

UČEBNÍ TEXT
ALGEBRA 2 PRO INFORMATIKY

DAVID STANOVSKÝ
DROBNÉ ZMĚNY: ZUZANA PATÁKOVÁ
stanovsk@karlin.mff.cuni.cz

Obsah

IV. Homomorfismy	4
1. Grupové homomorfismy	4
1.1. Základní vlastnosti	4
1.2. Izomorfismus	6
1.3. Neizomorfismus	7
1.4. Klasifikace vztahů	8
2. Faktorgrupy	9
2.1. Normální podgrupy	9
2.2. Konstrukce faktorgrupy	10
3. Ideály a dělitelnost	14
3.1. Ideály	14
3.2. Obory hlavních ideálů	15
4. Okružové homomorfismy a faktorokruhy	17
4.1. Homomorfismy	17
4.2. Izomorfismus	18
4.3. Konstrukce faktorokruhu podle ideálu	20
4.4. Faktorokruhy podle maximálních ideálů a prvoideálů	22
V. Číselná tělesa a kořeny polynomů	24
5. Okružová a tělesová rozšíření	24
5.1. Dělení	24
5.2. Tělesové rozšíření jako vektorový prostor	26
6. Algebraické prvky a rozšíření konečného stupně	27
6.1. Algebraická a transcendentní čísla	27
6.2. Minimální polynom a stupeň jednoduchého rozšíření	28
6.3. Vícenásobná rozšíření	31
7. Neoddělitelnost úloh pravítkem a kružítkem	32
8. Izomorfismy konečných a rozkladových nadtěles	35
9. Klasifikace konečných těles	37
9.1. Frobeniovův endomorfismus	37
9.2. Derivace a násobné kořeny	38
9.3. Klasifikace konečných těles	38
VI. Algoritmy polynomiální aritmetiky	40
10. Modulární reprezentace	40
10.1. Diskrétní Fourierova transformace	40
10.2. Rychlá Fourierova transformace	42
10.3. Primitivní odmocniny z jedné	44
11. Rychlé násobení a dělení polynomů	45
11.1. Rychlé násobení	45
11.2. Rychlé dělení polynomů	47
11.3. Výpočet inverzní mocninné řady	49
12. Rozklady polynomů nad konečnými tělesy	50
12.1. Bezètercová faktorizace	50
12.2. Berlekampův algoritmus	55
VII. Další třídy algebraických struktur	59
13. Obecné algebraické struktury	59
13.1. Algebraické struktury	59
13.2. Podstruktury	60
13.3. Homomorfismy a izomorfismus	61
13.4. Kongruence a faktorstruktury	62

14. Uspořádání a svazy	64
14.1. Uspořádané množiny	64
14.2. Svazy a Booleovy algebry	67

Homomorfismy

1. Grupové homomorfismy

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, v teorii grup to budou zobrazení zachovávající základní grupové operace.

Jak si ukážeme, homomorfismy přenášejí řadu dalších vlastností, např. v obou grupách panuje úzký vztah mezi podgrupami a jejich prvky.

1.1. Základní vlastnosti.

V celé sekci budeme uvažovat dvě grupy $\mathbf{G} = (G, \cdot, {}^{-1}, 1)$ a $\mathbf{H} = (H, \cdot, {}^{-1}, e)$.

Definice. Buď \mathbf{G}, \mathbf{H} grupy. Zobrazení $\varphi : \mathbf{G} \rightarrow \mathbf{H}$ je homomorfismus směr těchto grup, pokud pro každé $a, b \in G$ platí

$$\varphi(ab) = \varphi(a)\varphi(b), \quad \varphi(a^{-1}) = \varphi(a)^{-1}, \quad \varphi(1) = e.$$

Fakt, že je zobrazení mezi grupami homomorfismus, budeme zapisovat $\varphi : \mathbf{G} \rightarrow \mathbf{H}$.

Hned na začátku je dobré si všimnout, že druhá a třetí rovnost plynou z té první, což zřetelně zjednodušuje ověřování, zda je dané zobrazení homomorfismus.

Lemma 1.1. Buď \mathbf{G}, \mathbf{H} grupy a $\varphi : \mathbf{G} \rightarrow \mathbf{H}$ zobrazení. Pak φ je homomorfismus směr těchto grup právě tehdy, když $\varphi(ab) = \varphi(a)\varphi(b)$ 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 $\varphi(a^{-1}) = \varphi(a)^{-1}$ 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 $\varphi(a^{-1}) = \varphi(a)^{-1}$.

Buď $\varphi : \mathbf{G} \rightarrow \mathbf{H}$ homomorfismus. Jeho obrazem nazýváme jeho obor hodnot, tj. množinu

$$\text{Im}(\varphi) = \{\varphi(a) : a \in G\}.$$

Jeho jádro definujeme jako množinu

$$\text{Ker}(\varphi) = \{a \in G : \varphi(a) = e\}.$$

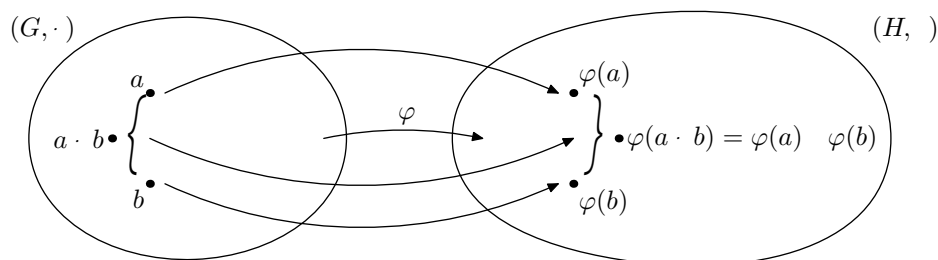
Tvrzení 1.2 (jádro a obraz jsou podgrupy). Buď \mathbf{G}, \mathbf{H} grupy a $\varphi : \mathbf{G} \rightarrow \mathbf{H}$ homomorfismus. Pak

- (1) $\text{Im}(\varphi)$ tvoří podgrupu grupy \mathbf{H} ;
- (2) $\text{Ker}(\varphi)$ tvoří podgrupu grupy \mathbf{G} .

Důkaz. (1) $e \in \text{Im}(\varphi)$, protože $e = \varphi(1)$. Pokud $\varphi(a), \varphi(b) \in \text{Im}(\varphi)$, pak $\varphi(a)^{-1} = \varphi(a^{-1}) \in \text{Im}(\varphi)$ a $\varphi(a)\varphi(b) = \varphi(ab) \in \text{Im}(\varphi)$.

(2) $1 \in \text{Ker}(\varphi)$, protože $\varphi(1) = e$. Pokud $a, b \in \text{Ker}(\varphi)$, pak a^{-1} a ab také, protože $\varphi(a^{-1}) = \varphi(a)^{-1} = e^{-1} = e$ a $\varphi(ab) = \varphi(a)\varphi(b) = e \cdot e = e$.

Tvrzení 1.3. Buď \mathbf{G}, \mathbf{H} grupy a $\varphi : \mathbf{G} \rightarrow \mathbf{H}$ homomorfismus. Pak φ je prostý právě tehdy, když $\text{Ker}(\varphi) = \{1\}$.

Obrázek 1. Homomorfismus $\varphi : G \rightarrow H$.

Důkaz. Je-li φ prostý, dva různé prvky se nemohou zobrazovat na e , takže $\text{Ker}(\varphi)$ musí obsahovat jen jeden prvek, a tím je 1. Naopak, $\varphi(a) = \varphi(b)$ právě tehdy, když $e = \varphi(a) \varphi(b)^{-1} = \varphi(a b^{-1})$, takže neprostá zobrazení obsahují nejednotkový prvek v jádru.

Příklad. Řada známých zobrazení v matematice je homomorfismem určitých grup.

Uvažujme zobrazení $z \mapsto |z|$ na komplexních číslech. Toto zobrazení je homomorfismem grup $\mathbb{C} \rightarrow \mathbb{R}$, protože $|ja + bj| = |ja| |bj|$. Jeho jádrem je podgrupa komplexních jednotek, jeho obrazem podgrupa kladných čísel. Naopak, toto zobrazení není homomorfismem grup $\mathbb{C} \rightarrow \mathbb{R}$, protože obecně $|ja + bj| \neq |ja| + |bj|$.

Uvažujme zobrazení $z \mapsto e^z$ na komplexních číslech. Toto zobrazení je homomorfismem grup $\mathbb{C} \rightarrow \mathbb{C}$, protože $e^{a+b} = e^a e^b$. Jeho jádrem je podgrupa $i\mathbb{Z} = \{2\pi i k : k \in \mathbb{Z}\}$, jeho obrazem celé \mathbb{C} .

Uvažujme zobrazení $A \mapsto \det(A)$ na maticích. Toto zobrazení je homomorfismem grup $\text{GL}_n(\mathbb{T}) \rightarrow \mathbb{T}$, protože $\det(AB) = \det(A) \det(B)$. Jeho jádrem je podgrupa $\text{SL}_n(\mathbb{T})$, jeho obrazem celé \mathbb{T} .

Uvažujme zobrazení $\pi \mapsto \text{sgn}(\pi)$ na permutacích. Toto zobrazení je homomorfismem grup $\text{S}_n \rightarrow \mathbb{Z}$, protože $\text{sgn}(\pi \sigma) = \text{sgn}(\pi) \text{sgn}(\sigma)$. Jeho jádrem je podgrupa A_n , jeho obrazem celé \mathbb{Z} .

Příklad. Je-li G grupa a $a \in G$ řádu n , pak diskretní exponenciála $Z_n \rightarrow G$, $k \mapsto a^k$, je prostý homomorfismus. Obrazem je podgrupa $\langle a \rangle$.

Příklad. Působení grupy G na množině X není nic jiného než homomorfismus $G \rightarrow \text{S}_X$, srovnajte obě dvě věci!

Homomorfismy jsou určeny svými hodnotami na generátorech: uvažujme homomorfismus $\varphi : G \rightarrow H$, nechť $G = \langle X \rangle$ a označme hodnoty $\varphi(a) = h_a$ pro všechna $a \in X$. Obecný prvek grupy G lze napsat ve tvaru $g = a_1^{k_1} \dots a_n^{k_n}$, kde $a_1, \dots, a_n \in X$ a $k_1, \dots, k_n \in \mathbb{Z}$. Hodnota zobrazení pak bude

$$\varphi(g) = \varphi(a_1)^{k_1} \dots \varphi(a_n)^{k_n} = h_{a_1}^{k_1} \dots h_{a_n}^{k_n}.$$

Avšak pozor, na rozdíl od vektorových prostorů není možné volit obrazy generátorů libovolně, jak ukazuje například následující tvrzení.

Tvrzení 1.4 (řád prvku a jeho obrazu). Buď $\varphi : G \rightarrow H$ homomorfismus grup. Pak, pro každé $a \in G$,

$$\text{ord}(\varphi(a)) \mid \text{ord}(a).$$

Je-li navíc φ prostý, pak

$$\text{ord}(\varphi(a)) = \text{ord}(a).$$

Důkaz. Označme $\text{ord}(a) = n$. Pak $\varphi(a)^n = \varphi(a^n) = \varphi(1) = e$, èli nutnì $\varphi(a)^k = e$ pro nìjaké $k \neq n$. Je-li navíc φ prostý, pro vèchna $k < n$ musí platit $\varphi(a)^k = \varphi(a^k) \neq e$, protože $a^k \neq 1$.

Úloha. Popište vèchny homomor smy $Z_{10} \rightarrow S_3$.

Øeñení. Grupa Z_{10} je cyklická, $Z_{10} = \langle 1 \rangle$, èli stačí urèit pøípušné hodnoty $\varphi(1)$: potom $\varphi(k) = \varphi(1 + \dots + 1) = \varphi(1) \dots \varphi(1) = \varphi(1)^k$. Øád prvku 1 v Z_{10} je 10, èli øád prvku $\varphi(1)$ v S_3 musí èíslo 10 dìlit. Avšak v S_3 jsou pouze prvky øádu 1, 2, 3, èli máme nejvýše ètyøí možnosti: $\varphi(1) \in \{id, (1\ 2), (1\ 3), (2\ 3)\}$. Je snadné ovìdit, že vèchna ètyøí zobrazení $k \mapsto id$ a $k \mapsto (i\ j)^k$ jsou homomor smy $Z_{10} \rightarrow S_3$. (Rozmyslete si, proč zobrazení $k \mapsto (1\ 2\ 3)^k$ nespløuje de nici homomor smu, bez odkazu na Tvzení 1.4.)

Tvrzení 1.5. Buò G, H, K grupy a $\varphi : G \rightarrow H$, $\psi : H \rightarrow K$ homomor smy. Pak

- (1) $\psi \circ \varphi$ je homomor smus $G \rightarrow K$,
- (2) je-li φ bijektivní, pak φ^{-1} je homomor smus $H \rightarrow G$.

Důkaz. (1) Oznaème $K = (K, +, \cdot, 0)$. Pro $a, b \in G$ platí

$$(\psi \circ \varphi)(a \cdot b) = \psi(\varphi(a \cdot b)) = \psi(\varphi(a) \cdot \varphi(b)) = \psi(\varphi(a)) \cdot \psi(\varphi(b)) = (\psi \circ \varphi)(a) \cdot (\psi \circ \varphi)(b)$$

postupným použitím faktu, že φ a ψ jsou homomor smy.

