, že je prvek a algebraický nad $\mathbf{T}$, stačí najít rozšíření $\mathbf{S} \geq \mathbf{T}$ konečného stupně, v němž a leží. Typickým příkladem je důkaz, že součet, rozdíl, součin a podíl algebraických prvků je algebraický prvek.

Věta 22.8 (algebraické prvky tvoří podtěleso). *Buď* $\mathbf{T} \leq \mathbf{S}$ *rozšíření těles. Prvky* $\mathbf{S}$*, které jsou algebraické nad* $\mathbf{T}$*, tvoří podtěleso tělesa* $\mathbf{S}$*.*

Důkaz. Uvažujme prvky $a, b \in S$ algebraické nad $\mathbf{T}$. Rozšíření $\mathbf{T} \leq \mathbf{T}(a,b)$ je konečného stupně (Důsledek 22.6), a tedy algebraické (Tvrzení 22.7). 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 $\mathbf{S}$. $\square$

Jinou aplikací Tvrzení 22.7 je fakt, že každé rozšíření $\mathbf{T} \leq \mathbf{S}$ konečného stupně lze napsat jako $\mathbf{S} = \mathbf{T}(a_1, \ldots, a_n)$, kde $a_1, \ldots, a_n$ jsou nějaké algebraické prvky nad $\mathbf{T}$. Tento fakt lze dokázat snadno indukcí podle $k = [\mathbf{S} : \mathbf{T}]$. Pro $k = 1$ je $\mathbf{S} = \mathbf{T}$. V indukčním kroku zvolme prvek $a \in S \setminus T$. Ten musí být algebraický, rozšíření rozbijeme jako $\mathbf{T} < \mathbf{T}(a) \leq \mathbf{S}$ a aplikujeme indukční předpoklad na rozšíření $\mathbf{T}(a) \leq \mathbf{S}$. Detaily důkazu si doplňte jako cvičení.

Tento fakt má menší význam než se zdá. Již jsme dokázali, že každé konečné těleso $\mathbf{T}$ charakteristiky p lze napsat jako $\mathbf{T} = \mathbb{Z}_p(a)$: stačí vzít generátor a cyklické grupy $\mathbf{T}^*$ (Věta 16.7). Na tělesa charakteristiky 0 se pak aplikuje tzv. *Artinova věta o primitivním prvku*, která říká, že (za jistých předpokladů) každé rozšíření $\mathbf{T} \leq \mathbf{S}$ konečného stupně lze napsat jako $\mathbf{S} = \mathbf{T}(a)$, kde a je nějaký algebraický prvek nad $\mathbf{T}$. Důkaz této věty je poměrně komplikovaný, čtenáře odkazujeme do libovolné učebnice komutativní algebry. (Pro nekonečná tělesa nenulové charakteristiky existuje řada protipříkladů, můžete si zkusit nějaký najít třeba pro podílové těleso oboru $\mathbb{Z}_p[x,y]$.)

23. Neřešitelnost úloh pravítkem a kružítkem

Mezi klasické matematické úlohy s kořeny v antickém Řecku patří konstrukce pomocí pravítka a kružítka. Některé úlohy jsou snadné a učí se na základní škole: například zdvojení čtverce či půlení úhlu. Jsou úlohy, které odolávaly tisíciletí: například konstrukci pravidelného sedmnáctiúhelníka objevil Gauss roku 1796. Už v té době se tušilo, že některé úlohy zřejmě řešit nepůjdou, ale byl to až rozvoj algebry počátkem 19. století, který to umožnil dokázat. Mezi nejznámější takové úlohy patří:

- *zdvojení krychle*: k dané úsečce sestrojit úsečku, která je $\sqrt[3]{2}$-krát delší (původní formulace: k dané úsečce u sestrojit úsečku v takovou, že krychle s hranou v má dvakrát větší objem, než krychle s hranou u);
- *trisekce úhlu*: k danému úhlu sestrojit třetinový úhel;
- *rektifikace kružnice* a *kvadratura kruhu*: k dané úsečce sestrojit sestrojit úsečku, která je π-krát delší (původní formulace: k dané kružnici k sestrojit úsečku, která je stejně dlouhá jako obvod k, resp. úsečku takovou, že čtverec nad ní sestrojený má stejný obsah kruh daný k; obě úlohy lze snadno převést na konstrukci π-krát delší úsečky).
- *konstrukce pravidelných n-úhelníků*, pro některá n.

V této sekci si ukážeme důkazovou metodu, kterou vymyslel Pierre Wantzel roku 1837. Její pomocí lze dokázat neřešitelnost všech uvedených úloh (v některých případech za pomoci další teorie, jako je důkaz transcendence čísla π).

Předně musíme upřesnit, co vlastně rozumíme konstrukcí pravítkem a kružítkem. 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.

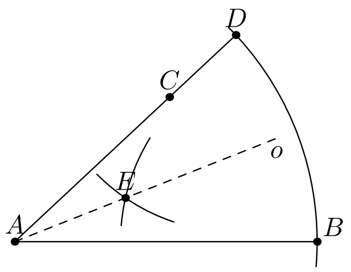
Formálně, *konstrukce pravítkem a kružítkem* je posloupnost $\mathcal{M}_0 \subseteq \mathcal{M}_1 \subseteq \ldots \subseteq \mathcal{M}_n$ konečných množin bodů v rovině taková, že $\mathcal{M}_{i+1} = \mathcal{M}_i \cup \{X\}$, kde X vznikne jako

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

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: místo množin bodů budeme uvažovat tělesa souřadnic. Zvolme v rovině souřadnice a uvažujme nejmenší těleso $\mathbf{T}_i \leq \mathbb{R}$, které obsahuje x-ové i y-ové souřadnice všech bodů z $\mathcal{M}_i$. Čili, pokud $\mathcal{M}_i$ obsahuje body $A_1, \ldots, A_k$ se souřadnicemi $(a_1, b_1), \ldots, (a_k, b_k)$, pak $\mathbf{T}_i = \mathbb{Q}(a_1, b_1, \ldots, a_k, b_k)$. Přidáním bodu X se souřadnicemi (u, v) dostaneme $\mathbf{T}_{i+1} = \mathbf{T}_i(u, v)$. Výsledkem je 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.

Lemma 23.1. *Pro každou konstrukci pravítkem a kružítkem je $[\mathbf{T}_{i+1} : \mathbf{T}_i] \in \{1, 2\}$ pro každé i.*

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

(1) Jde-li o průsečík dvou různoběžných přímek, získáme souřadnice nového bodu řešením soustavy dvou lineárních rovnic o dvou neznámých nad tělesem $\mathbf{T}_i$. Konkrétně, přímka určená body A, B se souřadnicemi (a, b), (c, d), kde $a, b, c, d \in T_i$, má rovnici

$$(b - d)x + (c - a)y = bc - ad$$

a vidíme, že všechny tři koeficienty jsou v tělese $\mathbf{T}_i$. Řešením soustavy lineárních rovnic dvou proměnných nad tělesem $\mathbf{T}_i$ je dvojice (u, v) prvků tělesa $\mathbf{T}_i$, takže $\mathbf{T}_{i+1} = \mathbf{T}_i(u, v) = \mathbf{T}_i$ 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 $\mathbf{T}_i$. Přímku jsme si rozebrali výše, a kružnice $k(A, |BC|)$ určená body A, B, C se souřadnicemi (a, b), (c, d), (e, f), kde $a, b, c, d, e, f \in T_i$, má rovnici

$$(x - a)^2 + (y - b)^2 = (c - e)^2 + (d - f)^2$$

a vidíme, že všechny koeficienty jsou v tělese $\mathbf{T}_i$. Vyjádříme-li z rovnice přímky y a dosadíme jej do kvadratické, dostaneme kvadratickou rovnici pro x, jejíž koeficienty jsou z $\mathbf{T}_i$ a řešením je $x = u + v\sqrt{s}$ pro nějaká $u, v, s \in T_i$. Dosazením do lineární rovnice zjistíme, že $y = u' + v'\sqrt{s}$ pro nějaká $u', v' \in T_i$. Čili $\mathbf{T}_{i+1} = \mathbf{T}_i(u, v) = \mathbf{T}_i(\sqrt{s})$, 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. (Proveďte popsaný výpočet podrobně a ověřte, že skutečně obě řešení náleží $\mathbf{T}_i(\sqrt{s})$!)

(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, vše nad tělesem $\mathbf{T}_i$. Stejným argumentem jako v (2) dostaneme

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

(Proveďte popsaný výpočet podrobně sami!) □

Tvrzení 23.2 (stupeň rozšíření pro konstrukce pravítkem a kružítkem)**.** *Pro každou konstrukci pravítkem a kružítkem je $[\mathbf{T}_n : \mathbf{T}_0] = 2^k$ pro nějaké $k \leq n$.*

Důkaz. Podle Tvrzení 22.5 je

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

což je součin jedniček a dvojek. □

Příklad (zdvojení krychle)**.** Zvolme souřadnice tak, že krajní body zadané úsečky (symbolizující hranu krychle) 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 $\mathbf{T}_n$, čili $\mathbb{Q} \leq \mathbb{Q}(\sqrt[3]{2}) \leq \mathbf{T}_n$ a podle Tvrzení 22.5 je

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

což je ve sporu s Tvrzením 23.2.

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

Příklad (rektifikace kružnice a kvadratura kruhu)**.** Analogicky, 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 π (resp. 2π a $\sqrt{\pi}$ v původním zadání). Čili transcendentní číslo π by mělo být prvkem tělesa $\mathbf{T}_n$, ale to je podle Tvrzení 23.2 rozšířením $\mathbb{Q}$ konečného stupně, a tedy podle Tvrzení 22.7 obsahuje pouze algebraická čísla, spor.

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

Příklad (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 22.3 nalézt minimální polynom čísla $\cos 20°$ nad tělesem $\mathbb{Q}(\sqrt{3})$, tj. nějaký ireducibilní polynom, jehož je číslo $\cos 20°$ kořenem. Prolistujeme-li nějakou sbírku goniometrických vzorců, najdeme vztah

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

z kterého plyne, že $\cos 20°$ je kořenem polynomu $4x^3 - 3x - \frac{1}{2} \in \mathbb{Q}[x]$. 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 $[\mathbb{Q}(\sqrt{3}, \cos 20°) : \mathbb{Q}(\sqrt{3})] = \deg m_{\cos 20°, \mathbb{Q}(\sqrt{3})} = 3$.

24. Izomorfismy tělesových rozšíření

24.1. **Galoisova grupa rozšíření.**

Galoisova teorie studuje grupy symetrií tělesových rozšíření. Konkrétně, půjde o automorfismy většího tělesa, které zachovávají menší těleso.

Definice. Buď $\mathbf{T}, \mathbf{S}, \mathbf{U}$ tělesa taková, že $\mathbf{T} \leq \mathbf{S}$ a $\mathbf{T} \leq \mathbf{U}$. Okruhový izomorfismus $\varphi : \mathbf{S} \to \mathbf{U}$ se nazývá $\mathbf{T}$-*izomorfismus*, pokud $\varphi(t) = t$ pro každé $t \in T$.

Všimněte si, že $\mathbf{T}$-izomorfismus je lineárním zobrazením vektorových prostorů $\mathbf{S}_\mathbf{T} \to \mathbf{U}_\mathbf{T}$: obě definice vyžadují $\varphi(a + b) = \varphi(a) + \varphi(b)$ pro všechna $a, b \in S$ a pro skalární násobení platí $\varphi(t \cdot a) = \varphi(t) \cdot \varphi(a) = t \cdot \varphi(a)$ pro všechna $t \in T$, $a \in S$. Opačná implikace samozřejmě neplatí: je řada lineárních zobrazení, které nezachovávají násobení.

Definice. Buď $\mathbf{T} \leq \mathbf{S}$ rozšíření těles. $\mathbf{T}$-izomorfismy $\mathbf{S} \to \mathbf{S}$ se nazývají $\mathbf{T}$-*automorfismy*. Je snadné nahlédnout, že jsou uzavřeny na skládání a invertování (drobné rozšíření Tvrzení 20.3), a tedy tvoří podgrupu symetrické grupy na množině S. Tato grupa se nazývá *Galoisova grupa rozšíření* $\mathbf{T} \leq \mathbf{S}$ a značí se $\mathbf{Gal}(\mathbf{S}/\mathbf{T})$.

Příklad. Spočteme grupu $\mathbf{Gal}(\mathbb{C}/\mathbb{R})$. Báze vektorového prostoru $\mathbb{C}_\mathbb{R}$ je $1, i$. Uvažujme $\mathbb{R}$-automorfismus φ. Nutně $\varphi(1) = 1$, neboť $1 \in \mathbb{R}$. Dále $\varphi(i)^2 = \varphi(i^2) = \varphi(-1) = -1$, a tedy $\varphi(i) \in \{i, -i\}$. Protože $\varphi(a + bi) = a + b\varphi(i)$, dostáváme přesně dvě možnosti: $\varphi = id$ a $\varphi = {}^-$, komplexní sdružení. Obě zobrazení jsou okruhovými homomorfismy, a tedy

$$\mathrm{Gal}(\mathbb{C}/\mathbb{R}) = \{id, {}^-\}, \qquad \mathbf{Gal}(\mathbb{C}/\mathbb{R}) \simeq \mathbb{Z}_2.$$

Příklad. Obecně je výpočet Galoisových grup obtížný a výsledek předem nejasný. Například, platí, ale není snadné dokázat, že $\mathbf{Gal}(\mathbb{R}/\mathbb{Q})$ je jednoprvková, zatímco $\mathbf{Gal}(\mathbb{C}/\mathbb{Q})$ je nekonečná.

V dalším textu se soustředíme na případ $\mathbf{S} = \mathbf{T}(a_1, \ldots, a_n)$, kde $a_1, \ldots, a_n$ jsou algebraické prvky nad $\mathbf{T}$. Základním pozorováním je, že $\mathbf{T}$-automorfismy jsou určeny hodnotami na prvcích $a_1, \ldots, a_n$. Buď φ nějaký $\mathbf{T}$-automorfismus a označme $\varphi(a_i) = u_i$. Obecný prvek $s \in S$ můžeme vyjádřit jako součet

$$s = \sum c_{i_1,\ldots,i_n} a_1^{i_1} \cdots a_n^{i_n}$$

pro nějaká $c_{i_1,\dots,i_n} \in T$ (pro $n = 1$ viz Tvrzení 3.1 a 22.2, zobecnění na více prvků je přímočaré) a jeho obraz pak bude

$$\varphi(s) = \sum \varphi(c_{i_1,\dots,i_n})\varphi(a_1)^{i_1} \cdots \varphi(a_n)^{i_n} = \sum c_{i_1,\dots,i_n} u_1^{i_1} \cdots u_n^{i_n}.$$

Ovšem pozor, ne každá volba hodnot u_i dává $\mathbf{T}$-automorfismus. Jak jsme viděli na příkladu výše, pro $\mathbf{S} = \mathbb{Q}(i)$ jsou jediné možnosti $\varphi(i) \in \{\pm i\}$. Obecný princip formuluje následující tvrzení.

Tvrzení 24.1 (Galoisova grupa a kořeny polynomů)**.** *Buď* $\mathbf{T} \leq \mathbf{S}$ *rozšíření těles,* $f \in T[x]$ *a* $A \subseteq S$ *množina všech kořenů polynomu* f *v* $S \setminus T$*. Pro každé* $\varphi \in \mathrm{Gal}(\mathbf{S}/\mathbf{T})$ *je* $\varphi|_A$ *permutací množiny* A *a zobrazení*

$$\mathbf{Gal}(\mathbf{S}/\mathbf{T}) \to \mathbf{S}_A, \qquad \varphi \mapsto \varphi|_A$$

je grupovým homomorfismem.

Důkaz. Označme $f = \sum c_i x^i$ a uvažujme jeho kořen $a \in S$. Pak $\varphi(a)$ je také kořenem f, protože

$$f(\varphi(a)) = \sum c_i \varphi(a)^i = \sum \varphi(c_i)\varphi(a)^i = \varphi\left(\sum c_i a^i\right) = \varphi(f(a)) = \varphi(0) = 0,$$

kde druhá rovnost využívá faktu, že $\varphi|_T$ je identita. Přitom kořeny, které leží v T, se musí zobrazit na sebe, a z prostosti φ plyne, že se na ně nezobrazí žádný kořen z $S \setminus T$. Čili $\varphi|_A$ zobrazuje A do A, je prosté, množina A je konečná, takže musí být permutací. Uvedené zobrazení pak očividně zachovává skládání permutací. □

Příklad. Spočteme grupu $\mathbf{Gal}(\mathbb{Q}(\sqrt{s})/\mathbb{Q})$, kde s není čtvercem. Buď φ nějaký $\mathbb{Q}$-automorfismus. Prvek $\sqrt{s}$ je kořenem polynomu $f = x^2 - s$ a podle Tvrzení 24.1 se musí zobrazit na některý z kořenů f v $\mathbb{Q}(\sqrt{s})$. Ty jsou pouze dva, čili $\varphi(\sqrt{s}) \in \{\pm\sqrt{s}\}$ a dostáváme zobrazení $\varphi(a + b\sqrt{s}) = a \pm b\sqrt{s}$. Snadno ověříme, že jde o $\mathbb{Q}$-automorfismy, a tedy

$$\mathbf{Gal}(\mathbb{Q}(\sqrt{s})/\mathbb{Q}) \simeq \mathbb{Z}_2.$$

Příklad. Spočteme grupu $\mathbf{Gal}(\mathbb{Q}(\sqrt[3]{s})/\mathbb{Q})$, kde $\sqrt[3]{s} \notin \mathbb{Q}$. Buď φ nějaký $\mathbb{Q}$-automorfismus. Prvek $\sqrt[3]{s}$ je kořenem polynomu $f = x^3 - s$ a podle Tvrzení 24.1 se musí zobrazit na některý z kořenů f v $\mathbb{Q}(\sqrt[3]{s})$. Ale v tomto tělese f žádný jiný kořen nemá (oba zbylé kořeny v $\mathbb{C}$ jsou imaginární), čili máme pouze jeden $\mathbf{T}$-automorfismus, identitu. Galoisova grupa je jednoprvková.

24.2. Izomorfismy kořenových a rozkladových nadtěles.

Pro Galoisovu teorii jsou stěžejní Galoisovy grupy rozkladových nadtěles. V této podsekci doplníme teorii rozkladových nadtěles, která je nutná k jejich výpočtu.

Buď $\mathbf{T}$ těleso a f polynom z $\mathbf{T}[x]$ stupně ≥ 1. Připomeňme, že

- *kořenovým nadtělesem* pro f nad $\mathbf{T}$ rozumíme minimální rozšíření, ve kterém má polynom f kořen (tj. rozšíření $\mathbf{S}$, kde existuje $a \in S$ takové, že $\mathbf{S} = \mathbf{T}(a)$ a $f(a) = 0$);
- *rozkladovým nadtělesem* rozumíme minimální rozšíření, kde se rozkládá na lineární činitele (tj. rozšíření $\mathbf{S}$, kde existují $a_1, \dots, a_n \in S$ taková, že $\mathbf{S} = \mathbf{T}(a_1, \dots, a_n)$ a $f \parallel (x - a_1) \cdot \ldots \cdot (x - a_n)$).

Důsledek 9.7 prokazuje existenci těchto rozšíření.

Příklad. Díky základní větě algebry (Věta 12.1) víme, že kořenové i rozkladové nadtěleso polynomu f nad tělesem $\mathbb{Q}$ lze nalézt uvnitř tělesa $\mathbb{C}$: kořenovým bude libovolné $\mathbb{Q}(a)$, kde a je nějaký komplexní kořen f, a rozkladovým bude $\mathbb{Q}(a_1, \dots, a_m)$, kde $a_1, \dots, a_m$ jsou všechny komplexní kořeny f.

- Uvažujme polynom $x^2 + 1$. Jediným kořenovým nadtělesem obsaženým v $\mathbb{C}$ je těleso $\mathbb{Q}(i) = \mathbb{Q}(-i)$, které obsahuje oba kořeny $\pm i$, a tedy je i nadtělesem rozkladovým.

Připomeňme značení $\zeta_3 = e^{2\pi i/3}$.

- Uvažujme polynom $x^3 - 1$. Tento polynom má dvě různá kořenová nadtělesa v $\mathbb{C}$, a to $\mathbb{Q} = \mathbb{Q}(1)$ a $\mathbb{Q}(\zeta_3) = \mathbb{Q}(\zeta_3^2)$. Tato tělesa jistě nejsou $\mathbb{Q}$-izomorfní. To větší je rozkladové, neboť obsahuje všechny tři kořeny.

- Uvažujme polynom x^3-2. Tento polynom má dvě různá kořenová nadtělesa, $\mathbb{Q}(\sqrt[3]{2})$ a $\mathbb{Q}(\sqrt[3]{2}\cdot\zeta_3)$ (to druhé obsahuje oba imaginární kořeny). Ač to není vidět na první pohled, tato tělesa jsou $\mathbb{Q}$-izomorfní. Rozkladovým nadtělesem pak bude těleso $\mathbb{Q}(\sqrt[3]{2},\sqrt[3]{2}\cdot\zeta_3) = \mathbb{Q}(\sqrt[3]{2},\zeta_3)$.

Rozložitelné polynomy typicky nemají izomorfní kořenová nadtělesa: mimo jiné proto, že ireducibilní dělitelé různých stupňů vynucují různý stupeň příslušných kořenových nadtěles. Na druhou stranu, možná trochu překvapivě, pro ireducibilní polynomy jsou všechna kořenová nadtělesa izomorfní. Pro rozkladová nadtělesa máme izomorfismus také, tentokrát již bez předpokladu ireducibility.