(2) Napišme $u, v \in H$ jako $u = \varphi(a)$ a $v = \varphi(b)$ pro jistá $a, b \in G$. Pak

$$\varphi^{-1}(u \cdot v) = \varphi^{-1}(\varphi(a) \cdot \varphi(b)) = \varphi^{-1}(\varphi(a \cdot b)) = a \cdot b = \varphi^{-1}(u) \cdot \varphi^{-1}(v)$$

použitím faktu, že φ je homomor smus a $\varphi^{-1} \circ \varphi = id$.

1.2. Izomorfismus.

Definice. Bijektivní homomor smy nazýváme izomor smy.

Z Tvzení 1.5 ihned plyne, že složení izomor smù je izomor smus a inverzní zobrazení k izomor smu je také izomor smus.

Na izomor smus je možné pohlížet jako na økopírování algebraické struktury: máme-li grupu G a bijektivní zobrazení $\varphi : G \rightarrow H$, můžeme na množinu H økopírovat grupové operace øedpisem

$$e = \varphi(1), \quad a^0 = \varphi((\varphi^{-1}(a))^{-1}), \quad a \cdot b = \varphi(\varphi^{-1}(a) \cdot \varphi^{-1}(b)).$$

Vidíme, že zobrazení φ^{-1} bude izomor smem mezi novou grupou $H = (H, \cdot^0, e)$ a starou grupou G . Jedna grupa je kopií druhé, došlo pouze k øejmenování prvkù kopírovacím zobrazením φ . Na každý izomor smus lze pohlížet tímto způsobem.

Dvì grupy G, H nazveme izomorfní, pokud existuje izomor smus $G \rightarrow H$, tento fakt znaèíme $G \cong H$. Neformálnì, jedna grupa je økopíjí druhé. Tvzení 1.5 implikuje, že izomor smus dává ekvivalenci na třídì vèch grup:

- reflexivita: $G \cong G$ je zaruèeno izomor smem $id : G \rightarrow G$;
- symetrie: je-li $G \cong H$ pomocí izomor smu φ , pak $H \cong G$ pomocí izomor smu φ^{-1} ;
- tranzitivita: je-li $G \cong H$ pomocí izomor smu φ a $H \cong K$ pomocí izomor smu ψ , pak $G \cong K$ pomocí izomor smu $\psi \circ \varphi$.

Na prosté homomor smy lze nahlížet jako na izomor smy mezi výchozí grupou a obrazem, tj. prostý homomor smus $\varphi : G \rightarrow H$ je izomor smem $G \rightarrow \text{Im}(\varphi)$. Takovým homomor smem se říká vnoøení grupy G do grupy H , tj. grupa H obsahuje izomorfní kopii G jako podgrupu.

Příklad. Grupy Z_2 a Z jsou izomorfní. Podívejme se na tabulky jejich operací:

$$\begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \quad \begin{array}{c|cc} & 1 & 1 \\ \hline 1 & 1 & 1 \\ 1 & 1 & 1 \end{array}$$

Tyto tabulky vypadají podobně: jedna je kopií druhé, pokud přepíšeme $0 \leftrightarrow 1, 1 \leftrightarrow 0$.

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 třeba ověřit, že zobrazení $a + bi \mapsto (a, b)$ je grupový izomorfismus $\mathbb{C} \rightarrow \mathbb{R} \times \mathbb{R}$.

Příklad. Grupy Z_n a $C_n = \langle \zeta_n \rangle$, 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 třeba ověřit, že zobrazení $k \mapsto \zeta_n^k$ je grupový izomorfismus $Z_n \rightarrow C_n$.

Příklad. Všechny tyto grupy Z_2, Z_4, Z_8 a $\text{id}, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3) \in S_4$ jsou navzájem izomorfní. Není to vidět na první pohled, ale intuice se dá vybudovat přes generátory: všechny tyto grupy lze napsat jako $G = \langle a, b \rangle$, kde $a^2 = 1, b^2 = 1$ a $ab = ba$ je ten třetí prvek různý od jednotky. Formální důkaz si udělejte jako cvičení.

1.3. Neizomorfismus.

Viděli jsme, že zobrazení $a + bi \mapsto (a, b)$ je izomorfismem grup $\mathbb{C} \rightarrow \mathbb{R} \times \mathbb{R}$, 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 $Z_{mn} \cong Z_m \times Z_n$. Jak je tomu pro m, n soudělná? Zobrazení $x \mapsto (x \bmod m, x \bmod n)$ není ani prosté, ani na: čísla 0 i $\text{NSN}(m, n)$ 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 $G \cong H$ platí, že pokud má grupa G vlastnost V , pak má i grupa H vlastnost V .

Příkladem invariantu je počet prvků daného řádu: je-li φ izomorfismus, podle Tvzení 1.4 je řád a a $\varphi(a)$ vždy stejný.

Příklad.

Grupa Z_{mn} obsahuje prvek řádu mn . Avšak v grupě $Z_m \times Z_n$ mají všechny prvky řád nejvýše $\text{NSN}(m, n)$. Ěli 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 $1, 2, \infty$. Ěli tyto grupy nemohou být izomorfní.

Kvaternionová grupa Q_8 i dihedralní grupa D_8 obsahují prvky řádů $1, 2, 4$. Avšak Q_8 obsahuje šest prvků řádu 4 , zatímco D_8 pouze dva, takže nemohou být izomorfní.

Jiným příkladem invariantu je minimální počet generátorů.

Tvrzení 1.6. Buď $\varphi : G \rightarrow H$ homomorfismus grup, který je na. Je-li $G = \langle X \rangle$, pak $H = \langle \varphi(X) \rangle$.

Důkaz. Prvek $b \in H$ napíšeme jako $b = \varphi(a)$ pro nějaký $a \in G$, prvek a vyjádříme v generátorech jako $a = u_1^{k_1} \dots u_n^{k_n}$, kde $u_i \in X$, a prvek b dostaneme jako $b = \varphi(a) = \varphi(u_1)^{k_1} \dots \varphi(u_n)^{k_n} \in \langle \varphi(X) \rangle$.

Na rozdíl od vektorových prostorů, v grupách mohou být minimální generující množiny různě velké, např. $Z = \langle 1 \rangle = \langle 2 \rangle = \langle 3 \rangle$. Invariantem je nejmenší počet prvků, který je potřeba k nagenarování dané grupy.

Příklad. Grupy \mathbb{Z} a $\mathbb{Z} \times \mathbb{Z}$ nejsou izomorfní, protože grupu $\mathbb{Z} \times \mathbb{Z}$ nelze nagenarovat jedním prvkem: podgrupa $\{h(a,b) \mid h(ka, kb) : k \in \mathbb{Z}\}$ obsahuje dvojici $(1,1)$ pouze pro $(a,b) = (1,1)$, ale ani jedna z těchto dvojic $\mathbb{Z} \times \mathbb{Z}$ nengeneruje. O něco složitější argument by prošel i pro úlohu $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \times \mathbb{Z}_n$ pro souditelná m, n .

Uvedené dva invarianty umožňují prokázat neizomorfnost i ve spoustě případů, ale ne ve všech. Příkladem je dvojice grup \mathbb{Q} a $\mathbb{Q}^+ = \{a \in \mathbb{Q} : a > 0\}$, 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 izomorfnost $\varphi : G \rightarrow H$ a předpokládejme, že tato vlastnost platí v grupě G . Prvek $u \in H$ napíšeme jako $u = \varphi(a)$, vezmeme $b \in G$ takové, že $a = b^2$ 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 $0 < a \in \mathbb{Q}$ existuje $0 < b \in \mathbb{Q}$ 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), podrobně to najdete v nějakých učebnicích matematické logiky. Ani tyto invarianty však nemusí pomoci. Příkladem jsou grupy \mathbb{Q} a $\mathbb{Q} \setminus \mathbb{Q}$, které nelze odlišit žádnou formulí prvního řádu, ale přesto nejsou izomorfní (viz cvičení).

1.4. Klasifikační věty.

Jedním ze základních cílů každé algebraické teorie je tzv. klasifikace objektů, tj. úplný seznam všech příkladů až na izomorfnost. 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 izomorfnost, všechny příklady cyklických grup.

Věta 1.7 (klasifikace cyklických grup). Buď G cyklická grupa.

- (1) Je-li G nekonečná, pak je izomorfní grupě \mathbb{Z} .
- (2) Je-li G konečná řádu n , pak je izomorfní grupě \mathbb{Z}_n .

Důkaz. Buď $G = \langle a \rangle$ cyklická grupa.

- (1) Předpokládejme, že je G nekonečná, tedy $\text{ord}(a) = \infty$, a uvažujme zobrazení

$$\mathbb{Z} \rightarrow G, \quad k \mapsto a^k.$$

Toto zobrazení je homomorfismus, nebo $a^k a^l = a^{k+l}$. Přitom jádro je triviální, protože $a^k \neq 1$ pro všechna $k \neq 0$, takže podle tvrzení 1.3 jde o prosté zobrazení. Z prvního semestru víme, že je toto zobrazení na G .

- (2) Předpokládejme, že je G řádu n , tedy $\text{ord}(a) = n$, a uvažujme zobrazení

$$\mathbb{Z}_n \rightarrow G, \quad k \mapsto a^k.$$

Toto zobrazení je homomorfismus, nebo $a^k a^l = a^{k+l} = a^{k+l \bmod n}$, 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 $a^{k+l \bmod n} = a^{k+l} a^{-n} = a^{k+l} 1^{-1} = a^{k+l}$. Podobně jako pro nekonečnou grupu dostáváme, že jádro je triviální a že jde o zobrazení na G .

Mnohem komplikovanější je klasifikace konečně generovaných abelovských grup, která říká, že každá abelovská grupa s konečnou množinou generátorů je izomorfní

direktnímu součinu konečnì mnoha cyklických grup. Navíc, použitím ěínské vĕty o zbytcích ve formì Tvzení 4.4, konečnì cyklické grupy staĕi uva%ovat pouze ěádu mocniny prvoěísla. Tyto komponenty jsou navíc jednoznačnì určenné (a% na pořadí), tj. volbou neizomorfních cyklických grup dostaneme neizomorfní direktní součiny.

Vĕta 1.8 (klasi kace konečných abelovských grup). Buĕ G konečnì generovaná abelovská grupa, $|G| > 1$. Pak existují $m, n \in \mathbb{N}$, prvoěísla p_1, \dots, p_m (ne nutnì po dvou ruzná) a pĕirozená ěísla k_1, \dots, k_m taková, %e

$$G \cong \mathbb{Z}^n \times \mathbb{Z}_{p_1}^{k_1} \times \mathbb{Z}_{p_2}^{k_2} \times \dots \times \mathbb{Z}_{p_m}^{k_m}.$$

ěísla m, n jsou určena jednoznačnì a ěísla $p_1^{k_1}, \dots, p_m^{k_m}$ jednoznačnì a% na pořadí.

Důkaz této vĕty je pomĕrnì zdouhavý, najdete jej v ka%dě učebnici teorie grup.

Přříklad. Podle Vĕty 1.8 je ka%dá ětyĕprvková abelovská grupa izomorfní buĕ grupĕ \mathbb{Z}_4 , nebo grupĕ $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Grupa \mathbb{Z}_5 je také ětyĕprvková. Vidíme, %e $\text{ord}(2) = 4$, takže $\mathbb{Z}_5 \cong \mathbb{Z}_4$.

Grupa \mathbb{Z}_8 je také ětyĕprvková. Vidíme, %e $\text{ord}(3) = \text{ord}(5) = \text{ord}(7) = 2$, takže $\mathbb{Z}_8 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Oblíbenou kratochvílí grupaou je hledání malých grup, což je svým způsobem také klasi kačnĕ vĕta. V současně době je znám seznam vĕch grup a% do velikosti $2047 = 2^{11} - 1$. Následující tabulka obsahuje klasi kaci vĕch grup ěádu n pro $n \leq 15$ a pro $n = p, 2p, p^2$, kde p je prvoěíslo.

n	grupy ěádu 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	$\mathbb{Z}_8, \mathbb{Z}_2 \times \mathbb{Z}_4, \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2, \mathbb{Z}_2 \times \mathbf{D}_8, \mathbf{Q}_8$
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	$\mathbb{Z}_3 \times \mathbb{Z}_5$

Připad $n = p$ je důsledkem Lagrangeovy vĕty: grupa prvoěíselné velikosti nemu%e mít vlastní podgrupy, takže musí být generovaná libovolným svým prvkem (kromĕ jednotky) a podle klasi kace cyklických grup musí být izomorfní \mathbb{Z}_p . Ostatní případy jsou znatelnĕ tĕ%í.

2. Faktorgrupy

2.1. Normální podgrupy.

Předmĕtem této sekce je velmi důležitá konstrukce faktorgrup. Jako parametr slou%í jistý typ podgrup, zvaných normální. S tímto pojmem se nyní seznámíme.

Tvzení 2.1 (ekvivalentní de nice normální podgrupy). Buĕ G grupa a H její podgrupa. Následující tvzení 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) \Rightarrow (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) \Rightarrow (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 Tvzení 2.1. Tento fakt značíme $H \triangleleft G$.

V abelovských grupách je každá podgrupa normální, obě ekvivalentní podmínky jsou triviálně splněny. Z triviálních důvodů platí také $g \in G$ a $G \triangleleft G$.