Věta 24.2 (jednoznačnost kořenových a rozkladových nadtěles)**.** *Buď* $\mathbf{T}$ *těleso a* $f\in T[x]$ *stupně* ≥ 1.

(1) *Je-li* f *ireducibilní, pak každá dvě kořenová nadtělesa pro* f *nad* $\mathbf{T}$ *jsou* $\mathbf{T}$*-izomorfní.*
(2) *Každá dvě rozkladová nadtělesa pro* f *nad* $\mathbf{T}$ *jsou* $\mathbf{T}$*-izomorfní.*

V dalším výkladu (konkrétně k výpočtu Galoisových grup a jednoznačnosti algebraického uzávěru) budeme potřebovat silnější tvrzení o rozšiřování částečných izomorfismů mezi kořenovými a rozkladovými nadtělesy. Věta 24.2 tak bude speciálním případem Lemmat 24.3 a 24.4. K jejich formulaci je potřeba následující značení a pozorování.

Buď $\mathbf{T}\leq\mathbf{T}_1$, $\mathbf{T}\leq\mathbf{T}_2$ rozšíření těles a $\varphi:\mathbf{T}_1\to\mathbf{T}_2$ $\mathbf{T}$-izomorfismus. Zobrazení φ lze rozšířit na $\mathbf{T}$-izomorfismus oborů polynomů nad těmito tělesy (budeme jej opět značit φ):

$$\varphi:\mathbf{T}_1[x]\to\mathbf{T}_2[x],\qquad \sum a_ix^i\mapsto\sum\varphi(a_i)x^i.$$

Označme $f=\sum a_ix^i$, $g=\sum b_ix^i$. Koeficienty součtu $f+g$ jsou a_i+b_i, koeficienty součtu $\varphi(f)+\varphi(g)$ jsou $\varphi(a_i)+\varphi(b_i)=\varphi(a_i+b_i)$ a vidíme, že $\varphi(f+g)=\varphi(f)+\varphi(g)$. Koeficienty součinu fg jsou $\sum_{i+j=k}a_ib_j$, koeficienty součinu $\varphi(f)\varphi(g)$ jsou $\sum_{i+j=k}\varphi(a_i)\varphi(b_j)=\varphi(\sum_{i+j=k}a_ib_j)$ a vidíme, že $\varphi(fg)=\varphi(f)\varphi(g)$. Bijektivita zobrazení je zřejmá. Okamžitým důsledkem součinové vlastnosti je, že

- $f\mid g$ v $\mathbf{T}_1[x]$ právě tehdy, když $\varphi(f)\mid\varphi(g)$ v $\mathbf{T}_2[x]$;
- polynom f je ireducibilní v $\mathbf{T}_1[x]$ právě tehdy, když $\varphi(f)$ je ireducibilní v $\mathbf{T}_2[x]$.

Lemma 24.3. *Buď* $\mathbf{T}\leq\mathbf{T}_1$, $\mathbf{T}\leq\mathbf{T}_2$ *rozšíření těles a* $\varphi:\mathbf{T}_1\to\mathbf{T}_2$ $\mathbf{T}$*-izomorfismus. Buď* $f\in T_1[x]$ *ireducibilní polynom,* $\mathbf{T}_1(a)$ *kořenové nadtěleso pro* f *nad* $\mathbf{T}_1$ *a* $\mathbf{T}_2(b)$ *kořenové nadtěleso pro* $\varphi(f)$ *nad* $\mathbf{T}_2$. *Pak existuje* $\mathbf{T}$*-izomorfismus* $\psi:\mathbf{T}_1(a)\to\mathbf{T}_2(b)$ *takový, že* $\psi(a)=b$ *a* $\psi|_{T_1}=\varphi$.

Důkaz. Podle Tvrzení 22.2 je $T_1(a)=T_1[a]=\{g(a):\ g\in T_1[x]\}$ a $T_2(b)=T_2[b]=\{g(b):\ g\in T_2[x]\}$. Uvažujme tedy zobrazení

$$\psi:T_1(a)\to T_2(b),\qquad g(a)\mapsto g(b).$$

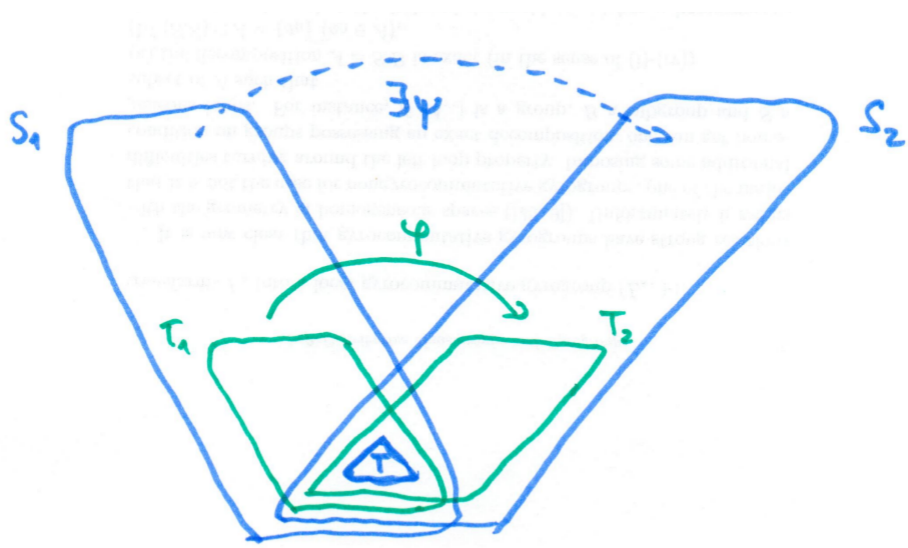
Předně je třeba dokázat, že to je dobře definované zobrazení. Označme $\tilde a=\varphi(a)$. Uvědomte si, že $f=m_{a,\mathbf{T}_1}$, protože f je ireducibilní polynom a a je jeho kořen, a zrovna tak $\varphi(f)=m_{\tilde a,\mathbf{T}_2}$, protože $\varphi(f)$ je ireducibilní polynom a $\tilde a$ je jeho kořen. Čili

$$g(a)=h(a)\ \Leftrightarrow\ (g-h)(a)=0\ \Leftrightarrow\ f\mid g-h$$

a analogicky

$$\varphi(g)(\tilde a)=\varphi(h)(\tilde a)\ \Leftrightarrow\ \varphi(g-h)(\tilde a)=0\ \Leftrightarrow\ \varphi(f)\mid\varphi(g-h).$$

Ekvivalence obou tvrzení na pravé straně plyne z pozorování výše. Dokázali jsme, že φ je dobře definované zobrazení a navíc prosté. Očividně jde o bijekci a je snadné ověřit, že jde o okruhový homomorfismus: pro každé $g,h\in T_1[x]$ platí $\psi(g(a)+h(a))=\psi((g+h)(a))=\varphi(g+h)(b)=\varphi(g)(b)+\varphi(h)(b)=\psi(g(a))+\psi(h(a))$ a analogicky pro násobení. Prvky tělesa $\mathbf{T}_1$ odpovídají volbě konstantního polynomu c, pro takový polynom platí $\psi(c)=\psi(c(a))=\varphi(c)(b)=\varphi(c)$, čili $\psi|_{\mathbf{T}_1}=\varphi$. Volbou $g=x$ ověříme, že $\varphi(a)=b$. $\square$



Obrázek 22. Ilustrace důkazu jednoznačnosti rozkladového nadtělesa.

Lemma 24.4. *Buď* $\mathbf{T} \leq \mathbf{T}_1$, $\mathbf{T} \leq \mathbf{T}_2$ *rozšíření těles a* $\varphi : \mathbf{T}_1 \to \mathbf{T}_2$ $\mathbf{T}$*-izomorfismus. Buď* $f \in T_1[x]$ *polynom stupně* ≥ 1 *a označme* $\mathbf{S}_1$ *rozkladové nadtěleso polynomu* f *nad* $\mathbf{T}_1$ *a* $\mathbf{S}_2$ *rozkladové nadtěleso polynomu* $\varphi(f)$ *nad* $\mathbf{T}_2$*. Pak existuje* $\mathbf{T}$*-izomorfismus* $\psi : \mathbf{S}_1 \to \mathbf{S}_2$ *takový, že* $\psi|_{T_1} = \varphi$.

Důkaz. Budeme postupovat indukcí podle stupně polynomu f. Je-li $\deg f = 1$, pak $\mathbf{S}_1 = \mathbf{T}_1$, $\mathbf{S}_2 = \mathbf{T}_2$ a $\psi = \varphi$. V indukčním kroku uvažujme ireducibilní dělitel g polynomu f a jeho kořen a v $\mathbf{S}_1$. Pak $\varphi(g)$ je ireducibilní dělitel polynomu $\varphi(f)$ a uvažujme jeho kořen b v $\mathbf{S}_2$. Podle Lemmatu 24.3 existuje zobrazení $\psi : \mathbf{T}_1(a) \to \mathbf{T}_2(b)$ takové, že $\psi(a) = b$ a $\psi|_{T_1} = \varphi$. Napišme $f = (x-a)\cdot h$ pro nějaký $h \in T_1[x]$, čili také $\psi(f) = (x-b)\cdot\psi(h)$. Pak $\mathbf{S}_1$ je rozkladové nadtěleso polynomu h nad $\mathbf{T}_1(a)$ a $\mathbf{S}_2$ je rozkladové nadtěleso polynomu $\psi(h)$ nad $\mathbf{T}_2(b)$. Protože $\deg h < \deg f$, podle indukčního předpokladu existuje $\mathbf{T}$-izomorfismus $\rho : \mathbf{S}_1 \to \mathbf{S}_2$ takový, že $\rho|_{T_1(a)} = \psi$, čili také $\rho|_{T_1} = \varphi$. □

Volbou $\mathbf{T}_1 = \mathbf{T}_2 = \mathbf{T}$ a $\varphi = id$ v obou lemmatech dostaneme Větu 24.2.

24.3. Galoisova grupa polynomu.

Definice. Buď f polynom z $\mathbf{T}[x]$ stupně ≥ 1. *Galoisovou grupou polynomu* f nad tělesem $\mathbf{T}$, značíme $\mathbf{Gal}(f/\mathbf{T})$, rozumíme jakoukoliv grupu $\mathbf{Gal}(\mathbf{S}/\mathbf{T})$, kde $\mathbf{S}$ je rozkladové nadtěleso polynomu f nad $\mathbf{T}$.

Dává tento pojem smysl, když je rozkladové nadtěleso určeno jednoznačně pouze *až na izomorfismus*? Uvažujme $\mathbf{T}$-izomorfismus $\psi : \mathbf{S}_1 \to \mathbf{S}_2$ dvou rozkladových nadtěles pro f. Je snadné ověřit (proveďte jako cvičení!), že

$$\mathbf{Gal}(\mathbf{S}_1/\mathbf{T}) \to \mathbf{Gal}(\mathbf{S}_2/\mathbf{T}), \qquad \varphi \mapsto \psi \circ \varphi \circ \psi^{-1}$$

je izomorfismem příslušných Galoisových grup. Čili Galoisova grupa polynomu jsou určena *až na izomorfismus.*

Příklad. V příkladech v sekci 24.1 jsme ukázali, že $\mathbf{Gal}(x^2 - s/\mathbb{Q}) \simeq \mathbb{Z}_2$. Ale pozor, $\mathbf{Gal}(x^3 - 2/\mathbb{Q})$ není jednoprvková grupa, neboť $\mathbb{Q}(\sqrt[3]{2})$ není rozkladové nadtěleso polynomu $x^3 - 2$.

Následující tvrzení umožňují určit Galoisovy grupy některých jednodušších polynomů.

Tvrzení 24.5 (základní vlastnosti Galoisových grup)**.** *Buď* $\mathbf{T}$ *těleso,* f *polynom z* $\mathbf{T}[x]$ *stupně* ≥ 1 *a* $\mathbf{S}$ *jeho rozkladové nadtěleso. Pak*

(1) $\mathbf{Gal}(\mathbf{S}/\mathbf{T})$ *je izomorfní podgrupě symetrické grupy* $\mathbf{S}_m$*, kde* m *je počet různých kořenů polynomu* f *v* $S \setminus T$*;*
(2) *je-li* f *ireducibilní, pak pro každé dva kořeny* $a, b \in S$ *existuje* $\varphi \in \mathrm{Gal}(\mathbf{S}/\mathbf{T})$ *takový, že* $\varphi(a) = b$*;*
(3) *pro každé rozšíření* $\mathbf{T} \leq \mathbf{S} \leq \mathbf{U}$ *takové, že* $\mathbf{U}$ *je také rozkladovým nadtělesem nějakého polynomu nad* $\mathbf{T}$*, platí* $\mathbf{Gal}(\mathbf{U}/\mathbf{S}) \trianglelefteq \mathbf{Gal}(\mathbf{U}/\mathbf{T})$ *a*

$$\mathbf{Gal}(\mathbf{U}/\mathbf{T}) \,/\, \mathbf{Gal}(\mathbf{U}/\mathbf{S}) \simeq \mathbf{Gal}(\mathbf{S}/\mathbf{T}).$$

Důkaz. (1) Označme $A = \{a_1, \dots, a_m\}$ množinu kořenů polynomu f v $S \setminus T$. Tvrzení 24.1 říká, že pro každé $\varphi \in \mathrm{Gal}(\mathbf{S}/\mathbf{T})$ je $\varphi|_A$ permutace na A a zobrazení

$$\mathbf{Gal}(\mathbf{S}/\mathbf{T}) \to \mathbf{S}_A, \qquad \varphi \mapsto \varphi|_A$$

je dobře definovaný homomorfismus. Dokážeme, že je prostý. Protože je $\mathbf{S}$ rozkladové pro f, platí $\mathbf{S} = \mathbf{T}(a_1, \dots, a_m)$. Čili každé $\varphi \in \mathrm{Gal}(\mathbf{S}/\mathbf{T})$ je jednoznačně určené svými hodnotami na prvcích $a_1, \dots, a_m$, a tedy také svojí restrikcí $\varphi|_A$.

(2) Podle Lemmatu 24.3 existuje $\mathbf{T}$-izomorfismus kořenových nadtěles $\psi : \mathbf{T}(a) \to \mathbf{T}(b)$ takový, že $\psi(a) = b$. Ten se podle Lemmatu 24.4 rozšiřuje do $\mathbf{T}$-izomorfismu $\rho : \mathbf{S} \to \mathbf{S}$ takového, že $\rho|_{T(a)} = \psi$, speciálně tedy $\rho(a) = b$.

(3) Podobně jako v části (1), Tvrzení 24.1 říká, že každé $\varphi \in \mathrm{Gal}(\mathbf{U}/\mathbf{T})$ permutuje kořeny polynomu f v $\mathbf{U}$, ovšem tyto kořeny generují těleso $\mathbf{S}$, takže $\varphi(S) = S$ a restrikce $\varphi|_S$ je $\mathbf{T}$-automorfismem tělesa $\mathbf{S}$. Čili zobrazení

$$\Phi : \mathbf{Gal}(\mathbf{U}/\mathbf{T}) \to \mathbf{Gal}(\mathbf{S}/\mathbf{T}), \qquad \varphi \mapsto \varphi|_S$$

je dobře definovaný homomorfismus. Dokážeme, že jeho jádrem je $\mathbf{Gal}(\mathbf{U}/\mathbf{S})$ a obrazem celé $\mathbf{Gal}(\mathbf{S}/\mathbf{T})$. Dokazované tvrzení pak plyne z faktu, že jádro je normální podgrupou, a z 1. věty o izomorfismu.

Jádro $\mathbf{Ker}(\Phi)$ obsahuje právě ty automorfismy φ, pro které $\varphi|_S$ je identita, tedy právě všechny $\mathbf{S}$-automorfismy tělesa $\mathbf{U}$, tedy $\mathbf{Ker}(\Phi) = \mathbf{Gal}(\mathbf{U}/\mathbf{S})$. Co se týče obrazu, je-li dáno $\psi \in \mathrm{Gal}(\mathbf{S}/\mathbf{T})$, čili $\mathbf{T}$-izomorfismus $\mathbf{S} \to \mathbf{S}$, podle Lemmatu 24.4 existuje $\mathbf{T}$-automorfismus φ tělesa $\mathbf{U}$ takový, že $\varphi|_S = \psi$, a tedy $\mathbf{Im}(\Phi) = \mathbf{Gal}(\mathbf{S}/\mathbf{T})$. □

Na několika příkladech ilustrujeme použití Tvrzení 24.5 k výpočtu Galoisových grup.

Příklad. Uvažujme ireducibilní polynom f stupně 2 nad tělesem $\mathbf{T}$. Podle bodu (1) je tato grupa nejvýše dvouprvková, podle bodu (2) musí mít alespoň dva prvky, čili $\mathbf{Gal}(f/\mathbf{T}) \simeq \mathbb{Z}_2$. Pro $\mathbf{T} = \mathbb{Q}$ je netriviálním prvkem $\mathbf{Gal}(f/\mathbb{Q})$ komplexní sdružení, tj. zobrazení $z \mapsto \bar{z}$.

Příklad. Uvažujme ireducibilní polynom f stupně 3 nad tělesem $\mathbb{Q}$ s dvěma imaginárními kořeny. Podle bodu (1) se $\mathbf{Gal}(f/\mathbb{Q})$ vnořuje do $\mathbf{S}_3$, podle bodu (2) je aspoň tříprvková. Vzhledem k tomu, že rozkladové nadtěleso obsahuje imaginární prvky, komplexní sdružení je jeho $\mathbb{Q}$-automorfismem řádu 2. Podle Lagrangeovy věty $\mathbf{Gal}(f/\mathbb{Q})$ nemůže být tříprvková, a tedy musí být izomorfní celé $\mathbf{S}_3$.

Platí, ale není úplně jednoduché dokázat, že ireducibilní polynom stupně 3 nad tělesem $\mathbb{Q}$ má Galoisovu grupu izomorfní $\mathbf{S}_3$ právě tehdy, když je jeho diskriminant D záporný (viz sekce 25.1); v opačném případě je Galoisova grupa tříprvková.

Příklad. Spočteme grupu

$$\mathbf{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}).$$

Aplikací Tvrzení 24.1 na polynomy $x^2 - 2$ a $x^2 - 3$ vidíme, že každý $\varphi \in \mathrm{Gal}(\mathbf{S}/\mathbb{Q})$ splňuje $\varphi(\sqrt{2}) = u\sqrt{2}$ a $\varphi(\sqrt{3}) = v\sqrt{3}$ pro nějaká $u, v \in \{\pm 1\}$, čili $\mathbb{Q}$-automorfismy jsou nejvýše čtyři a lze je zapsat vzorcem

$$\varphi(a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}) = a + ub\sqrt{2} + vc\sqrt{3} + uvd\sqrt{6}.$$

Z vzorce také vidíme, že $\varphi^2 = id$ pro všechny volby u, v.

Jsou to skutečně $\mathbb{Q}$-automorfismy? Je možné, ale velmi pracné, to ověřit přímo z definice. Jednodušší je využít bodu (2). V sekci 22.2 jsme ukázali, že

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

Aplikujeme-li bod (2) na minimální polynom $m_{\sqrt{2}+\sqrt{3},\mathbb{Q}}$, vidíme, že $|\mathbf{Gal}(\mathbf{S}/\mathbb{Q})| \geq 4$.

Shrnuto, $|\mathbf{Gal}(\mathbf{S}/\mathbb{Q})|$ je čtyřprvková grupa, a protože jsou všechny prvky řádu ≤ 2, platí $\mathbf{Gal}(\mathbf{S}/\mathbb{Q}) \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$.

Všimněte si, že ve všech příkladech vyšel stupeň rozkladového nadtělesa stejně, jako řád Galoisovy grupy. Pro tělesa charakteristiky 0 skutečně platí, že je-li $\mathbf{S}$ rozkladovým nadtělesem polynomu f nad $\mathbf{T}$, pak $[\mathbf{S} : \mathbf{T}] = |\mathbf{Gal}(\mathbf{S}/\mathbf{T})|$, nicméně důkaz této věty je značně nad rámec možností této učebnice.

Ve zbytku sekce se budeme věnovat dvěma speciálním případům. Za prvé, ukážeme, že rozkladová nadtělesa polynomů definujících n-té odmocniny mají řešitelné Galoisovy grupy; tento fakt je stěžejní v důkazu části Galoisovy věty (Věta 25.1), která říká, že polynomy, jejichž kořeny lze vyjádřit vzorcem, mají řešitelné Galoisovy grupy. Za druhé, ukážeme si polynom, jehož Galoisova grupa není řešitelná, a o němž tedy Galoisova věta říká, že jeho kořeny vzorcem vyjádřit nejdou.

Připomeňme značení $\zeta_n = e^{2\pi i/n}$.

Lemma 24.6. *Buď $0 \neq a \in \mathbb{Q}$. Rozkladovým nadtělesem polynomu $f = x^n - a$ nad tělesem $\mathbb{Q}$ je těleso $\mathbb{Q}(\zeta_n, b)$, kde b je libovolný komplexní kořen polynomu f.*

Důkaz. Komplexní kořeny polynomu f jsou právě čísla $b \cdot \zeta_n^k$, $k = 0, \ldots, n-1$: dosazením snadno ověříme, že každé z těchto čísel je kořenem, a víc kořenů být nemůže podle Tvrzení 2.4. Rozkladové nadtěleso tedy obsahuje jak prvek b (volbou $k = 0$), tak prvek ζ_n (součin $b^{-1} \cdot (b \cdot \zeta_n)$). Přitom každý kořen je součinem těchto dvou čísel, takže rozkladové nadtěleso můžeme napsat jako $\mathbb{Q}(\zeta_n, b)$. □

Tvrzení 24.7 (Galoisovy grupy pro odmocniny)**.** *Buď $\mathbb{Q} \leq \mathbf{T} \leq \mathbb{C}$ těleso, $n \in \mathbb{N}$, $a \in T$. Pak*

(1) $\mathbf{Gal}(x^n - 1/\mathbf{T})$ *je abelovská grupa,*
(2) $\mathbf{Gal}(x^n - a/\mathbf{T}(\zeta_n))$ *je abelovská grupa,*
(3) $\mathbf{Gal}(x^n - a/\mathbf{T})$ *je metabelovská grupa.*

Důkaz. Označme $\mathbf{S} = \mathbf{T}(\zeta_n)$ a $\mathbf{U} = \mathbf{T}(\zeta_n, b)$, kde b je nějaký komplexní kořen polynomu $x^n - a$.