Příklad.

Podgrupa $SL_n(\mathbb{T})$ matic s determinanem 1 je normální v grupě $GL_n(\mathbb{T})$, jak plyne z podmínky (2) užitím součinného vzorce pro determinanty: $\det(AHA^{-1}) = (\det A)(\det H)(\det A)^{-1} = \det H$.

Podgrupa A_n sudých permutací je normální v grupě S_n , jak plyne ze součinného vzorce pro znaménko: $\text{sgn}(aha^{-1}) = (\text{sgn } a)(\text{sgn } h)(\text{sgn } a)^{-1} = \text{sgn } h$.

Podgrupa D_{2n} není normální v grupě S_n .

Na závěr uvedeme jedno důležité pozorování doplňující Tvzení 1.2.

Tvzení 2.2. Jádru homomorfismu je normální podgrupa.

Důkaz. Uvažujme $\varphi : G \rightarrow H$. V Tvzení 1.2 jsme dokázali, že $\text{Ker}(\varphi)$ je podgrupa. Normalita plyne z toho, že pro libovolné $a \in \text{Ker}(\varphi)$ a $u \in G$ platí $\varphi(uaa^{-1}) = \varphi(u)\varphi(a)\varphi(a)^{-1} = \varphi(u)\varphi(a)^0 = \varphi(u)$.

2.2. 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, ně které prvky splynou (při pohledu pouhým okem například hvězdy v jedné galaxii). To, co vidíme, je faktorobjekt původního objektu (světelné body na obloze). O trochu formálněji, ztotožníme podobné objekty (na obloze ty, které jsou příliš blízko). Co se přesně myslí relací podobnosti už závisí na konkrétním typu objektu.

Definice. Buď G grupa a N její normální podgrupa. Definujeme relaci na množině G předpisem

$$a \sim b \iff a^{-1}b \in N.$$

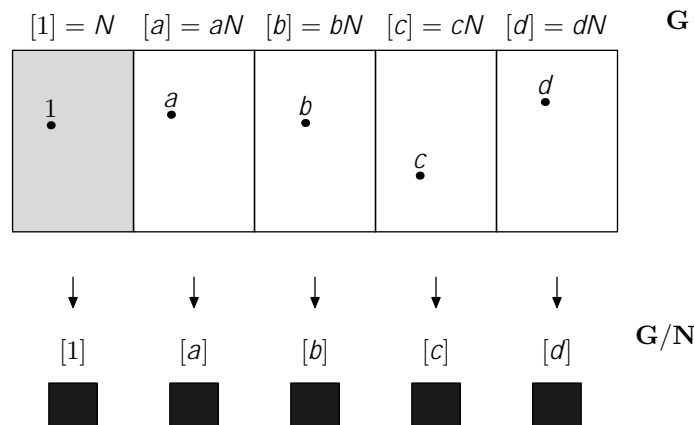
Podle jednoho z lemmat ze sekce o Lagrangeově větě je $a^{-1}b$ právě tehdy, když $Na = Nb$, ěli relace \sim je ekvivalencí na množině G . Její bloky jsou rozkladové třídy grupy G podle podgrupy N , a protože je N normální, levé i pravé rozkladové třídy jsou totéž (Tvzení 2.1), ěli

$$[a] = aN = Na.$$

Na těchto blocích definujeme operace předpisem

$$[a][b] = [ab] \quad a^{-1} \quad [a]^{-1} = [a^{-1}]$$

(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



Obrázek 2. Konstrukce faktorgrupy

vezmeme blok $[1] = N$. Množina bloků s výše uvedenými operacemi se nazývá faktorgrupa grupy \mathbf{G} podle podgrupy \mathbf{N} ,

$$\mathbf{G/N} = (\{[a] : a \in \mathbf{G}, [1], [1]\}.$$

Lemma 2.3. Buď \mathbf{G} grupa a \mathbf{N} její normální podgrupa.

- (1) Výše uvedené operace na blocích jsou dobře definovány.
- (2) Faktorgrupa $\mathbf{G/N}$ 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 $a = c$ a $b = d$, tj. $a = c^{-1} \cdot c$ a $b = d^{-1} \cdot d$, z uzavřenosti množiny \mathbf{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 $ab = cd$, tj. $[a \cdot b] = [c \cdot d]$. Pro inverz stačí využít faktu, že $ac^{-1} \in N$, $a^{-1}c \in N$, protože levé i pravé rozkladové třídy jsou stejné, a tedy $Na = Nc$, $aN = cN$.

(2) Ověříme, že $\mathbf{G/N}$ splňuje axiomy grup. Operace 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 $\mathbf{G} = \mathbf{Z}$ a normální podgrupu $\mathbf{H} = n\mathbf{Z}$. Platí

$$a \sim b \iff n \mid a - b \iff a \equiv b \pmod{n},$$

bloky této ekvivalence jsou rozkladové třídy

$$[a] = \{k \in \mathbf{Z} : k \equiv a \pmod{n}\} = a + n\mathbf{Z}, \quad a = 0, \dots, n-1.$$

Přitom $[a] + [b] = [a + b] = [a + b \pmod{n}]$ a $[a]^{-1} = [a^{-1}] = [n - a]$, čili operace na prvcích $\mathbf{Z}/n\mathbf{Z}$ jsou jako operace na číslech $0, \dots, n-1$ modulo n . Není těžké ověřit, že $[a] \cdot [b] = [a \cdot b]$ je izomorfní s $\mathbf{Z}/n\mathbf{Z} \cong \mathbf{Z}_n$.

Příklad. Uvažujme grupu $\mathbf{G} = \mathbf{S}_n$ a normální podgrupu $\mathbf{H} = \mathbf{A}_n$. Platí

$$\pi \sim \sigma \iff \pi \sigma^{-1} \in \mathbf{A}_n \iff \text{sgn}(\pi) = \text{sgn}(\sigma),$$

čili tato ekvivalence má právě dva bloky: množinu S sudých permutací a množinu L lichých permutací. Operace na těchto třídách je $S \cdot S = L$, $L \cdot S = S$ a $S \cdot L = L$, $L \cdot L = S$, jde o dvouprvkovou grupu izomorfní grupě \mathbf{Z}_2 .

Jak jednoduše, ale přitom formálně určit, jak vypadá faktorgrupa dané grupy? Pomůžeme nám následující věta, která dává do souvislosti faktorgrupy a homomorfní obrazy grup.

Věta 2.4 (věta o homomorfnímu). Buď $\varphi: G \rightarrow H$ homomorfní skupin.

(1) Je-li $N = \text{Ker}(\varphi)$ normální podgrupa grupy G , pak je zobrazení

$$\psi: G/N \rightarrow H, \quad [a] \mapsto \varphi(a)$$

dobře definované a je to grupový homomorfní skupin.

(2) (1. věta o izomorfnímu) $G/\text{Ker}(\varphi) \cong \text{Im}(\varphi)$.

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 $a \notin b$, a přitom se tímto blokem snažíme přiřadit dvě různé hodnoty $\varphi(a) \neq \varphi(b)$. Ovšem

$[a] = [b] \iff a b^{-1} \in N \iff a b^{-1} \in \text{Ker}(\varphi) \iff \varphi(a b^{-1}) = 1 \iff \varphi(a) = \varphi(b)$,
tedy ψ je dobře definované zobrazení. Protože $\psi([a b]) = \varphi(a b) = \varphi(a) \varphi(b) = \psi([a]) \psi([b])$, je to homomorfní skupin.

(2) Použijeme (1) pro $N = \text{Ker}(\varphi)$. Výsledný homomorfní skupin je prostý, nebož

$$[a] = [b] \iff a b^{-1} \in \text{Ker}(\varphi) \iff \varphi(a b^{-1}) = 1 \iff \varphi(a) = \varphi(b),$$

a uvažujeme-li jej jako zobrazení $G/\text{Ker}(\varphi) \rightarrow \text{Im}(\psi) = \text{Im}(\varphi)$, pak je také na.

1. věta o izomorfnímu je dobrým nástrojem, pokud chceme určit, jak vypadá daná faktorgrupa. Chceme-li dokázat, že $G/N \cong H$, stačí najít homomorfní skupin z G na H , jehož jádrem je N . Metodu ilustrujeme na několika příkladech.

Příklad. Jak vypadá faktorgrupa Z/nZ ? Analýzu situace jsme provedli výše a vidíme, že bychom měli hledat homomorfní skupin $Z \rightarrow Z_n$, jehož jádrem je podgrupa nZ . Situaci řeší zobrazení $a \mapsto a \bmod n$, které je zřejmým homomorfním skupin na Z_n , jehož jádrem je $\ker a \mapsto a \bmod n = 0 \bmod n = nZ$. Podle 1. věty o izomorfnímu

$$Z/nZ \cong Z_n.$$

Příklad. Jak vypadá faktorgrupa S_n/A_n ? Analýzu situace jsme provedli výše a vidíme, že bychom měli hledat homomorfní skupin $S_n \rightarrow Z$, jehož jádrem je podgrupa A_n . Situaci řeší zobrazení $\pi \mapsto \text{sgn}(\pi)$, které je zřejmým homomorfním skupin na Z , jehož jádro tvoří sudé permutace. Podle 1. věty o izomorfnímu

$$S_n/A_n \cong Z.$$

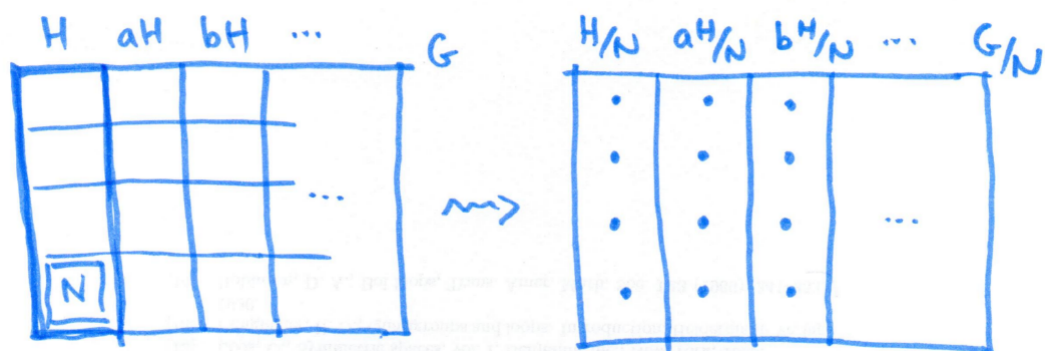
Příklad. Jak vypadá faktorgrupa $GL_n(\mathbb{T})/SL_n(\mathbb{T})$? Platí

$$A \in B \iff AB^{-1} \in SL_n(\mathbb{T}) \iff \det AB^{-1} = \det A(\det B)^{-1} = 1 \iff \det A = \det B.$$

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, èli faktorgrupa $GL_n(\mathbb{T})/SL_n(\mathbb{T})$ by měla být izomorfní grupì \mathbb{T} . Skutečně, zobrazení $\det: GL_n(\mathbb{T}) \rightarrow \mathbb{T}$ je homomorfním skupin na grupu \mathbb{T} , jehož jádro tvoří matice s determinantem 1, èli podgrupa $SL_n(\mathbb{T})$. Podle 1. věty o izomorfnímu je

$$GL_n(\mathbb{T})/SL_n(\mathbb{T}) \cong \mathbb{T}.$$

Jsou případy, kdy podobná analýza nedává žádný dobrý náhled. Leckdy je možné použít dalších triků, například úvah o počtu prvků a znalosti malých grup.



Obrázek 3. Ilustrace 2. vĕty o izomor smu. Vĕtĕí podgrupa H urĕuje hrubĕí ekvivalenci (s vĕtĕími bloky).

Přĕklad. Ukáŕeme, jak vypadá faktorgrupa S_4/\mathbf{K} , kde $K = \text{id}, (12)(34), (13)(24), (14)(23)g$. Podle Lagrangeovy vĕty je $|S_4/\mathbf{K}| = [S_4 : \mathbf{K}] = 24/4 = 6$, ěli faktorgrupa S_4/\mathbf{K} je izomorfnĕí buĕi grupĕi S_3 , nebo cyklickĕ grupĕi Z_6 . Dokáŕeme, ŕe není abelovská, coŕ potvrdĕí první variantu:

$$\begin{aligned} [(123)] [(1234)] &= [(123)(1234)] = [(1342)], \\ [(1234)] [(123)] &= [(1234)(123)] = [(1324)], \end{aligned}$$

ovĕem $[(1342)] \notin [(1324)]$, nebo $[(1342)(1324)]^{-1} = (124) \notin K$.

Jĕnĕm pĕĕkladem pouŕitĕí 1. vĕty o izomor smu je elegantnĕjĕí dĕkaz klasĕkace cyklickĕch grup.

Alternativnĕí dĕkaz Vĕty 1.7. Buĕi $G = \langle a \rangle$ cyklická grupa a uvaŕujme zobrazenĕí

$$\varphi : \mathbb{Z} \rightarrow G, \quad k \mapsto a^k.$$

To je zĕjmĕ na G . Je-li φ také prostĕ, pak je izomor smem $G \cong \mathbb{Z}$. V opaĕnĕm pĕĕpadĕ je $\text{Ker}(\varphi) = n\mathbb{Z}$, kde $n = \text{ord}(a)$, a podle 1. vĕty o izomor smu je $G \cong \mathbb{Z}/n\mathbb{Z} \cong Z_n$.