(1) Dokážeme, že $\mathbf{Gal}(x^n - 1/\mathbf{T})$ je izomorfní nějaké podgrupě grupy $\mathbb{Z}_n^*$, tedy jde o abelovskou grupu. Každý automorfismus $\varphi \in \mathrm{Gal}(\mathbf{S}/\mathbf{T})$ permutuje kořeny polynomu $x^n - 1$, čili $\varphi(\zeta_n) = \zeta_n^k$ pro nějaké $k \in \{0, \ldots, n-1\}$. Zároveň také permutuje kořeny všech polynomů $x^m - 1$, $m \mid n$, tedy zobrazení φ zachovává řády prvků v grupě $\mathbf{S}^*$, takže $\mathrm{ord}(\zeta_n^k) = n$, což nastane právě tehdy, když $\mathrm{NSD}(k, n) = 1$ (Tvrzení 16.3). Vidíme, že zobrazení $\mathbf{Gal}(\mathbf{S}/\mathbf{T}) \to \mathbb{Z}_n^*$, které automorfismu φ přiřadí toto k, je prostý homomorfismus: prostý díky tomu, že φ je jednoznačně určeno hodnotou na generátoru, a homomorfismus díky tomu, že skládání automorfismů odpovídá násobení příslušných exponentů: je-li $\varphi(\zeta_n) = \zeta_n^k$ a $\psi(\zeta_n) = \zeta_n^l$, pak $\varphi(\psi(\zeta_n)) = (\zeta_n^l)^k = \zeta_n^{kl}$.

(2) Dokážeme, že $\mathbf{Gal}(x^n - a/\mathbf{S})$ je izomorfní nějaké podgrupě grupy $\mathbb{Z}_n$, tedy jde o abelovskou grupu. Kořeny polynomu $x^n - a$ v $\mathbf{S}$ jsou právě čísla tvaru $b \cdot \zeta_n^k$, $k = 0, \ldots, n-1$. Každý automorfismus $\varphi \in \mathrm{Gal}(\mathbf{U}/\mathbf{S})$ fixuje prvek ζ_n a zobrazuje $b \mapsto b \cdot \zeta_n^k$ pro nějaké k. Vidíme, že zobrazení $\mathbf{Gal}(\mathbf{U}/\mathbf{S}) \to \mathbb{Z}_n$, které automorfismu φ přiřadí toto k, je prostý homomorfismus: prostý díky tomu, že φ je jednoznačně určeno hodnotou na generátoru, a homomorfismus díky tomu, že skládání automorfismů odpovídá sčítání příslušných exponentů: je-li $\varphi(b) = b \cdot \zeta_n^k$ a $\psi(b) = b \cdot \zeta_n^l$, pak $\varphi(\psi(b)) = (b \cdot \zeta_n^l) \cdot \zeta_n^k = b \cdot \zeta_n^{k+l}$.

(3) Uvažujme rozšíření $\mathbf{T} \leq \mathbf{S} \leq \mathbf{U}$. Obě větší tělesa jsou rozkladová, můžeme tedy aplikovat Tvrzení 24.5(3), které říká, že $\{id\} \leq \mathbf{Gal}(\mathbf{U}/\mathbf{S}) \trianglelefteq \mathbf{Gal}(\mathbf{U}/\mathbf{T})$. Přitom grupa $\mathbf{Gal}(\mathbf{U}/\mathbf{S})$ je abelovská podle bodu (2), grupa $\mathbf{Gal}(\mathbf{U}/\mathbf{T}) \,/\, \mathbf{Gal}(\mathbf{U}/\mathbf{S}) \simeq \mathbf{Gal}(\mathbf{S}/\mathbf{T})$ je abelovská podle bodu (1), čili grupa $\mathbf{Gal}(\mathbf{U}/\mathbf{T})$ je metabelovská. □

Přestože většina polynomů stupně ≥ 5 nemá řešitelnou Galoisovu grupu, není úplně snadné nějaké konkrétní ukázat. Asi nejjednodušší rodinu příkladů popisuje následující tvrzení.

Tvrzení 24.8 (polynomy s plnou Galoisovou grupou)**.** *Buď p prvočíslo a $f \in \mathbb{Q}[x]$ ireducibilní polynom stupně p, který má $p-2$ reálných a 2 imaginární kořeny. Pak $\mathbf{Gal}(f/\mathbb{Q}) \simeq \mathbf{S}_p$.*

Důkaz. Buď $\mathbf{U}$ rozkladové nadtěleso polynomu f nad $\mathbb{Q}$. Podle Tvrzení 24.5(1) je grupa $\mathbf{Gal}(\mathbf{U}/\mathbb{Q})$ izomorfní podgrupě $\mathbf{H} \leq \mathbf{S}_p$, jejíž prvky sestávají z restrikcí prvků $\mathbf{Gal}(\mathbf{U}/\mathbb{Q})$ na kořeny polynomu f. Dokážeme, že $\mathbf{H}$ obsahuje aspoň jednu transpozici a aspoň jeden p-cyklus. Pak stačí využít pozorování, že libovolná transpozice a libovolný p-cyklus generují celou grupu $\mathbf{S}_p$ (viz cvičení v sekci 14.1).

Komplexní sdružení je netriviálním $\mathbb{Q}$-automorfismem tělesa $\mathbf{U}$. Přitom $p-2$ kořenů fixuje a 2 prohazuje, jde tedy o transpozici na kořenech.

Uvažujme působení grupy $\mathbf{H}$ na množině kořenů polynomu f. Podle Tvrzení 24.5(2) jde o tranzitivní působení, má tedy jednu orbitu velikosti p. Avšak velikost orbity dělí řád působící grupy (Tvrzení 18.3), čili $p \mid |\mathbf{H}|$. Podle Cauchyho věty (Věta 18.6) obsahuje grupa $\mathbf{H}$ prvek řádu p, což může být pouze p-cyklus. □

Příklad. Příkladem polynomu, který splňuje předpoklady Tvrzení 24.8, je třeba $f = x^5 - 4x + 2$. Tento polynom je ireducibilní podle Eisensteinova kritéria a počet reálných kořenů snadno zjistíme pomocí diferenciálního kalkulu: $f' = 5x^4 - 4$, tato rovnice má dvě reálná řešení, tedy příslušná reálná funkce f má jedno lokální maximum a jedno lokální minimum, přičemž snadno dopočítáme, že maximum je kladné a minimum záporné. Protože jsou polynomiální funkce spojité, f musí mít právě tři reálné kořeny.

25. (Ne)řešitelnost polynomů v radikálech

25.1. **Cardanovy vzorce.**

Polynomiální rovnice byly předmětem studia od samého počátku matematiky: na kvadratické (resp. kubické) rovnice přirozeně vede řada geometrických úloh v rovině (resp. prostoru). Některé typy kvadratických rovnic uměli řešit již starověcí matematikové, kompletní návod pochází od učence Al-Chvárizmího z 9. století. Rovnice vyšších stupňů však dlouho odolávaly. První obecný postup pro řešení rovnic třetího stupně našel Niccolò Tartaglia okolo roku 1530 a pro rovnice čtvrtého stupně Lodovico Ferrari o málo později. Jejich postupy byly publikovány v roce 1545 v knize *Ars Magna* Gerolama Cardana, a tak se vžilo označení *Cardanovy vzorce.* Jejich postupy, převedené do moderního jazyka, si nyní ukážeme.

Při výkladu Cardanových vzorců budeme uvažovat racionální polynomy a pracovat v tělese komplexních čísel, i když většina tvrzení je platná v libovolném tělese, kde používané operace (odmocniny, dělení 2, 3) dávají smysl.

<u>*Kvadratické rovnice.*</u> Budeme řešit rovnici

$$ax^2 + bx + c = 0.$$

Substitucí $x = y - \frac{b}{2a}$ dostaneme rovnici

$$y^2 = \frac{b^2 - 4ac}{4a^2},$$

tedy

$$y = \pm\frac{\sqrt{b^2 - 4ac}}{2a}$$

a zpětným dosazením dostaneme

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Vidíme, že se oba kořeny nacházejí v tělese $\mathbb{Q}(\sqrt{b^2 - 4ac})$, které je rozkladovým nadtělesem polynomu $ax^2 + bx + c$ nad $\mathbb{Q}$.

Klíčovou fintou byla *substituce*, která nás zbavila prostředního členu. Stejný trik lze použít i pro rovnice vyššího stupně: máme-li rovnici n-tého stupně

$$a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0,$$

tedy ekvivalentně

$$x^n + \frac{a_{n-1}}{a_n}x^{n-1} + \ldots + \frac{a_1}{a_n}x + \frac{a_0}{a_n} = 0,$$

substituujeme $x = y - \frac{a_{n-1}}{na_n}$. Po roznásobení získáme ekvivalentní rovnici s monickým polynomem a nulovým koeficientem u y^{n-1}. Nadále se budeme věnovat jen tomuto speciálnímu tvaru.

Kubické rovnice. Budeme řešit rovnici

$$x^3 + bx + c = 0.$$

Všimněte si, že pro libovolné u, v platí

$$(u+v)^3 - 3uv(u+v) - (u^3 + v^3) = 0.$$

Řešení rovnice budeme hledat ve tvaru $x = u + v$, přičemž aby to sedělo, pro koeficienty dostáváme rovnosti

$$b = -3uv, \qquad c = -u^3 - v^3.$$

Nyní je snadné vyjádřit u, v pomocí koeficientů b, c: dosazením $v = -\frac{b}{3u}$ do druhé rovnosti dostáváme

$$u^6 + cu^3 - \frac{b^3}{27} = 0,$$

což je kvadratická rovnice s neznámou u^3. Označíme-li diskriminant $D = c^2 + \frac{4}{27}b^3$, řešením je $u^3 = \frac{-c\pm\sqrt{D}}{2}$. Ze dvou možností $\pm$ uvažujme například součet (druhá volba by přinesla ty samé tři kořeny, ale v jiném pořadí). Označíme-li ω libovolnou třetí odmocninu z $\frac{-c+\sqrt{D}}{2}$ a $\zeta = \zeta_3 = e^{2\pi i/3}$, máme pro u tři řešení,

$$u_k = \zeta^k \cdot \omega, \quad k = 0, 1, 2,$$

a k nim snadno dopočteme odpovídající hodnoty

$$v_k = -\frac{b}{3u_k} = \zeta^{-k} \cdot \frac{b}{3\omega}, \quad k = 0, 1, 2.$$

Řešením původní rovnice jsou pak všechny tři součty $u_k + v_k$.

Je-li zlomek $\frac{-c+\sqrt{D}}{2}$ reálný, je přirozené zvolit reálnou odmocninu $\omega = \sqrt[3]{\frac{-c+\sqrt{D}}{2}}$ a úpravou zlomku $\frac{b}{3\omega} = \sqrt[3]{\frac{-c-\sqrt{D}}{2}}$ dostaneme tři kořeny

$$\zeta^k \cdot \sqrt[3]{\frac{-c+\sqrt{D}}{2}} - \zeta^{-k} \cdot \sqrt[3]{\frac{c+\sqrt{D}}{2}}, \quad k = 0, 1, 2.$$

Vidíme, že se všechny tři kořeny $u_k + v_k$ nacházejí v tělese $\mathbb{Q}(\zeta, \omega)$. Rozkladové nadtěleso polynomu $x^3 + bx + c$ je jeho podtělesem.

Příklad. Vyřešíme rovnici

$$x^3 - 6x - 9 = 0.$$

Diskriminant je roven $D = 49$, čili $u^3 = \frac{9+7}{2} = 8$ a $v^3 = \frac{-9+7}{2} = -1$. Volbou $\omega = 2$ dostaneme kořeny

$$x_0 = u - v = 3, \quad x_1 = \zeta u - \zeta^2 v = \frac{-3+\sqrt{3}i}{2}, \quad x_2 = \zeta^2 u - \zeta v = \frac{-3-\sqrt{3}i}{2}.$$

Rozkladovým nadtělesem je $\mathbb{Q}(\zeta) = \mathbb{Q}(\sqrt{3}i)$ a v tomto případě je totožné s výše popsaným tělesem $\mathbb{Q}(\zeta, \omega)$.

Příklad. Vyřešíme rovnici

$$x^3 - 3x + 1 = 0.$$

Diskriminant je roven $D = -3$, čili $u^3 = \frac{-1+\sqrt{3}i}{2}$ a $v^3 = \frac{1+\sqrt{3}i}{2}$, z čehož snadno vyjádříme kořeny jako rozdíly třetích odmocnin z jistých imaginárních čísel. Přesto, jak se snadno přesvědčíme vyšetřením průběhu funkce, všechny tři kořeny jsou reálné. Čili rozkladové nadtěleso neobsahuje žádný z prvků $\sqrt{D}, \zeta, \omega$.

Případu, kdy jsou reálné kořeny vyjádřeny pomocí odmocnin z imaginárních čísel, se říká *casus irreducibilis*. Pro některé polynomy se tomuto popisu nelze vyhnout, tj. neexistuje zápis kořenů pomocí vzorce, který by používal pouze reálná čísla. Tento fakt přesvědčil tehdejší matematiky k přijetí komplexních čísel jako smysluplného číselného oboru.

Kvartické rovnice. Budeme řešit rovnici

$$x^4 + bx^2 + cx + d = 0.$$

Napišme rovnici ve tvaru

$$x^4 + 2ux^2 + u^2 = -bx^2 - cx - d + 2ux^2 + u^2 = (2u - b)x^2 - cx + (u^2 - d),$$

kde u je jakýsi zatím neznámý parametr. Všimněte si, že levou stranu lze napsat jako $(x^2 + u)^2$. Kdybychom i pravou stranu uměli napsat jako druhou mocninu nějakého polynomu v proměnné x, mohli bychom obě strany odmocnit a získat tak kvadratickou rovnici pro x. Aby pravá strana byla mocninou, diskriminant musí být roven nule, tj.

$$c^2 - 4(2u - b)(u^2 - d) = 0.$$

Tím dostáváme rovnici třetího stupně pro u, přičemž nějaký její kořen u_0 nalezneme pomocí Tartagliova vzorce. S tímto u_0 můžeme obě strany dané rovnice odmocnit a získáme dvě kvadratické rovnice

$$x^2 + u_0 = rx + s \qquad \text{a} \qquad x^2 + u_0 = -rx - s,$$

kde r, s splňují $(rx + s)^2 = (2u_0 - b)x^2 - cx + (u_0^2 - d)$, čili

$$r = \sqrt{2u_0 - b}, \quad s = \sqrt{u_0^2 - d}$$

(je-li výraz pod odmocninou imaginární, vezmeme libovolnou odmocninu). Ačkoliv popsaný postup připomíná spíše algoritmus než vzorec, v principu je možné vyjádřit všechna čtyři řešení vzorcem, který používá koeficienty daného polynomu a základní operace $+, -, \cdot, /$ a odmocniny.

Příklad. Vyřešíme rovnici

$$x^4 + x^2 + 4x - 3 = 0.$$

Diskriminant vede na rovnici $-2u^3 + u^2 - 6u + 7 = 0$, která má řešení např. $u_0 = 1$. Původní rovnici upravíme na tvar $(x^2 + 1)^2 = (x - 2)^2$, a tak stačí řešit rovnice

$$x^2 + 1 = x - 2 \qquad \text{a} \qquad x^2 + 1 = -x + 2.$$

Řešením jsou čísla

$$\frac{1 \pm \sqrt{11}i}{2} \qquad \text{a} \qquad \frac{-1 \pm \sqrt{5}}{2}.$$

25.2. **Galoisova věta.**

Problém řešení rovnic stupně 5 a více zůstal otevřený dalších téměř 300 let po Cardanovi. Existují vzorce, které by vyjadřovaly kořeny polynomů stupně n pomocí jejich koeficientů za použití základních aritmetických operací $+, -, \cdot, /$ a n-tých odmocnin?

V roce 1799 přišel Paolo Ruffini s nápadem, jak dokázat, že vzorce pro rovnice stupně ≥ 5 neexistují. Jeho argument byl sice neúplný, ale na základě jeho myšlenek Niels Henrik Abel publikoval v roce 1823 kompletní důkaz. Tuto tzv. *Abel-Ruffiniho větu* si teď dokážeme. Půjdeme ale jinou cestou, kterou odhalil o 10 let později Évariste Galois. Jeho metoda je elegantnější a navíc umožňuje dokázat kritérium, které popisuje právě ty polynomy, jejichž kořeny lze vyjádřit vzorcem. Tedy nejen že neexistuje vzorec, který by fungoval pro všechny polynomy daného stupně zároveň, ale pro některé polynomy neexistuje ani jednorázové vyjádření kořenů.

Nejprve si však musíme ujasnit, co přesně znamená „vyjádřitelnost kořenů vzorcem“.

Definice. Buď $\mathbf{T} \leq \mathbf{U}$ rozšíření těles a $a \in U$. Řekneme, že prvek a je *vyjádřitelný v radikálech* nad tělesem $\mathbf{T}$, pokud existuje řada rozšíření $\mathbf{T} = \mathbf{T}_0 \leq \mathbf{T}_1 \leq \ldots \leq \mathbf{T}_k$ taková, že $a \in T_k$ a každé $\mathbf{T}_i$ je rozkladovým nadtělesem nějakého polynomu $x^{n_i} - a_i \in T_{i-1}[x]$ nad tělesem $\mathbf{T}_{i-1}$.

Příklad. Prvek
$$\frac{\sqrt[5]{\sqrt[3]{2}+1}}{i+1}$$
je vyjádřitelný nad $\mathbb{Q}$, neboť je prvkem rozšíření $\mathbb{Q} \leq \mathbf{T}_1 \leq \mathbf{T}_2 \leq \mathbf{T}_3$, kde postupně použijeme polynomy $x^3 - 2 \in \mathbb{Q}[x]$, $x^5 - (\sqrt[3]{2}+1) \in T_1[x]$ a $x^2 + 1 \in T_2[x]$.

Definice. Buď $\mathbf{T}$ těleso a f polynom z $\mathbf{T}[x]$. Řekneme, že polynom f je *řešitelný v radikálech* nad tělesem $\mathbf{T}$, pokud je každý kořen polynomu f vyjádřitelný v radikálech nad $\mathbf{T}$. Jinými slovy, pokud existuje řada rozšíření $\mathbf{T} = \mathbf{T}_0 \leq \mathbf{T}_1 \leq \ldots \leq \mathbf{T}_k$ taková, že každé $\mathbf{T}_i$ je rozkladovým nadtělesem nějakého polynomu $x^{n_i} - a_i \in T_{i-1}[x]$ nad tělesem $\mathbf{T}_{i-1}$, a rozkladové nadtěleso polynomu f je obsaženo v $\mathbf{T}_k$.

Příklad. Ukážeme si, jak se Cardanovy vzorce interpretují v rámci formální definice řešitelnosti.

- Kořeny polynomu $ax^2 + bx + c$ najdeme v tělese $\mathbb{Q}(\sqrt{b^2 - 4ac})$, které je rozkladové pro $x^2 - (b^2 - 4ac) \in \mathbb{Q}[x]$.
- Kořeny polynom $x^3 + bx + c$ najdeme v tělese $\mathbb{Q}(\zeta, \omega)$, které dostaneme jako řadu dvou rozkladových rozšíření
$$\mathbb{Q} \leq \mathbb{Q}(\sqrt{D}) \leq \mathbb{Q}(\zeta, \omega),$$
první pro polynom $x^2 - D \in \mathbb{Q}[x]$, druhé pro polynom $x^3 - \frac{-c+\sqrt{D}}{2} \in \mathbb{Q}(\sqrt{D})[x]$.
- Sledujíce Ferrariho postup pro polynom $x^4 + bx^2 + cx + d$ postupně budujeme řadu rozkladových rozšíření $\mathbb{Q} \leq \mathbf{T}_1 \leq \mathbf{T}_2 \leq \mathbf{T}_3$, kde $\mathbf{T}_1$ je rozkladové pro polynom $c^2 - 4(2x - b)(x^2 - d)$ nad $\mathbb{Q}$ (výpočet u_0), $\mathbf{T}_2$ je rozkladové pro polynomy $x^2 - (2u_0 - b)$, $x^2 - (u_0^2 - d)$ nad $\mathbf{T}_2$ (výpočet r, s) a $\mathbf{T}_3$ je rozkladové pro polynomy $x^2 \pm rx + u_0 \pm s$ (finální výpočet kořenů zadaného polynomu). Tuto řadu lze dále rozepsat tak, aby všechna nadtělesa byla rozkladová pro polynomy tvaru $x^n - a$, přičemž s trochou práce lze ukázat, že n bereme postupně $2, 3, 2, 2$ (což souvisí s tím, že v grupě $\mathbf{S}_4$ existuje řada podgrup $\mathbf{S}_4 = \mathbf{H}_0 \trianglerighteq \mathbf{H}_1 \trianglerighteq \mathbf{H}_2 \trianglerighteq \mathbf{H}_3 \trianglerighteq \mathbf{H}_4 = \{id\}$, kde $|\mathbf{H}_i/\mathbf{H}_{i+1}|$ je postupně 2,3,2,2, ale tak daleko se ve výkladu Galoisovy teorie nedostaneme).