Jak vypadajĕí podgrupy faktorgrup? O tom hovoŕĕí 2. vĕta o izomor smu.

Tvrzĕnĕí 2.5 (2. vĕta o izomor smu). Buĕi G grupa a N jeĕí normálnĕí podgrupa.

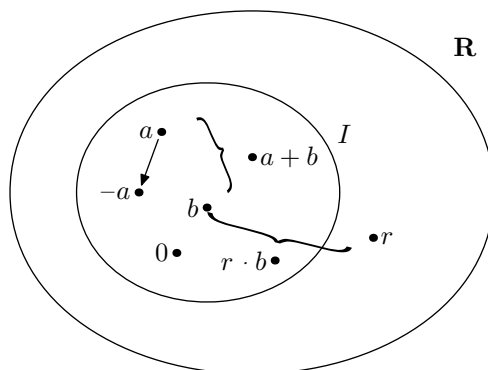
- (1) Je-li $N \subseteq H \subseteq G$, pak H/N je normálnĕí podgrupa v G/N .
- (2) Je-li $K \subseteq G/N$, pak existuje normálnĕí podgrupa $H \subseteq G$ taková, ŕe $K = H/N$.
- (3) Pro $N \subseteq H \subseteq G$ platĕí

$$(G/N)/(H/N) \cong G/H.$$

Dĕkaz. (1) Buĕi $[a], [b]$ prvky H/N , ěli $a, b \in H$, a buĕi $[g]$ prvek G/N . Pak $[a][b] = [ab]$ je prvek H/N , protoŕe $ab \in H$, a ze stejnĕho dĕvodu jsou H/N také $[1], [a]^{-1} = [a^{-1}]$ a $[g][a][g]^{-1} = [gag^{-1}]$.

(2) Buĕi $H = \text{fĕa} \in G : [a] \in Kg$. Pro $a, b \in H$ a $g \in G$ platĕí $ab \in H$, protoŕe $[ab] = [a][b] \in K$, a ze stejnĕho dĕvodu jsou prvky H také $1, a^{-1}$ a gag^{-1} . Zjevĕnĕ $K = H/N$.

(3) Uvaŕujme homomor smus $\varphi : G/N \rightarrow G/H, [a]_N \mapsto [a]_H$. Je dobĕe de novanĕy, protoŕe $N \subseteq H$, a tedy $[a]_N = [b]_N$ implikuje $[a]_H = [b]_H$. Je to homomor smus, $\varphi([a]_N[b]_N) = \varphi([ab]_N) = [ab]_H = [a]_H[b]_H = \varphi([a]_N)\varphi([b]_N)$. Jeho obraz je

Obrázek 4. Ideál I v oboru R .

celé G/H a jeho jádro sestává z těch $[a]_N$, pro které je $a \in H$, tedy $\text{Ker}(\varphi) = H/N$. Aplikací 1. vĕty o izomorfismu dostaneme uvedený vztah.

3. Ideály a dĕlitelnost

Než se dostaneme k okruhovým homomorfismům a ke konstrukci faktorokruhu, je potřeba pochopit pojem ideálu, který hraje v okruzích stejnou roli jako v grupách normální podgrupy.

3.1. Ideály.

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

$$\begin{aligned} \text{pokud } a, b \in I, \text{ pak } & -a \in I \text{ a } a+b \in I, \\ \text{pokud } a \in I \text{ a } r \in R, \text{ pak } & r \cdot a \in I. \end{aligned}$$

Příklad. Množiny $n\mathbb{Z} = \{nz : z \in \mathbb{Z}\} = \{u \in \mathbb{Z} : n \mid u\}$ jsou ideály v oboru \mathbb{Z} . Žádné jiné ideály v oboru \mathbb{Z} nejsou, což si můžete dokázat jako cvičení. Plyne to také z Vĕty 3.2.

Konstrukci ideálů z předchozího příkladu lze zobecnit.

Tvrzení 3.1 (definice hlavních ideálů). Buď R komutativní okruh a $a \in R$. Pak

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

je ideál v R . Obsahuje-li R jednotku, pak aR je nejmenší ideál (nejmenší vzhledem k inkluzi) obsahující prvek a .

Důkaz. Součet a rozdíl dvou prvků dĕlitelných a je dĕlitelný a , a pokud $a \mid u$, pak $a \mid ru$ pro libovolné $r \in R$. Ěli aR je ideál.

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

Definice. Ideály z Tvrzení 3.1 se nazývají hlavní. Speciálně, $0R = 0R$ a $R = 1R$ jsou hlavní ideály v libovolném komutativním okruhu s jednotkou; říká se jim nevlastní.

Hlavní ideály pĕkně odrážejí dĕlitelnost: z tranzitivity relace dĕlitelnosti ihned plyne, že

$$\begin{aligned} a \mid b \text{ právě tehdy, kdy\u017e } bR &= aR; \\ a \nmid b \text{ právě tehdy, kdy\u017e } aR &\neq bR. \end{aligned}$$

Původní motivací studia ideálů bylo řešení problému nejednoznačných ireducibilních rozkladů. Idea byla, že nejednoznačnost je způsobena jakýmsi chybějícím prvky, které by dělily činitele v těchto nejednoznačných rozkladech. Například v $\mathbb{Z}[\sqrt{5}]$ máme $4 = 2^2 = (1 + \sqrt{5})(1 - \sqrt{5})$, kdyby ovšem existovaly nějaké prvoideální 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 jeden rozklad $4 = p^2q^2$. Tímto prvoideálními prvky se nakonec ukázaly být tzv. prvoideály, de nicí uvidíme později. 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.

3.2. Obory hlavních ideálů.

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

Cílem této podsekcce je zařadit obory hlavních ideálů do hierarchie oborů integrity z hlediska dělitelnosti.

Příklad. Obory \mathbb{Z} nebo $\mathbb{T}[x]$, \mathbb{T} těleso, jsou obory hlavních ideálů. Obecněji, eukleidovské obory mají pouze hlavní ideály (Věta 3.2).

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 zimním semestru 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 $au + av \in aR$, čili $au + av \in aR$, a tedy $aR = R$.

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

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

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

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

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

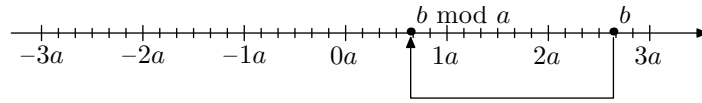
Je vidět, že jde o ideál. Přitom I obsahuje polynomy x_1 a x_2 , které jsou nesoudělné, nemůže tedy být hlavní.

Věta 3.2. V eukleidovských oborech je každý ideál hlavní.

Důkaz. Buď I ideál v eukleidovském oboru \mathbb{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.

Obrázek 5. Ilustrace důkazu Věty 3.2 v případě $\mathbf{R} = \mathbf{Z}$.

Tvrzení 3.3 (ideály v tělesech). Buď \mathbf{R} komutativní okruh s jednotkou. Pak \mathbf{R} je těleso právě tehdy, když má pouze nevlastní ideály.

Důkaz. () Tělesa jsou eukleidovské obory, tedy každý ideál je hlavní. Pro každé $a \notin 0$ platí $a \mid 1$, ěli pro každý nenulový ideál aR platí $aR = 1R = R$.

() Pro každý hlavní ideál aR , $a \notin 0$, platí $aR = R = 1R$, ěli každý nenulový prvek a je invertibilní.

Nyní si dokážeme jedno pomocné tvrzení o obecných ideálech.

Tvrzení 3.4 (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 $I + J = \{fa + gb : a \in I, b \in J\}$ je také ideál v \mathbf{R} . Tento ideál je nejmenší takový, ě obsahuje $I \cup J$.
- (3) Jsou-li $I_j, j \in \mathbf{N}$, ideály v \mathbf{R} takové, ě $I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$, pak $\bigcup_{j \in \mathbf{N}} I_j$ je také ideál v \mathbf{R} .

Důkaz. (1) Buď $a, b \in I \cap J$ a $r \in R$. Pak $a + b, ar$ náležejí 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 $(a + b) + (c + d) = (a + c) + (b + d) \in I + J$, $(a + b)r = (ar) + (br) \in I + J$ 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 $a, b \in \bigcup_{j \in \mathbf{N}} I_j$ a $r \in R$. Pak existují $j, k \in \mathbf{N}$ taková, ě $a \in I_j$ a $b \in I_k$, ěli $a, b \in I_{\max(j,k)}$, a tedy $a + b, ar \in I_{\max(j,k)} \subseteq \bigcup_{j \in \mathbf{N}} I_j$.

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

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

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

$$a_1R + \dots + a_nR = \left\{ \sum_{i=1}^n a_i r_i : r_1, \dots, r_n \in R \right\}$$

Tímto prvkům se také říká báze ideálu (bez nároku na nezávislost, jak je zvykem v lineární algebře). Výše uvedené nehlavní ideály pak můžeme napsat jako $\mathbf{Z}[x] + x\mathbf{Z}[x]$, resp. $x_1\mathbf{R}[x_1, \dots, x_k] + \dots + x_k\mathbf{R}[x_1, \dots, x_k]$.

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

Věta 3.5. Obory hlavních ideálů jsou gaussovské a platí v nich Bézoutova rovnost.

Důkaz. Buď \mathbf{R} obor hlavních ideálů. Podle věty ze zimního semestru stačí dokázat, ě v \mathbf{R} (1) existují NSD a (2) neexistují nekonečné posloupnosti vlastních dělitelů. Pripomeňme, ě pro libovolná u, v platí $u \mid v, v \mid uR$.

(1) Zvolme $a, b \in R$ a označme $I = aR + bR$. Každý ideál je hlavní, existuje tedy $c \in R$ takové, ě $I = cR$. 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 $I \subseteq dR$ a dostáváme $d \mid c$. Vidíme, ě $c = \text{NSD}(a, b)$ 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 R existuje nekonečná posloupnost vlastních dělitelů a_1, a_2, \dots (tj. $a_{i+1} \nmid a_i$ a $a_i \nmid a_{i+1}$). Pak $a_1R \supset a_2R \supset a_3R \supset \dots$ 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 $i \in \mathbb{N}$. Ale pak $bR \supset a_iR \supset a_{i+1}R \supset \dots \supset I = bR$, spor.

Shrnutí:

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

4. Okruhové homomorfismy a faktorokruhy

4.1. Homomorfismy.

V celé podsekcí budou R, S značit dva okruhy. Než přistoupíme k definici homomorfismu, definujme pojem ideálu obecně, i pro okruhy které nejsou nutně komutativní.

Definice. (Oboustranným) ideálem v okruhu R nazýváme každou neprázdnou podmnožinu $I \subseteq R$ takovou, že

$$\begin{aligned} \text{pokud } a, b \in I, \text{ pak } a + b \in I \text{ a } a - b \in I, \\ \text{pokud } a \in I \text{ a } r \in R, \text{ pak } ra \in I \text{ i } ar \in I. \end{aligned}$$

Okruhové homomorfismy jsou zobrazení zachovávající základní okruhové operace. Většina faktů v této sekci je přímou analogií situace, kterou jsme viděli v grupách.

Definice. Zobrazení $\varphi: R \rightarrow S$ se nazývá homomorfismus těchto okruhů, zapisujeme $\varphi: R \rightarrow S$, pokud pro každé $a, b \in R$ platí

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

Z Lemmatu 1.1 ihned plyne, že $\varphi(-a) = -\varphi(a)$ pro všechna $a \in R$ a $\varphi(0) = 0$.

Obrazem homomorfismu nazýváme jeho obor hodnot, tj. množinu

$$\text{Im}(\varphi) = \{ \varphi(a) \in S : a \in R \}$$

Jádro homomorfismu definujeme jako množinu

$$\text{Ker}(\varphi) = \{ a \in R : \varphi(a) = 0 \}$$

Následující tvrzení mimo jiné říká, že ideály hrají vzhledem k homomorfismům roli normálních podgrup.

Tvrzení 4.1 (jádro je ideál, obraz je podokruh). Buď R, S okruhy a $\varphi: R \rightarrow S$ homomorfismus. Pak

- (1) $\text{Im}(\varphi)$ tvoří podokruh okruhu S ;
- (2) $\text{Ker}(\varphi)$ tvoří ideál okruhu R .

Důkaz. Okruhový homomorfismus je zároveň grupový homomorfismus vzhledem k operacím $+$, \cdot , 0 , 1 , ěli můžeme použít Tvzení 1.2 a ihned dostaneme uzavřenost jádra a obrazu na operaci $+$, \cdot , 0 . Uzavřenost obrazu na násobení se dokáže stejně jako pro sčítání. Zbývá dokonit část (2): 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 $a \cdot r$, ěli $\text{Ker}(\varphi)$ tvoří ideál v R .

Z Tvzení 1.3 ihned plyne jeho analogie pro okruhy.

Tvzení 4.2. Buď R, S okruhy a $\varphi : R \rightarrow S$ homomorfismus. Pak φ je prostý právě tehdy, když $\text{Ker}(\varphi) = \{0\}$.

Podobně jako pro grupy se dokáže také následující tvrzení (proveďte jako cvičení!).