Nyní můžeme zformulovat slavnou Galoisovu větu.

Věta 25.1 (Galoisova věta)**.** *Buď $\mathbf{T}$ těleso charakteristiky 0 a f polynom z $\mathbf{T}[x]$ stupně ≥ 1. Polynom f je řešitelný v radikálech právě tehdy, když je grupa $\mathbf{Gal}(f/\mathbf{T})$ řešitelná.*

Podle Tvrzení 24.5 se Galoisova grupa polynomu stupně n vnořuje do grupy $\mathbf{S}_n$. Existence vzorců na řešení polynomiálních rovnic stupně n se tak dostává do přímé souvislosti s řešitelností grupy $\mathbf{S}_n$.

Důsledek 25.2 (Al Chvárizmí, Tartaglia, Ferrari)**.** *Všechny polynomy stupně ≤ 4 jsou řešitelné v radikálech.*

Důkaz. Buď f polynom stupně n. Podle Tvrzení 24.5(1) je $\mathbf{Gal}(f/\mathbf{T})$ podgrupou grupy $\mathbf{S}_n$. V sekci 19.3 jsme ukázali, že grupy $\mathbf{S}_n$, $n \leq 4$, i jejich podgrupy (Tvrzení 19.8(1)) jsou řešitelné. Z Galoisovy věty tedy plyne, že polynomy stupně ≤ 4 jsou řešitelné v radikálech. □

Důsledek 25.3 (Abel-Ruffiniho věta)**.** *Existují racionální polynomy stupně 5 a více, které nejsou řešitelné v radikálech nad tělesem $\mathbb{Q}$.*

Důkaz. V sekci 19.3 jsme si řekli, že grupy $\mathbf{S}_n$, $n \geq 5$, nejsou řešitelné. Podle Tvrzení 24.8 existuje polynom stupně 5, jehož Galoisova grupa je $\mathbf{S}_5$. □

V této učebnici dokážeme pouze jednu implikaci Galoisovy věty: tu, ze které plyne neexistence vzorců. Opačná implikace je složitější, k důkazu Abel-Ruffiniho věty není potřeba a existenci Cardanových vzorců jsme ukázali explicitně.

Idea důkazu je následující: pro řešitelný polynom f vezmeme rozšíření

$$\mathbb{Q} = \mathbf{T}_0 \leq \mathbf{T}_1 \leq \ldots \leq \mathbf{T}_k$$

taková, že $\mathbf{T}_i$ je rozkladové nadtěleso nějakého polynomu $x^{n_i} - a_i \in T_{i-1}[x]$ nad tělesem $\mathbf{T}_{i-1}$, a rozkladové nadtěleso polynomu f je obsaženo v $\mathbf{T}_k$. Za jistých okolností bude takové řadě odpovídat řada normálních podgrup

$$\mathbf{Gal}(\mathbf{T}_k/\mathbb{Q}) = \mathbf{Gal}(\mathbf{T}_k/\mathbf{T}_0) \geq \mathbf{Gal}(\mathbf{T}_k/\mathbf{T}_1) \geq \ldots \geq \mathbf{Gal}(\mathbf{T}_k/\mathbf{T}_k) = \{id\},$$

přičemž faktorgrupy $\mathbf{Gal}(\mathbf{T}_k/\mathbf{T}_i) \,/\, \mathbf{Gal}(\mathbf{T}_k/\mathbf{T}_{i+1})$ budou izomorfní $\mathbf{Gal}(x^{n_i} - a_i/\mathbf{T}_i)$, a tedy řešitelné podle Tvrzení 24.7. Potom Důsledek 19.9 zaručí, že celá grupa $\mathbf{Gal}(\mathbf{T}_k/\mathbb{Q})$ je řešitelná a pomocí Tvrzení 24.5(3) se ukáže řešitelnost i pro Galoisovu grupu rozkladového nadtělesa polynomu f, které je obsaženo v $\mathbf{T}_k$.

Aby tento postup fungoval, tělesa $\mathbf{T}_1, \ldots, \mathbf{T}_k$ by musela být rozkladová pro nějaké polynomy nad $\mathbf{T}$. To obecně pravda není, v důkazu tedy budeme konstruovat posloupnost o trochu větších těles, která tuto vlastnost mají a přitom Galoisova grupa největšího tělesa zůstává řešitelná.

Lemma 25.4. *Buď $\mathbf{S}$ rozkladové nadtěleso nějakého polynomu nad tělesem $\mathbf{T}$ a buď g ireducibilní polynom v $\mathbf{T}[x]$. Pokud má polynom g v tělese $\mathbf{S}$ nějaký kořen, pak se v $\mathbf{S}[x]$ rozkládá na lineární činitele.*

Důkaz. Označme f polynom, pro nějž je $\mathbf{S}$ rozkladovým nadtělesem, a uvažujme rozkladové nadtěleso $\mathbf{U}$ pro polynom fg nad $\mathbf{T}$. Označme a kořen polynomu g v tělese $\mathbf{S}$ a uvažujme jakýkoliv jiný kořen b tohoto polynomu v $\mathbf{U}$. Chceme dokázat, že b leží v $\mathbf{S}$. Podle Lemmatu 24.3 existuje $\mathbf{T}$-izomorfismus $\mathbf{T}(a) \to \mathbf{T}(b)$ zobrazující $a \mapsto b$, a ten se podle Lemmatu 24.4 rozšiřuje do $\mathbf{T}$-izomorfismu $\varphi : \mathbf{U} \to \mathbf{U}$, tj. prvku $\mathbf{Gal}(\mathbf{U}/\mathbf{T})$, který splňuje $\varphi(a) = b$. Podle Tvrzení 24.1 zobrazení φ permutuje kořeny polynomu f, ty generují těleso $\mathbf{S}$, a tedy $\varphi(S) \subseteq S$. Speciálně dostáváme, že $b = \varphi(a) \in S$. $\square$

Lemma 25.5. *Buď $\mathbf{T}$ těleso charakteristiky 0 a $\mathbf{T} \leq \mathbf{S} \leq \mathbf{U}$ rozšíření těles taková, že $\mathbf{S}$ je rozkladové nadtěleso nějakého polynomu nad $\mathbf{T}$ a $\mathbf{U}$ je rozkladové nadtěleso polynomu $x^n - a \in S[x]$ nad $\mathbf{S}$. Pak existuje rozšíření $\mathbf{U} \leq \mathbf{V}$ takové, že $\mathbf{V}$ je rozkladové nadtěleso nějakého polynomu nad $\mathbf{T}$ a $\mathbf{Gal}(\mathbf{V}/\mathbf{S})$ je řešitelná grupa.*

Poznamenejme, že kdyby bylo samo $\mathbf{U}$ rozkladovým nadtělesem nějakého polynomu nad $\mathbf{T}$, pak bychom mohli volit $\mathbf{V} = \mathbf{U}$ a řešitelnost by zajistilo Tvrzení 24.7.

Důkaz. Bez újmy na obecnosti můžeme předpokládat, že $\mathbf{U} \leq \mathbb{C}$ (rozkladová nadtělesa jsou izomorfní a jedno lze najít v $\mathbb{C}$). Označme f polynom, pro nějž je $\mathbf{S}$ rozkladovým nadtělesem. Definujeme polynom

$$g = m_{a,\mathbf{T}}(x^n) \in T[x]$$

(do minimálního polynomu $m_{a,\mathbf{T}}$ dosadíme mocninu proměnné x) a uvažujme rozkladové nadtěleso $\mathbf{V} \leq \mathbb{C}$ polynomu $fg \in T[x]$ nad tělesem $\mathbf{T}$.

Nejprve si všimneme, že $\mathbf{U} \leq \mathbf{V}$: v oboru $\mathbf{S}[x]$ platí $x - a \mid m_{a,\mathbf{T}}$, tedy také $x^n - a \mid m_{a,\mathbf{T}}(x^n) = g$, takže se polynom $x^n - a$ rozkládá ve $\mathbf{V}[x]$ na lineární činitele. Dokážeme, že $\mathbf{Gal}(\mathbf{V}/\mathbf{S})$ je řešitelná grupa.

Polynom $m_{a,\mathbf{T}}$ je ireducibilní, má kořen v tělese $\mathbf{S}$, a tedy se tam podle Lemmatu 25.4 rozkládá na lineární činitele. Označme tento rozklad $m_{a,\mathbf{T}} = (x - a_1) \cdots (x - a_m)$. Pak

$$g = m_{a,\mathbf{T}}(x^n) = (x^n - a_1) \cdots (x^n - a_m).$$

Definujeme sekvenci

$$\mathbf{S} = \mathbf{S}_0 \leq \mathbf{S}_1 \leq \ldots \leq \mathbf{S}_{m-1} \leq \mathbf{S}_m = \mathbf{V},$$

kde $\mathbf{S}_i$ je rozkladovým nadtělesem polynomu $x^n - a_i$ nad $\mathbf{S}_{i-1}$, čili také rozkladovým nadtělesem polynomu $(x^n - a_1)\cdots(x^n - a_i)$ nad $\mathbf{S}$, pro každé $i = 1, \dots, m$. Uvažujme řadu podgrup

$$\mathbf{Gal}(\mathbf{V}/\mathbf{S}) = \mathbf{Gal}(\mathbf{V}/\mathbf{S}_0) \geq \mathbf{Gal}(\mathbf{V}/\mathbf{S}_1) \geq \dots \geq \mathbf{Gal}(\mathbf{V}/\mathbf{S}_m) = \{id\}.$$

Protože jsou všechna mezitělesa $\mathbf{S}_i$ rozkladová nad $\mathbf{S}$, můžeme aplikovat Tvrzení 24.5(3). Aplikací na rozšíření $\mathbf{S} \leq \mathbf{S}_i \leq \mathbf{V}$ vidíme, že $\mathbf{Gal}(\mathbf{V}/\mathbf{S}_i) \trianglelefteq \mathbf{Gal}(\mathbf{V}/\mathbf{S})$. Aplikací na rozšíření $\mathbf{S} \leq \mathbf{S}_{i-1} \leq \mathbf{S}_i$ vidíme, že

$$\mathbf{Gal}(\mathbf{V}/\mathbf{S}_i) \big/ \mathbf{Gal}(\mathbf{V}/\mathbf{S}_{i+1}) \simeq \mathbf{Gal}(\mathbf{S}_{i+1}/\mathbf{S}_i),$$

přičemž tyto faktorgrupy jsou řešitelné podle Tvrzení 24.7, protože $\mathbf{S}_i$ je rozkladovým nadtělesem polynomu $x^n - a_i$ nad tělesem $\mathbf{S}_{i-1}$. Důsledek 19.9 zaručí, že celá grupa $\mathbf{Gal}(\mathbf{V}/\mathbf{S})$ je řešitelná. □

Důkaz Galoisovy věty 25.1, část ($\Rightarrow$).

Buď f polynom řešitelný v radikálech a uvažujme řadu rozšíření prokazující tento fakt, tj. mějme $\mathbf{T} = \mathbf{T}_0 \leq \mathbf{T}_1 \leq \dots \leq \mathbf{T}_k$ taková, že $\mathbf{T}_i$ je rozkladové nadtěleso nějakého polynomu $x^{n_i} - a_i \in T_{i-1}[x]$ nad tělesem $\mathbf{T}_{i-1}$ a rozkladové nadtěleso $\mathbf{W}$ polynomu f nad $\mathbf{T}$ je obsaženo v tělese $\mathbf{T}_k$. Dokážeme, že grupa $\mathbf{Gal}(f/\mathbf{T}) = \mathbf{Gal}(\mathbf{W}/\mathbf{T})$ je řešitelná.

Postavíme řadu rozšíření

$$\mathbf{T} = \mathbf{U}_0 = \mathbf{V}_0 \leq \mathbf{U}_1 \leq \mathbf{V}_1 \leq \dots \leq \mathbf{U}_k \leq \mathbf{V}_k$$

tak, že pro $i = 1, \dots, m$ vezmeme $\mathbf{U}_i$ rozkladové nadtěleso polynomu $x^{n_i} - a_i$ nad tělesem $\mathbf{V}_{i-1}$ a vezmeme $\mathbf{V}_i$ jako těleso $\mathbf{V}$ z Lemmatu 25.5 aplikovaného na rozšíření $\mathbf{T} \leq \mathbf{V}_{i-1} \leq \mathbf{U}_i$. Čili každé $\mathbf{V}_i$ je rozkladové nadtěleso nad $\mathbf{T}$ a grupa $\mathbf{Gal}(\mathbf{V}_i/\mathbf{V}_{i-1})$ je řešitelná.

Zbytek důkazu je podobný jako v předchozím lemmatu. Z řady rozšíření $\mathbf{T} = \mathbf{V}_0 \leq \mathbf{V}_1 \leq \dots \leq \mathbf{V}_k$ získáme řadu podgrup

$$\mathbf{Gal}(\mathbf{V}_k/\mathbf{T}) = \mathbf{Gal}(\mathbf{V}_k/\mathbf{V}_0) \geq \mathbf{Gal}(\mathbf{V}_k/\mathbf{V}_1) \geq \dots \geq \mathbf{Gal}(\mathbf{V}_k/\mathbf{V}_k) = \{id\}.$$

Protože jde o rozkladová nadtělesa nad $\mathbf{T}$, můžeme aplikovat Tvrzení 24.5(3). Aplikací na rozšíření $\mathbf{T} \leq \mathbf{V}_i \leq \mathbf{V}_k$ vidíme, že $\mathbf{Gal}(\mathbf{V}/\mathbf{V}_i) \trianglelefteq \mathbf{Gal}(\mathbf{V}/\mathbf{T})$. Aplikací na rozšíření $\mathbf{T} \leq \mathbf{V}_{i-1} \leq \mathbf{V}_i$ vidíme, že

$$\mathbf{Gal}(\mathbf{V}_k/\mathbf{V}_i) \big/ \mathbf{Gal}(\mathbf{V}_k/\mathbf{V}_{i+1}) \simeq \mathbf{Gal}(\mathbf{V}_{i+1}/\mathbf{V}_i),$$

což již víme, že jsou řešitelné grupy. Důsledek 19.9 zaručí, že celá grupa $\mathbf{Gal}(\mathbf{V}_k/\mathbf{T})$ je řešitelná.

Zbývá dokázat, že grupa $\mathbf{Gal}(\mathbf{W}/\mathbf{T})$ je také řešitelná. Znovu použijeme Tvrzení 24.5(3) na rozšíření $\mathbf{T} \leq \mathbf{W} \leq \mathbf{V}_k$ a vidíme, že

$$\mathbf{Gal}(\mathbf{W}/\mathbf{T}) \simeq \mathbf{Gal}(\mathbf{V}_k/\mathbf{T}) \,/\, \mathbf{Gal}(\mathbf{V}_k/\mathbf{W}).$$

Nyní stačí použít Tvrzení 19.8, které říká, že faktorgrupa řešitelné grupy $\mathbf{Gal}(\mathbf{V}_k/\mathbf{T})$ je také řešitelná. □

Galoisova teorie dále rozvíjí vztah mezi podtělesy daného rozkladového tělesa a podgrupami příslušné Galoisovy grupy: mezi nimi existuje vzájemně jednoznačná korespondence přenášející řadu důležitých vlastností (například stupeň rozšíření souvisí s indexem podgrupy). Její výklad najdete ve většině učebnic komutativní algebry. Také bychom mohli zmínit, že předpoklad charakteristiky 0 je zbytečně silný, většina Galoisovy teorie platí i pro konečná tělesa a obecně všechna rozšíření, která jsou tzv. *separabilní.*

26. Konečná tělesa

Aplikací Vět 9.6 a 24.2 o existenci a jednoznačnosti rozkladových nadtěles ukážeme, že pro každou mocninu prvočísla p^k existuje, až na izomorfismus, právě jedno těleso velikosti p^k. Princip důkazu je v tom, že těleso má přesně p^k prvků právě tehdy, když je rozkladovým nadtělesem polynomu $x^{p^k} - x$ nad tělesem $\mathbb{Z}_p$. Z existence a jednoznačnosti rozkladových nadtěles pak plyne existence a jednoznačnost konečných těles.

Lemma 26.1. *Rozkladové nadtěleso polynomu $x^{p^k} - x$ nad tělesem $\mathbb{Z}_p$ má právě p^k prvků.*

Důkaz. Označme $q = p^k$. Buď $\mathbf{T}$ rozkladové nadtěleso polynomu $f = x^q - x$ nad $\mathbb{Z}_p$. Ukážeme, že kořeny f tvoří v $\mathbf{T}$ podtěleso. Tvrzení 20.4 o Frobeniově endomorfismu říká, že zobrazení $\varphi : a \mapsto a^p$ je homomorfismem $\mathbf{T} \to \mathbf{T}$. Jeho k-násobné složení, φ^k, je také homomorfismem a zobrazuje $a \mapsto (((a^p)^p)\dots)^p = a^{p^k} = a^q$, čili

$$(a+b)^q = a^q + b^q \quad \text{a} \quad (a \cdot b)^q = a^q \cdot b^q$$

pro každé $a, b \in T$. Tedy, jsou-li a, b kořeny polynomu f, tj. $a^q = a$ a $b^q = b$, pak $(a+b)^q = a^q + b^q = a + b$ je také kořen f a stejně tak $(a \cdot b)^q = a^q \cdot b^q = a \cdot b$, $(-a)^q = -a^q = -a$ a $(a^{-1})^q = (a^q)^{-1} = a^{-1}$. Čili kořeny tvoří podtěleso. Z požadavku minimality pak plyne, že rozkladové nadtěleso $\mathbf{T}$ sestává právě z kořenů f, a tedy má *nejvýše* $\deg f = q$ prvků.

Abychom dokázali, že $\mathbf{T}$ má *přesně* q prvků, stačí ověřit, že polynom f nemá vícenásobné kořeny. Kdyby byl prvek a vícenásobným kořenem, podle Věty 2.7 by platilo $f'(a) = 0$. Ovšem $f' = qx^{q-1} - 1 = -1$, a tedy žádné kořeny nemá. □

Lemma 26.2. *Buď $\mathbf{T}$ konečné těleso, $|T| = p^k$. Pak $\mathbf{T}$ je rozkladovým nadtělesem polynomu $x^{p^k} - x$ nad tělesem $\mathbb{Z}_p$ a v $\mathbf{T}[x]$ platí*

$$x^{p^k} - x = \prod_{a \in T} (x - a).$$

Důkaz. Označme $q = p^k$. Nejprve si všimněte, že každý prvek $a \in T$ je kořenem polynomu $f = x^q - x$. Pro 0 to platí triviálně a pro nenulový prvek a využijeme Lagrangeovu větu (Věta 14.9): $\operatorname{ord}(a) \mid |T^*| = q - 1$, čili $a^{q-1} = 1$ a $a^q = a$. Tedy $\prod_{a \in T}(x - a) \mid f$, a z rovnosti stupňů i vedoucích koeficientů dostáváme rovnost těchto polynomů. Podle předchozího lemmatu má rozkladové nadtěleso polynomu f právě q prvků, čili $\mathbf{T}$ je tímto tělesem. □

Věta 26.3 (klasifikace konečných těles)**.**

(1) *Konečné těleso velikosti n existuje právě tehdy, když $n = p^k$ pro nějaké prvočíslo p a přirozené číslo k.*
(2) *Konečná tělesa stejné velikosti jsou izomorfní.*

Důkaz. (1) ($\Rightarrow$) plyne z Tvrzení 21.1, ($\Leftarrow$) plyne z Lemmatu 26.1 a Věty 9.6 o existenci rozkladových nadtěles. (2) plyne z Lemmatu 26.2 a Věty 24.2 o jednoznačnosti rozkladových nadtěles. □

V sekci 10.1 jsme představili konečná tělesa ve formě faktorokruhů $\mathbb{Z}_p[\alpha]/(m)$. Lze každé konečné těleso tímto způsobem reprezentovat?

Věta 26.4 (reprezentace konečných těles)**.** *Pro každé prvočíslo p a přirozené číslo k existuje ireducibilní polynom $m \in \mathbb{Z}_p[\alpha]$ stupně k a*

$$\mathbb{F}_{p^k} \simeq \mathbb{Z}_p[\alpha]/(m).$$

Důkaz. Podle Věty 26.3 existuje nějaké těleso $\mathbf{T} \geq \mathbb{Z}_p$ velikosti p^k. Podle Věty 16.7 existuje nějaký generátor a cyklické grupy $\mathbf{T}^*$. Uvažujme minimální polynom $m_{a,\mathbb{Z}_p}$. Ten je jistě ireducibilní a jeho stupeň je

$$\deg m_{a,\mathbb{Z}_p} = [\mathbb{Z}_p(a) : \mathbb{Z}_p] = [\mathbf{T} : \mathbb{Z}_p] = k,$$

přičemž první rovnost plyne z Tvrzení 22.3, druhá z faktu, že $\mathbf{T} = \mathbb{Z}_p(a)$, protože $\mathbf{T}$ sestává z mocnin prvku a, a třetí z toho, že vektorový prostor s p^k prvky má dimenzi k. Z jednoznačnosti ve Větě 26.3 plyne, že $\mathbf{T} \simeq \mathbb{Z}_p[\alpha]/(m_{a,\mathbb{Z}_p})$. □

Všimněte si, jakým obratem jsme prokázali existenci ireducibilního polynomu stupně k v $\mathbb{Z}_p[x]$: nejprve jsme prokázali existenci nějakého tělesa velikosti p^k, abychom mohli vzít generátor jeho multiplikativní grupy a jeho minimální polynom. Přímý důkaz existence těchto polynomů je možný, ale mnohem techničtější a dává menší vhled do celé situace.

Obsah