Tvzení 4.3. Buď R, S, T okruhy a $\varphi : R \rightarrow S$, $\psi : S \rightarrow T$ homomorfy. Pak

- (1) $\psi \circ \varphi$ je homomorfismus $R \rightarrow T$,
- (2) je-li φ bijektivní, pak ψ^{-1} je homomorfismus $S \rightarrow R$.

Příklad. Důležitou rodinou jsou tzv. modulární homomorfy. V číselné variantě zvolme číslo $m > 0$ a definujme zobrazení

$$\varphi_m : \mathbb{Z} \rightarrow \mathbb{Z}_m, \quad a \mapsto a \pmod{m}.$$

V polynomiální variantě zvolme polynom $0 \neq m \in \mathbb{T}[x]$ a definujme zobrazení

$$\varphi_m : \mathbb{T}[x] \rightarrow \mathbb{T}[x]/(m), \quad f \mapsto f \pmod{m}.$$

Není těžké ověřit, že jde o homomorfismus, jejich jádra jsou $m\mathbb{Z}$, resp. $m\mathbb{T}[x]$. Podobně bychom mohli postupovat v každém oboru, kde je definováno jednoznačné dělení se zbytkem.

Příklad. Důležitou rodinou jsou tzv. dosazovací homomorfy. Uvažujme komutativní okruh $R = S$ a prvek $a \in S$ a definujme zobrazení

$$\varphi_a : R[x] \rightarrow S, \quad f \mapsto f(a).$$

Není těžké ověřit, že jde o homomorfismus. Je-li $R = S$, jeho jádrem je hlavní ideál $(x - a)R[x]$ a obrazem celé R (díky konstantním polynomům). Obecně může jádro a obraz vycházet v libovolný, např. pro $R = \mathbb{Z}$, $S = \mathbb{C}$, $a = i$ dostaneme jádro $(x^2 + 1)\mathbb{Z}[x]$ a obraz $\mathbb{Z}[i]$.

Příklad. Oba výše uvedené typy lze kombinovat, například zobrazení

$$\varphi : \mathbb{Z}[x] \rightarrow \mathbb{Z}_2, \quad f \mapsto f(0) \pmod{2}$$

je také homomorfismus, jeho jádro sestává z těch polynomů, jejichž absolutní člen je sudý, což není hlavní ideál.

4.2. Izomorfismy.

Bijektivní homomorfismus se nazývají izomorfismus. Dva okruhy nazveme izomorfní, pokud mezi nimi vede izomorfismus, známe $R \cong S$. Víe, co bylo v sekci 1.2 řešeno o izomorfismech grup, platí analogicky i o izomorfismech okruhů.

Na izomorfismus je možné pohlížet jako na přepokopování: máme-li okruh R a bijektivní zobrazení $\varphi : R \rightarrow S$, můžeme na množinu S přepokopovat obě operace $\in f+$, g předpisem

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

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

Příklad. Je snadné nahlédnout, že množina matic

$$S = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$$

tvorí podokruh okruhu $M_2(\mathbb{R})$ matic 2×2 nad reálnými čísly. Zobrazení

$$\varphi: \mathbb{C} \rightarrow S, \quad a + bi \mapsto \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

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

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

a podobně

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

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

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

$$\varphi: \mathbb{Z}_n \rightarrow \mathbb{P}, \quad 0 \mapsto 0, \quad k \mapsto \underbrace{1 + \dots + 1}_k,$$

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

Důležitým příkladem izomorfismu je modulární zobrazení z důkazu čínského věty o zbytcích. Na všech prvcích se chová jako okruhový izomorfismus. Vztahovaný na invertibilní prvky se chová jako grupový izomorfismus.

Tvrzení 4.4 (algebraická verze čínského věty o zbytcích). Buďte m_1, \dots, m_n po dvou nesoudělná přirozená čísla a označme $M = m_1 \dots m_n$. Zobrazení

$$\begin{aligned} \varphi: \mathbb{Z}_M &\rightarrow \mathbb{Z}_{m_1} \times \dots \times \mathbb{Z}_{m_n} \\ a &\mapsto (a \bmod m_1, \dots, a \bmod m_n). \end{aligned}$$

je izomorfismem těchto okruhů. Restrikce $\varphi|_{\mathbb{Z}_M^\times}$ je grupovým izomorfismem

$$\mathbb{Z}_M^\times \rightarrow \mathbb{Z}_{m_1}^\times \times \dots \times \mathbb{Z}_{m_n}^\times.$$

Důkaz. Pohledem do důkazu čínského věty o zbytcích zjistíme, že je zobrazení φ bijektivní. Ověříme, že to je homomorfismus: pro obě operace $\in \mathbb{F}_+$, g platí

$$\begin{aligned} \varphi(a) \varphi(b) &= (a \bmod m_1, \dots, a \bmod m_n) (b \bmod m_1, \dots, b \bmod m_n) \\ &= ((a \cdot b) \bmod m_1, \dots, (a \cdot b) \bmod m_n) = \varphi(a \cdot b \bmod M), \end{aligned}$$

při němž v poslední rovnosti využíváme faktu, že všechna m_i dělí M .

Druhou část důkazu necháváme jako cvičení: dokažte, že invertibilní prvky modulo M jsou zobrazeny na invertibilní prvky modulo jednotlivá m_i , čili $\varphi|_{\mathbb{Z}_M^\times}$ je skutečně bijekcí na množinu $\mathbb{Z}_{m_1}^\times \times \dots \times \mathbb{Z}_{m_n}^\times$.

Analogicky bychom mohli interpretovat čínskou větu o zbytcích pro polynomy (cvičení).

4.3. Konstrukce faktorokruhu podle ideálu.

V zimním semestru 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 . Vzhledem k tomu, že ideál tvoří normální podgrupu aditivní grupy daného okruhu, můžeme v konstrukci faktorokruhu vyjít vte, co jsme udělali pro faktorgrupy.

Definice. Buď R okruh a I jeho ideál. Definujeme ekvivalenci na množině R předpisem

$$a \sim b, \quad a - b \in I.$$

Platí $a \sim b$ právě tehdy, když $a + I = b + I$, ěli bloky jsou rozkladové třídy $[a] = a + I$. Na těchto blocích definujeme operace předpisem

$$[a] + [b] = [a + b], \quad [a] = [a], \quad [a][b] = [ab].$$

(v následujícím lemmatu ověříme, že je tato definice korektní). Množina bloků s výše uvedenými operacemi se nazývá faktorokruh okruhu R podle ideálu I ,

$$R/I = (\{[a] : a \in R, +, \cdot, 0, 1\}).$$

Lemma 4.5. Buď R okruh a I jeho ideál.

- (1) Výše uvedené operace na blocích jsou dobře definovány.
- (2) Faktorokruh R/I je skutečně okruh.

Důkaz. Z konstrukce faktorgrupy již víme, že je sětání a odětání dobře definované. Pro násobení předpokládejme $[a] = [c]$ a $[b] = [d]$, ověříme, že $[ab] = [cd]$. Předpokládáme $a - c \in I$ a $b - d \in I$, spoěteme

$$ab - cd = \underbrace{a(b - d)}_{\in I} + \underbrace{(a - c)d}_{\in I},$$

ěli $ab - cd \in I$, tj. $[ab] = [cd]$. Podobně jako pro grupy se ukáže, že takto definované operace splňují všechny axiomy okruhů, větěně komutativity a existence jednotky, pokud tyto platily v původním okruhu R (pozor, vlastnost běti oborem integrity zachována bět nemusí, viz sekce 4.4).

Přěklad. UvaŹujme komutativně okruh $R = Z$ a ideál $I = nZ$. Platě

$$a \sim b, \quad n \mid a - b, \quad a \sim b \pmod{n}.$$

Stejně jako pro grupy není problěm ukázat, že $Z/nZ \cong Z_n$.

Podobně jako pro grupy platě věta o homomor smu a 1. věta o izomor smu.

Věta 4.6 (věta o homomor smu). Buď $\varphi : R \rightarrow S$ homomor smus okruhů.

- (1) Je-li $I = \text{Ker}(\varphi)$ ideál v R , pak je zobrazeně

$$\psi : R/I \rightarrow S, \quad [a] \mapsto \varphi(a)$$

dobře definované a je to okruhový homomor smus.

- (2) [1. věta o izomor smu] $R/\text{Ker}(\varphi) \cong \text{Im}(\varphi)$.

Důkaz. Plyne přímo z grupové verze (Věta 2.4), pouze je třeba si uvědomit, že věchna dobře definovaná zobrazeně jsou i okruhové homomor smy.

Přěklad. Modulární homomor smus $Z \rightarrow Z_n, a \mapsto a \pmod{n}$, prokazuje, že $Z/nZ \cong Z_n$.

Je-li \mathbf{R} komutativní okruh a $I = mR$ jeho hlavní ideál, pak

$$a + b, \quad m \mid a - b, \quad a - b \pmod{m}.$$

Je-li v okruhu \mathbf{R} dle nováno jednoznačně dělení se zbytkem (např. $\mathbf{R} = \mathbb{Z}$ nebo $\mathbf{R} = \mathbb{T}[x]$, \mathbb{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 , nebo

$$[a] + [b] = [a + b] = [a + b \pmod{m}], \quad [a] \cdot [b] = [a \cdot b] = [a \cdot b \pmod{m}].$$

Speciálně, pro $\mathbf{R} = \mathbb{T}[x]$, \mathbb{T} těleso, vidíme, že konstrukce faktorokruhu ve smyslu této sekce a ve smyslu zimního semestru jsou v podstatě totožné. 1. věta o izomorismu použitá na modulární zobrazení $\mathbb{T}[x] \rightarrow \mathbb{T}[x]/(m)$ dává izomorismus

$$\mathbb{T}[x]/m\mathbb{T}[x] \cong \mathbb{T}[x]/(m),$$

v němž polynom f stupně $< \deg m$ jednoznačně odpovídá příslušnému bloku $[f]$. Značení se zpravidla smíhuje, používají se oba zápisy $\mathbb{T}[x]/m\mathbb{T}[x]$ i $\mathbb{T}[x]/(m)$ pro obě formálně různé, nicméně izomorfní konstrukce.

Příklad. Jak vypadá faktorokruh $\mathbb{T}[x]/(x - a)$, kde $a \in \mathbb{T}$? Uvažujme dosazovací homomorismus

$$\mathbb{T}[x] \rightarrow \mathbb{T}, \quad f \mapsto f(a).$$

Je to zobrazení na \mathbb{T} a jeho jádro je

$$\{f \in \mathbb{T}[x] : f(a) = 0\} = \{f \in \mathbb{T}[x] : (x - a) \mid f\} = (x - a)\mathbb{T}[x].$$

Podle 1. věty o izomorismu je $\mathbb{T}[x]/(x - a) \cong \mathbb{T}$.

Pro polynomy vyššího stupně je situace složitější.

Příklad. Jak vypadá faktorokruh $\mathbb{Q}[x]/(x^2 + 1)$? Uvažujme homomorismus

$$\mathbb{Q}[x] \rightarrow \mathbb{Q}(i), \quad f \mapsto f(i).$$

Je to zobrazení na $\mathbb{Q}(i)$ a jeho jádro je

$$\begin{aligned} \{f \in \mathbb{Q}[x] : f(i) = 0\} &= \{f \in \mathbb{Q}[x] : f(-i) = 0\} \\ &= \{f \in \mathbb{Q}[x] : (x - i) \mid f, (x + i) \mid f\} \\ &= \{f \in \mathbb{Q}[x] : (x - i)(x + i) \mid f\} = (x^2 + 1)\mathbb{Q}[x] \\ &= (x^2 + 1)\mathbb{Q}[x]. \end{aligned}$$

Podle 1. věty o izomorismu je $\mathbb{Q}[x]/(x^2 + 1) \cong \mathbb{Q}(i)$.

Příklad. Jak vypadá faktorokruh $\mathbb{Q}[x]/(x^2 - 1)$? Uvažujme homomorismus

$$\mathbb{Q}[x] \rightarrow \mathbb{Q} \times \mathbb{Q}, \quad f \mapsto (f(1), f(-1)).$$

Je to zobrazení na $\mathbb{Q} \times \mathbb{Q}$ a jeho jádro je

$$\begin{aligned} \{f \in \mathbb{Q}[x] : f(1) = f(-1) = 0\} &= \{f \in \mathbb{Q}[x] : (x - 1) \mid f, (x + 1) \mid f\} \\ &= \{f \in \mathbb{Q}[x] : (x - 1)(x + 1) \mid f\} = (x^2 - 1)\mathbb{Q}[x] \\ &= (x^2 - 1)\mathbb{Q}[x]. \end{aligned}$$

Podle 1. věty o izomorismu je $\mathbb{Q}[x]/(x^2 - 1) \cong \mathbb{Q} \times \mathbb{Q}$.

Na závěr jeden příklad s ideálem, který není hlavní.

Příklad. Jak vypadá faktorokruh $Z[x]/I$, kde $I = \{f \in Z[x] : m \mid f(0)\}$? 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 $f(0) - g(0) \pmod{m}$. Existuje tedy přesná m rozkladových tříd. Není třeba nahlédnout, jaké zobrazení

$$\varphi: Z[x] \rightarrow Z_m, \quad \varphi(f) = f(0) \pmod{m}$$

je homomorfismus, jehož jádro je I , a tedy $Z[x]/I \cong Z_m$.

Podobně jako pro grupy se dokáže i 2. věta o izomorfismu (cvičení!).

Tvrzení 4.7 (2. věta o izomorfismu). Buď R okruh a I jeho ideál.

- (1) Je-li $J \subseteq I$ ideál v R , pak $J/I = \{f[a] : a \in J\}$ je ideál v R/I .
- (2) Je-li K ideál v R/I , pak existuje ideál J v R takový, že $K = J/I$.
- (3) V obou případech platí

$$(R/I)/(J/I) \cong R/J.$$

4.4. Faktorokruhy podle maximálních ideálů a prvoideálů.

V této části si ukážeme analogii k tvrzení, 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 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 $I \subsetneq J \subseteq R$.

Příklad. Jako cvičení si dokažte:

Ideál nZ v oboru Z je maximální právě tehdy, když to je prvoideál, což je právě tehdy, když n je prvočíslo.

Ideál $fT[x]$ v oboru $T[x]$ je maximální právě tehdy, když to je prvoideál, což je právě tehdy, když f je ireducibilní v $T[x]$.

Analogické tvrzení platí v každém oboru hlavních ideálů, ale obecně jsou maximální ideály a prvoideály různé pojmy. Například ideál $I = xZ[x]$ v oboru $Z[x]$

je prvoideálem, protože je polynom x prvočinitelem, není maximální, například ideál $\{f \in Z[x] : 2 \mid f(0)\}$ je větší.

Věta 4.8 (faktor podle prvoideálu a maximálního ideálu). Buď R komutativní okruh s jednotkou a I jeho ideál. Pak

- (1) R/I je obor integrity právě tehdy, když I je prvoideál;
- (2) R/I je těleso právě tehdy, když I je maximální ideál.

Důkaz. (1) Faktorokruh R/I je oborem integrity právě tehdy, když pro každé $a, b \in R$ platí, že kdykoliv $[ab] = [a][b] = [0]$, 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 3.3 je R/I tělesem právě tehdy, když neobsahuje žádné vlastní ideály. 2. věta o izomorfismu říká, že jakýkoliv vlastní ideál v R/I je tvaru J/I , kde $I \subsetneq J \subseteq R$ je ideál v R . Ěli neexistence vlastního ideálu v R/I je ekvivalentní tomu, že I je maximální ideál.

Příklad. Připomeňte si příklady $\mathbb{Q}[x]/(f)$ ze sekce 4.3:

polynom $x^2 + 1$ je ireducibilní, tedy ideál $(x^2 + 1)\mathbb{Q}[x]$ je maximální, a skutečně, $\mathbb{Q}[x]/(x^2 + 1) \cong \mathbb{Q}(i)$ 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, $\mathbb{Q}[x]/(x^2 - 1) \cong \mathbb{Q} \times \mathbb{Q}$ není těleso.

Příklad. Vzpomeňte si na příklad s oborem $\mathbb{Z}[x]$: ideál $I = x\mathbb{Z}[x]$ je prvoideálem, který není maximální, a vskutku, faktorokruh $\mathbb{Z}[x]/I \cong \mathbb{Z}$ (dosazovací homomorfismus) je oborem integrity, ale ne tělesem.

Číselná tělesa a kořeny polynomů

5. Okruhová a tělesová rozšíření

5.1. Definice.

V této sekci definujeme obecný pojem rozšíření a ukážeme, že jeho prvky lze vyjádřit pomocí hodnot polynomů.

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

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

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

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

$$\begin{aligned} \mathbb{Z}[\sqrt{s}] &= \{fa + b\sqrt{s} : a, b \in \mathbb{Z}\} \subset \mathbb{C}, \\ \mathbb{Q}[\sqrt{s}] &= \mathbb{Q}(\sqrt{s}) = \{fa + b\sqrt{s} : a, b \in \mathbb{Q}\} \subset \mathbb{C}. \end{aligned}$$

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

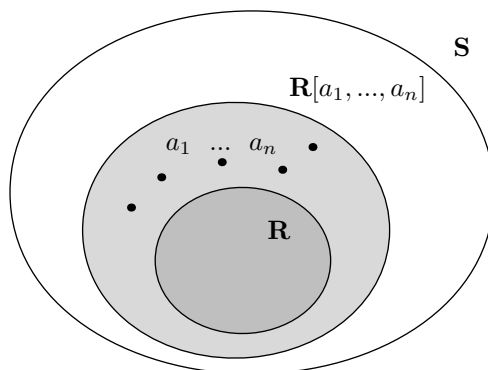
Příklad (cyklotomická rozšíření). Pro $\zeta_n = e^{2\pi i/n}$ (tzv. primitivní n -tá odmocnina z jedné) uvažujme n -té cyklotomické rozšíření

$$\begin{aligned} \mathbb{Z}[\zeta_n] &= \{fa_0 + a_1\zeta_n + a_2\zeta_n^2 + \dots + a_{n-1}\zeta_n^{n-1} : a_0, \dots, a_{n-1} \in \mathbb{Z}\} \subset \mathbb{C}, \\ \mathbb{Q}[\zeta_n] &= \mathbb{Q}(\zeta_n) = \{fa_0 + a_1\zeta_n + a_2\zeta_n^2 + \dots + a_{n-1}\zeta_n^{n-1} : a_0, \dots, a_{n-1} \in \mathbb{Q}\} \subset \mathbb{C}. \end{aligned}$$

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. Později si ukážeme obecné tvrzení 6.4, z něhož tento fakt ihned plyne.

Příklad. Můžeme uvažovat i rozšíření o více prvků, například

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

Obrázek 6. Rozšíření $R[a_1, \dots, a_n]$ S .

Tvrzení 5.1 (struktura okruhových rozšíření). Buď R S komutativní okruhy s jednotkou a $a \in S$. Pak

$$\begin{aligned} R[a] &= \{f(a) : f \in R[x]\} \\ &= \{f_0 + u_1 a + \dots + u_n a^n : n \in \mathbb{N}, u_0, \dots, u_n \in R\}. \end{aligned}$$

Jsou-li R S tělesa, pak

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

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

- (1) tvoří podokruh okruhu S ,
- (2) obsahuje R $\{f(a)g\}$,
- (3) je nejmenší podmnožinou okruhu 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 U obsahující R $\{f(a)g\}$. Tento podokruh musí obsahovat všechny mocniny a^i , jejich libovolné násobky prvky z R , a také součty těchto násobků. ěili musí obsahovat všechny prvky tvaru $u_0 + u_1 a + \dots + u_n a^n$, kde $u_0, \dots, u_n \in R$, a tedy $M \subseteq U$.

Vyjádření tělesových rozšíření se dokáže analogicky, navíc musíme dát pozor na inverzní prvky (cvičení!).

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

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

Pro každé těleso T platí $T[a] = T(a)$, protože v okruhu $T(a)$ navíc požadujeme přítomnost inverzních prvků. Za jakých podmínek platí $T[a] = T(a)$? V sekci 6 si ukážeme, že to nastane právě tehdy, když a tzv. algebraický prvek, tedy když je kořenem nějakého nenulového polynomu z $T[x]$. Jednu implikaci si můžeme dokázat hned.

Tvrzení 5.2. Buď T S tělesa a $a \in S$ prvek, který není kořenem žádného nenulového polynomu z $T[x]$. Pak $T[a] \neq T(a)$.

Důkaz. Podle Tvzení 5.1 je $T[a] = fT[a] : f \in T[x]$. Kdyby se v této množině nacházel prvek a^{-1} , pak by existoval polynom $f \in T[x]$ takový, že $f(a) = a^{-1}$, eili $af(a) = 1$, a tedy a by bylo kořenem nenulového polynomu $xf - 1 \in T[x]$, spor.

5.2. Tělesové rozšíření jako vektorový prostor.

Hlavním tématem této kapitoly je studium tzv. stupně rozšíření, tj. dimenze vřtřního tělesa jakožto vektorového prostoru nad svým podtělesem. V jakém smyslu, dimenze?

Rozšířením těles budeme rozumět dvojici komutativních těles T, S takovou, že $T \subseteq S$. Říkáme, že T je podtělesem S , nebo že S je rozšířením T .

Klíčem k pochopení celé kapitoly je myšlenka, že těleso S lze považovat za vektorový prostor nad tělesem T : sčítání a odčítání ponecháme a místo násobení jakožto operace $S \cdot S \rightarrow S$ uvažujeme pouze restriktci $T \cdot S \rightarrow S$. Neformálně, prvky vřtřního tělesa S považujeme za vektory, prvky menřního tělesa T za skaláry a uvažujeme pouze násobení skalár krát vektor. Tento vektorový prostor budeme značit S_T .

Uvřdomte si, že jde skutečně o vektorový prostor: aditivní struktura $(S, +, \cdot, 0)$ je abelovskou grupou a pro vřechna $a, b \in T$ (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 S_T se nazývá stupeň rozšíření a značí se

$$[S : T] = \dim S_T.$$

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

Přklady.

$[C : R] = 2$. Každé komplexní číslo lze zapsat právě jedním způsobem jako $a + bi$, $a, b \in R$, eili prvky $1, i$ tvoř bází prostoru C_R .

Analogicky, pro s , které není čtvercem, je stupeň $[Q(\sqrt{s}) : Q] = 2$, prvky $1, \sqrt{s}$ tvoř bází prostoru $Q(\sqrt{s})_Q$.

$[Q(\sqrt{2}, \sqrt{3}) : Q] = 4$, bází prostoru $Q(\sqrt{2}, \sqrt{3})_Q$ tvoř například prvky $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$.

Pozor, pro $\zeta_3 = e^{2\pi i/3}$ je stupeň $[Q(\zeta_3) : Q] = 2$ 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 $[Q(u) : Q]$ je nekonečný (spočetný): lineárně nezávislou množinu tvoř třeba prvky $1, u, u^2, \dots$ (viz Vřta 6.6).

Stupeň $[R : Q]$ je dokonce nespočetný: prostory spočetné dimenze nad spočetným tělesem jsou spočetné, zatímco reálných čísel je nespočetně.

Ěasem se nám bude hodit pojem prvookruhu a prvotělesa. Pro libovolný okruh R s jednotkou uvažujme zobrazení

$$Z \rightarrow R, \quad n \mapsto \underbrace{1 + \dots + 1}_n.$$

Je vidět, že jde o homomor smus, jehož obrazem je tzv. prvookruh okruhu R a jehož jádrem je ideál nZ , kde n je charakteristika okruhu R . Použitím 1. vřty o izomor smu dostáváme, že prvookruh libovolného okruhu je izomorfní buř okruhu Z v přpadě charakteristiky 0, nebo okruhu $Z/nZ \cong Z_n$ v přpadě charakteristiky n .

Nyní uvažujme těleso T . Jeho prvotělesem se rozumí nejmenří podtěleso. To v sobě jistě obsahuje prvookruh, ale navíc musí ke každému nenulovému prvku obsahovat jeho inverz. Charakteristika tělesa je 0 nebo prvočíslo p . V druhém přpadě je prvookruh již tělesem (izomorfním Z_p), eili pojmy splývají. V přpadě charakteristiky 0 prvotěleso sestává ze vřech zlomků ab^{-1} , kde a, b jsou prvky prvookruhu, eili je izomorfní tělesu Q .

Každé těleso je samozřejmě rozšířením svého prvotního tělesa. Speciálně, pro konečná tělesa dostáváme velmi zajímavý důsledek vektorového pohledu na tělesová rozšíření.

Tvrzení 5.3. Počet prvků konečného tělesa je mocnina prvočísla.

Důkaz. Konečné těleso T charakteristiky p je rozšířením svého prvotního tělesa $P \cong \mathbb{Z}_p$. ělí vektorový prostor T_P je izomorfní prostoru $(\mathbb{Z}_p)^k$, kde $k = [T : P]$, ělí má p^k prvků.

6. Algebraické prvky a rozšíření konečného stupně

6.1. 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, která charakterizuje polynomy, jejichž kořeny lze vyjádřit vzorci. Pro polynomy stupně ≤ 4 existují tzv. Cardanovy vzorce, ale pro některé polynomy stupně ≥ 5 žádné vzorce neexistují a v praxi se kořeny hledají pouze přibližně, pomocí numerických metod. V této podsekcí 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 lze vyjádřit algebraickými.

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 6.10).

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

V roce 1840 dokázal Joseph Liouville, že iracionální algebraická čísla nelze, v jistém smyslu, dobře aproximovat racionálními čísly. Z toho důvodu nemůže být algebraické například číslo $\sum_{i=1}^7 10^{-i!}$ (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 transcendentních čísel, aniž by musel nějaké nalézt. Jeho argument je založen na počítání: transcendentních čísel je mnohem víc (nespočetní) než těch algebraických (těch je jen spočetní). Cantorův důkaz, který byl jednou z hlavních motivací vzniku teorie množin, nyní ukážeme.

Připomeďme, 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 v'ímňte, že sjednocení dvou spočetných množin je spočetné: je-li $A = \{a_1, a_2, \dots, g\}$ a $B = \{b_1, b_2, \dots, g\}$, pak $A \cup B = \{a_1, b_1, a_2, b_2, \dots, g\}$.

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 6.1. Množina algebraických čísel je spočetná.

Důkaz. Deňujme index polynomu $f = a_0 + a_1x + \dots + a_nx^n \neq 0$ jako součet $|a_0| + |a_1| + \dots + |a_n| + n$. V'ímňte si, že existuje jen konečně mnoho celočíselných polynomů daného indexu (např. index 1: $f = 1$; index 2: $f = 2, f = x$; index 3: $f = 3, f = 2x, f = x - 1, f = x^2$), 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 6.2. 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

$$a_1 = 0, a_{11}a_{12}a_{13} \dots$$

$$a_2 = 0, a_{21}a_{22}a_{23} \dots$$

$$a_3 = 0, a_{31}a_{32}a_{33} \dots$$

...

Nyní deňujme číslo $b = 0, b_1b_2b_3 \dots$ 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 rozvojem 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 dalším textu dáme do souvislosti algebraičnost daného čísla se stupněm jistého tělesového rozšíření a také si ůekneme, jak pro algebraická čísla hledat polynomy, jejichž jsou kořenem.

6.2. Minimální polynom a stupeň jednoduchého rozšíření.

V této sekci uvedeme do souvislosti pojem algebraického čísla s vlastnostmi tzv. jednoduchých rozšíření, tj. rozšíření tvaru $\mathbb{T}(a)$, 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.

Zaňeme obecněji deňnicí algebraičnosti, nad libovolným tělesem.

Definice. Buď $\mathbb{T} \supset \mathbb{S}$ rozšíření tělesa $a \in \mathbb{S}$. ůekneme, že prvek a je algebraický nad \mathbb{T} , pokud existuje nenulový polynom $z \in \mathbb{T}[x]$, jehož je a kořenem. V opačném případě se prvek a nazývá transcendentní nad \mathbb{T} .

Uv'ídte si, že číslo je algebraické ve smyslu sekce 6.1 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 $\mathbf{T} \subseteq \mathbf{S}$ rozšíření těles a $a \in \mathbf{S}$ algebraický nad \mathbf{T} . Minimálním polynomem prvku a nad \mathbf{T} rozumíme ireducibilní monický polynom $m_{a,\mathbf{T}} \in \mathbf{T}[x]$, jehož je a kořenem.

Tvrzení 6.3 (vlastnosti minimálních polynomů). Buď $\mathbf{T} \subseteq \mathbf{S}$ rozšíření těles a $a \in \mathbf{S}$ algebraický nad \mathbf{T} . Pak

- (1) minimální polynom $m_{a,\mathbf{T}}$ existuje a je jednoznačně určený;
- (2) prvek a je kořenem polynomu $f \in \mathbf{T}[x]$ právě tehdy, když $m_{a,\mathbf{T}} \mid f$.

Důkaz. Množina $I = \{f \in \mathbf{T}[x] : f(a) = 0\}$ tvoří ideál v oboru $\mathbf{T}[x]$, a protože je $\mathbf{T}[x]$ oborem hlavních ideálů (Věta 3.2), existuje monický polynom $m \in \mathbf{T}[x]$ takový, že $I = m\mathbf{T}[x]$. Vidíme, že $f(a) = 0$ právě tehdy, když $m \mid f$. Kdyby polynom m nebyl ireducibilní v $\mathbf{T}[x]$, tj. kdyby $m = fg$, kde $f, g \in \mathbf{T}[x]$, pak $0 = m(a) = f(a)g(a)$, èli prvek a by byl kořenem alespoň jednoho z polynomů f, g , ale $m \nmid f, g$, spor. Èli m je minimální polynom prvku a nad \mathbf{T} . Pro jakýkoliv jiný monický ireducibilní polynom $\tilde{m} \in \mathbf{T}[x]$, jehož je a kořenem, platí $m \mid \tilde{m}$ a z ireducibility a moničnosti dostáváme $\tilde{m} = m$.

Příklad. Je ihned vidět, že

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

nebo 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 $x^3 - 1$, nebo tento polynom není ireducibilní. Platí $x^3 - 1 = (x - 1)(x^2 + x + 1)$, ζ_3 je kořenem druhého èinitele, ten je ireducibilní, a tedy $m_{\zeta_3,\mathbb{Q}} = x^2 + x + 1$.

Příklad. Spočteme minimální polynom prvku $a = \sqrt[3]{2} + \sqrt[3]{3}$. Platí

$$a^2 = 5 + 2\sqrt[3]{6}, \quad a^3 = 11\sqrt[3]{2} + 9\sqrt[3]{3}, \quad a^4 = 49 + 20\sqrt[3]{6}$$

a vidíme, že $a^4 = 10a^2 - 1$. Èli a je kořenem polynomu $x^4 - 10x^2 + 1$. Tento polynom je ireducibilní: díky kritériu existence racionálního kořene vidíme, že nemá racionální kořen, a na součin dvou polynomů stupňů 2 se rozkládat nemůže, nebo $\sqrt[3]{2} + \sqrt[3]{3}$ není řešením žádné kvadratické rovnice.

Tvrzení 6.4 (struktura jednoduchých rozšíření). Buď $\mathbf{T} \subseteq \mathbf{S}$ rozšíření těles a $a \in \mathbf{S}$ algebraický prvek nad \mathbf{T} . Pak

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

Důkaz. Podle Tvrzení 5.1 je

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

Dokážeme, že tyto prvky tvoří těleso. Mějme tedy nějaký prvek $0 \neq f(a) \in \mathbf{T}[a]$, hledáme jeho inverz, tedy polynom $g \in \mathbf{T}[x]$ 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 $\text{NSD}(m_{a,\mathbf{T}}, f) = 1$, èli podle Bézoutovy rovnosti existují polynomy $u, g \in \mathbf{T}[x]$ takové, že $1 = um_{a,\mathbf{T}} + gf$. Dosazením prvku a dostáváme

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

èli $g(a)$ je inverzní prvek k $f(a)$.

Alternativní důkaz. Uvažujme homomorismus $\varphi : \mathbf{T}[x] \rightarrow \mathbf{T}[a]$, $f \mapsto f(a)$. Ten je zřejmě na, jeho jádro je ideál $m_{a,\mathbf{T}}\mathbf{T}[x]$, a tak podle 1. věty o izomorfismu $\mathbf{T}[x]/(m_{a,\mathbf{T}}) \cong \mathbf{T}[a]$. Protože je $m_{a,\mathbf{T}}$ ireducibilní, tento ideál je maximální a tudíž je $\mathbf{T}[a]$ těleso, èli $\mathbf{T}[a] = \mathbf{T}(a)$.

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

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

Pro rozšíření vyšších stupňů vycházejí vzorce ošklivě (zkuste si to!) a Tvzení 6.4 má svoji cenu.

Poznámka. Je-li a transcendentní prvek nad \mathbb{T} , pak $\mathbb{T}[a] \neq \mathbb{T}(a)$. Kdyby $\frac{1}{a} \in \mathbb{T}[a]$, pak by existoval polynom $f \in \mathbb{T}[x]$ takový, že $f(a) = a^{-1}$, èli $af(a) = 1$, a tedy a by bylo koèenem polynomu $xf - 1 \in \mathbb{T}[x]$, spor.

Tvzení 6.5 (stupeò jednoduchých rozšíření). Buò $\mathbb{T} \subset \mathbb{S}$ rozšíření těles a $a \in \mathbb{S}$ algebraický prvek nad \mathbb{T} . Pak

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

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

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

Nyní dokážeme, že prvky $1, a, a^2, \dots, a^{n-1}$ generují vektorový prostor $\mathbb{T}(a)_{\mathbb{T}}$. Uvažujme prvek $f(a)$ tělesa $\mathbb{T}(a) = \mathbb{T}[a]$, vyjádøíme jej jako lineární kombinaci. Buò $q, r \in \mathbb{T}[x]$ takové, že $f = q \cdot m_{a,\mathbb{T}} + r$ a $\deg r < \deg m_{a,\mathbb{T}} = n$. Pak

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

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

Příklad. Pomocí Tvzení 6.5 lze uròit stupeò 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 prvoèíslo p , protože uvedený polynom je podle Eisensteinova kritéria ireducibilní. (Pokud p není prvoèíslo, 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. Je-li n prvoèíslo, minimálním polynomem je $x^{n-1} + x^{n-2} + \dots + 1 = \frac{x^n - 1}{x - 1}$, jehož ireducibilitu lze po substituci ukázat z Eisensteinova kritéria.

Důsledek 6.6. Buò $\mathbb{T} \subset \mathbb{S}$ rozšíření těles a $a \in \mathbb{S}$. Prvek a je algebraický nad \mathbb{T} právě tehdy, když je stupeò $[\mathbb{T}(a) : \mathbb{T}]$ 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 $t_i \in \mathbb{T}$, aspoò jeden nenulový, bylo by a koèenem nenulového polynomu $\sum_{i=0}^n t_i x^i \in \mathbb{T}[x]$, spor. Opaèná implikace plyne z Tvzení 6.5.

Příklad. Ukážeme si strukturu tzv. kvadratických rozšíření, tj. rozšíření stupně 2. Dokážeme, že je-li $\mathbb{T} \subset \mathbb{S} \subset \mathbb{C}$ a $[\mathbb{S} : \mathbb{T}] = 2$, pak

$$\mathbb{S} = \mathbb{T}(\sqrt[s]{a}) \text{ pro nějaké } s \in \mathbb{Z}.$$

Buò $1, a$ báze prostoru $\mathbb{S}_{\mathbb{T}}$. Pak $\mathbb{S} = \mathbb{T}(a)$ a podle Tvzení 6.5 je a koèenem nějakého polynomu z $\mathbb{T}[x]$ stupně 2. Známý vzorec na výpočet koèenů kvadratického polynomu říká, že $a = u + v\sqrt[s]{a}$ pro nějaká $u, v, s \in \mathbb{T}$, a tak $\mathbb{S} = \mathbb{T}(u + v\sqrt[s]{a}) = \mathbb{T}(\sqrt[s]{a})$.

6.3. Vícenásobná rozšíření.

K výpočtu stupně vícenásobného rozšíření slouží následující obecné pravidlo.

Tvrzení 6.7 (stupeň vícenásobných rozšíření). Buď $\mathbf{T} \supset \mathbf{S} \supset \mathbf{U}$ rozšíření těles. Pak

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

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 bázi vektorového prostoru \mathbf{U}_T .

Nejprve dokážeme, že C generuje prostor \mathbf{U}_T (jistě $C \subset \mathbf{U}_T$, a tedy C generuje podprostor \mathbf{U}_T). Je-li $u \in \mathbf{U}_T$, pak $u = \sum_j s_j b_j$ pro nějaká $s_j \in \mathbf{S}$ a $b_j \in B$. Každé s_j lze napsat jako $s_j = \sum_i t_{ij} a_i$ pro nějaká $t_{ij} \in \mathbf{T}$ a $a_i \in A$, a dosazením druhé rovnosti do první dostaneme

$$u = \sum_j \left(\sum_i t_{ij} a_i \right) b_j = \sum_{i,j} t_{ij} a_i b_j.$$

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

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

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

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

$$[\mathbf{U} : \mathbf{T}] = |C| = |A| |B| = |\mathbf{S}| |\mathbf{T}| = [\mathbf{U} : \mathbf{S}] [\mathbf{S} : \mathbf{T}].$$

Tvrzení 6.5 a 6.7 můžeme aplikovat na výpočet stupně vícenásobných rozšíření typu $\mathbf{T}(a_1, a_2, \dots)$. Dvojitě rozšíření $\mathbf{T} \supset \mathbf{T}(a, b)$ můžeme rozbit na dvě jednoduchá rozšíření $\mathbf{T} \supset \mathbf{T}(a) \supset \mathbf{T}(a, b)$ a spočítáme

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

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 6.8. Buď $\mathbf{T} \supset \mathbf{S}$ rozšíření těles a $a_1, \dots, a_n \in \mathbf{S}$ 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

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

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

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

$$m_{\sqrt[3]{2}, \mathbb{Q}} = x^3 - 2;$$

$$m_{\sqrt[3]{3}, \mathbb{Q}(\sqrt[3]{2})} = x^3 - 3 \text{ (ověřte, že je opravdu ireducibilní v } \mathbb{Q}(\sqrt[3]{2})[x] \text{ !)}.$$

Podle Tvzení 6.5 a 6.7 dostáváme $[\mathbb{Q}(\sqrt[3]{2} + \sqrt[3]{3}) : \mathbb{Q}] = 4$ a $[\mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{3}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{3}) : \mathbb{Q}(\sqrt[3]{2})] [\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 2 \cdot 2 = 4$.

Je-li $\mathbf{T} \subseteq \mathbf{S}$ rozšíření tělesa a každý prvek tělesa \mathbf{S} je algebraický nad \mathbf{T} , hovoříme o algebraickém rozšíření. Tuto vlastnost mají všechna rozšíření konečného stupně.

Tvrzení 6.9. Rozšíření konečného stupně jsou algebraická.

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

Tvrzení 6.9 je principem nekonstruktivních důkazů algebraičnosti: k důkazu, že je prvek a algebraický nad \mathbf{T} , stačí najít rozšíření $\mathbf{S} \supseteq \mathbf{T}$ konečného stupně, v němž a leží. Typickým příkladem je důkaz, že součet, rozdíl, součin a podíl algebraických prvků je algebraický prvek, jak říká následující věta.

Věta 6.10 (algebraické prvky tvoří podtěleso). Buď $\mathbf{T} \subseteq \mathbf{S}$ rozšíření tělesa. Prvky \mathbf{S} , které jsou algebraické nad \mathbf{T} , tvoří podtěleso tělesa \mathbf{S} .

Důkaz. Uvažujme prvky $a, b \in \mathbf{S}$ algebraické nad \mathbf{T} . Rozšíření $\mathbf{T} \subseteq \mathbf{T}(a, b)$ je konečného stupně (Důsledek 6.8), a tedy algebraické (Tvrzení 6.9). ěilí všechny prvky $\mathbf{T}(a, b)$ jsou algebraické nad \mathbf{T} , speciálně také prvky $a + b, a - b, a \cdot b, a : b$ (pro $a \neq 0$). Tedy algebraické prvky tvoří podtěleso tělesa \mathbf{S} .

7. Neřešitelnost úloh pravítkem a kružítkem

Mezi klasické matematické úlohy s kořený v antickém Řecku patří konstrukce pomocí pravítka a kružítko. Některé úlohy jsou snadné a uě se na základní ěkole: například zdvojení ětverce ěi půlení úhlu. Jsou úlohy, které odolávaly tisíciletí: například konstrukci pravidelného sedmnáctiúhelníka objevil Gauss roku 1796. Už v té době se tuřilo, 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 úsece sestojit úseku, která je $\sqrt[3]{2}$ -krát delší (původní formulace: k dané úsece u sestojit úseku v takovou, že krychle s hranou v má dvakrát větší objem, než krychle s hranou u);

trisekce úhlu: k danému úhlu sestojit třetinový úhel;

rektifikace kružnice a kvadratura kruhu: k dané úsece sestojit sestojit úseku, která je π -krát delší (původní formulace: k dané kružnici k sestojit úseku, která je stejně dlouhá jako obvod k , resp. úseku takovou, že ětverec nad ní sestojený má stejný obsah jako kruh daný k ; obě úlohy lze snadno převést na konstrukci π -krát delší úseky).

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 transcendentnosti čísla π).

Předně musíme upřesnit, co vlastně rozumíme konstrukcí pravítkem a kružítkem. Na začátku je daná jistá konečná množina 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.

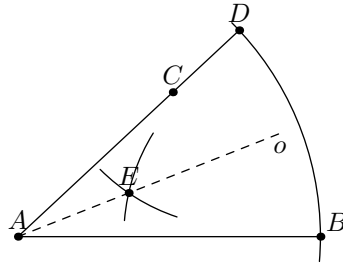
Formálně, konstrukce pravítkem a kružítkem je posloupnost $M_0 \subseteq M_1 \subseteq \dots \subseteq M_n$ konečných množin bodů v rovině taková, že $M_{i+1} = M_i \cup fXg$, kde X vznikne jako

- (1) průsečík přímky AB a přímky CD ;
- (2) průsečík přímky AB a kružnice $k(C, jDEj)$ se středem C a poloměrem $jDEj$;
- (3) průsečík kružnic $k(A, jBCj)$ a $k(D, jEFj)$

pro nějaké body $A, B, C, D, E, F \in M_i$.

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 $\mathbf{T}_i \supseteq \mathbb{R}$, které obsahuje x -ové i y -ové souřadnice všech bodů z M_i . Ěli, pokud M_i obsahuje body A_1, \dots, A_k se souřadnicemi $(a_1, b_1), \dots, (a_k, b_k)$, pak $\mathbf{T}_i = \mathbb{Q}(a_1, b_1, \dots, a_k, b_k)$. Přidáním bodu X se souřadnicemi (u, v) dostaneme $\mathbf{T}_{i+1} = \mathbf{T}_i(u, v)$. Výsledkem je řetězec rozšíření těles $\mathbf{T}_0 \subset \mathbf{T}_1 \subset \mathbf{T}_2 \subset \dots \subset \mathbf{T}_n$.

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



Sestrojíme body

$$D = k(A, jABj) \setminus AC \quad \text{a} \quad E = k(B, jBDj) \setminus k(D, jBDj),$$

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

$$M_0 = \{A, B, C\}, \quad M_1 = M_0 \cup \{D\}, \quad M_2 = M_1 \cup \{E\}.$$

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

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

Stejným krokem Wantzelovy metody je následující vlastnost.

Lemma 7.1. Pro každou konstrukci pravítkem a kružítkem je $[\mathbf{T}_{i+1} : \mathbf{T}_i] \in \{2, 3\}$ 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 $a, b, c, d \in \mathbf{T}_i$, 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 $k(A, jBCj)$ určená body A, B, C se souřadnicemi $(a, b), (c, d), (e, f)$, kde $a, b, c, d, e, f \in \mathbf{T}_i$, 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 je do kvadratické, dostaneme kvadratickou rovnici pro x , jejíž koeficienty jsou z \mathbf{T}_i a řešením je $x^0 = u + v\sqrt{s}$ pro nějaká $u, v, s \in \mathbf{T}_i$. Dosazením do lineární rovnice zjistíme, že $y^0 = u^0 + v^0\sqrt{s}$ pro nějaká $u^0, v^0 \in \mathbf{T}_i$. ěili $\mathbf{T}_{i+1} = \mathbf{T}_i(x^0, y^0) = \mathbf{T}_i(\sqrt{s})$, z čehož plyne, že

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

v závislosti na tom, zda je $\sqrt{s} \in \mathbf{T}_i$ nebo ne. (Proveřte popsany výpočet podrobně a ověřte, že skutečně obě řešení náležejí $\mathbf{T}_i(\sqrt{s})$!)

(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 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

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

(Proveřte popsany výpočet podrobně sami!)

Tvrzení 7.2 (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 6.7 je

$$[\mathbf{T}_n : \mathbf{T}_0] = [\mathbf{T}_n : \mathbf{T}_{n-1}] \cdots [\mathbf{T}_2 : \mathbf{T}_1] [\mathbf{T}_1 : \mathbf{T}_0],$$

což je součin jedniček a dvojek.

Příklad (zdvojení krychle). Zvolme souřadnice tak, že krajní body zadané úsečky (symbolizující hranu krychle) jsou $(0, 0)$ a $(1, 0)$; ěili $\mathbf{T}_0 = \mathbb{Q}$. 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 , ěili $\mathbb{Q}(\sqrt[3]{2}) \subset \mathbf{T}_n$ a podle Tvrzení 6.7 je

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

což je ve sporu s Tvrzením 7.2.

(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 $(1, 0)$; ěili $\mathbf{T}_0 = \mathbb{Q}$. Cílem úlohy je sestrojit úsečku délky π (resp. 2π a $\frac{\pi}{2}$ v původním zadání). ěili transcendentní ěíslo π by mělo být prvkem tělesa \mathbf{T}_n , ale to je podle Tvrzení 7.2 rozšířením \mathbb{Q} konečného stupně, a tedy podle Tvrzení 6.9 obsahuje pouze algebraická ěísla, 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 $(\frac{1}{2}, \frac{\sqrt{3}}{2})$; ěili $\mathbf{T}_0 = \mathbb{Q}(\sqrt{3})$. Dokážeme, že není možné sestrojit bod

$$(\cos 20^\circ, \sin 20^\circ).$$

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

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

můžeme použít stejný argument jako pro zdvojení krychle. K tomuto cíli stačí podle Tvrzení 6.5 nalézt minimální polynom ěísla $\cos 20^\circ$ nad tělesem $\mathbb{Q}(\sqrt{3})$, tj. nějaký

ireducibilní polynom, jehož je číslo $\cos 20^\circ$ kořenem. Prolistujeme-li nějakou sbírku goniometrických vzorců, najdeme vztah

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

z kterého plyne, že $\cos 20^\circ$ je kořenem polynomu $4x^3 - 3x - \frac{1}{2} \in \mathbb{Q}[x]$. Tento polynom je v $\mathbb{Q}(\sqrt[3]{3})[x]$ ireducibilní, neboť nemá v $\mathbb{Q}(\sqrt[3]{3})$ kořen (jak snadno zjistíme dosazením $x = a + b\sqrt[3]{3}$). Tedy

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

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

8. Izomorfismy kořenových a rozkladových nadtěles

Definice. Buď \mathbf{T} těleso a f polynom z $\mathbf{T}[x]$ stupně $n \geq 1$.

Kořenovým nadtělesem pro f nad \mathbf{T} rozumíme minimální rozšíření, ve kterém má polynom f kořen. Jinými slovy, nadtěleso \mathbf{S} , kde existuje $a \in \mathbf{S}$ takové, že $\mathbf{S} = \mathbf{T}(a)$ a $f(a) = 0$.

Rozkladovým nadtělesem pro f nad \mathbf{T} rozumíme minimální rozšíření, kde se f rozkládá na lineární faktory. Jinými slovy, nadtěleso \mathbf{S} , kde existují $a_1, \dots, a_n \in \mathbf{S}$ taková, že $\mathbf{S} = \mathbf{T}(a_1, \dots, a_n)$ a $f(x) = (x - a_1) \dots (x - a_n)$.

V zimním semestru jsme dokázali, že taková tělesa existují pro každé těleso \mathbf{T} a každý polynom f (kořenové se zkonstruuje jako faktorokruh $\mathbf{T}[\alpha]/(f(\alpha))$, rozkladové pak rekurzí aplikovanou na polynom $f/(x - \alpha)$). V této sekci si dokážeme jednoznačnost rozkladových nadtěles, ať už na izomorfním smyslu.

Příklad. Díky základní větě algebry víme, že kořenové i rozkladové nadtěleso polynomu f nad tělesem \mathbb{C} lze nalézt uvnitř tělesa \mathbb{C} : kořenovým bude libovolné \mathbb{C} , kde a je nějaký komplexní kořen f , a rozkladovým bude $\mathbb{C}(a_1, \dots, a_m)$, 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 $\mathbb{C}(i) = \mathbb{C}$, které obsahuje oba kořeny $\pm i$, a tedy je i nadtělesem rozkladovým.

Označme $\zeta = e^{2\pi i/3}$.

Uvažujme polynom $x^3 - 1$. Tento polynom má dvě různá kořenová nadtělesa v \mathbb{C} , a to $\mathbb{C} = \mathbb{C}(1)$ a $\mathbb{C}(\zeta) = \mathbb{C}(\zeta^2)$. Tato tělesa jistě nejsou izomorfní. To vidíme, že 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 $\mathbb{Q}(\sqrt[3]{2}\zeta)$ (to druhé obsahuje oba imaginární kořeny). Až to není vidět na první pohled, tato tělesa jsou izomorfní. Rozkladovým nadtělesem pak bude těleso $\mathbb{Q}(\sqrt[3]{2}, \sqrt[3]{2}\zeta) = \mathbb{Q}(\sqrt[3]{2}, \zeta)$.

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 dokážeme izomorfnost také, tentokrát již bez předpokladu ireducibility. Než začneme, uvedeme jednu pomocnou lemmu.

Definice. Buď $\mathbf{T} \subseteq \mathbf{S}, \mathbf{U}$ tři tělesa. \mathbf{T} -izomorfním smyslem $\mathbf{S} \cong \mathbf{U}$ rozumíme izomorfnost φ , pro který platí $\varphi(t) = t$ pro každé $t \in \mathbf{T}$.

Věta 8.1 (jednoznačnost kořenových a rozkladových nadtěles). Buď \mathbf{T} těleso a $f \in \mathbf{T}[x]$ stupně $n \geq 1$.

- (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.

Věta 8.1 je speciálním případem o trochu obecnějších Lemmat 8.2 a 8.3 o rozlišování ěastejných izomor smů. K jejich formulaci je potřeba následující značení a pozorování.

Buď $\mathbf{T} = \mathbf{T}_1, \mathbf{T} = \mathbf{T}_2$ rozlišování těles a $\varphi : \mathbf{T}_1 \rightarrow \mathbf{T}_2$ \mathbf{T} -izomor smus. Zobrazení φ lze rozlišit na \mathbf{T} -izomor smus oborů polynomů nad tělesy (budeme jej opít značit φ):

$$\varphi : \mathbf{T}_1[x] \rightarrow \mathbf{T}_2[x], \quad \sum a_i x^i \mapsto \sum \varphi(a_i) x^i.$$

Oznaěme $f = \sum a_i x^i, g = \sum b_i x^i$. Koeficienty souětu $f + g$ jsou $a_i + b_i$, koeficienty souětu $\varphi(f) + \varphi(g)$ jsou $\varphi(a_i) + \varphi(b_i) = \varphi(a_i + b_i)$ a vidíme, ěe $\varphi(f + g) = \varphi(f) + \varphi(g)$. Koeficienty souěinu fg jsou $\sum_{i+j=k} a_i b_j$, koeficienty souěinu $\varphi(f)\varphi(g)$ jsou $\sum_{i+j=k} \varphi(a_i)\varphi(b_j) = \varphi(\sum_{i+j=k} a_i b_j)$ a vidíme, ěe $\varphi(fg) = \varphi(f)\varphi(g)$. Bijektivita zobrazení je zřejmá. Okamžitým dśsledkem souěinové vlastnosti je, ěe

$f \mid g$ v $\mathbf{T}_1[x]$ právě tehdy, kdyě $\varphi(f) \mid \varphi(g)$ v $\mathbf{T}_2[x]$;
 polynom f je ireducibilní v $\mathbf{T}_1[x]$ právě tehdy, kdyě $\varphi(f)$ je ireducibilní v $\mathbf{T}_2[x]$.

Lemma 8.2. Buď $\mathbf{T} = \mathbf{T}_1, \mathbf{T} = \mathbf{T}_2$ rozlišování těles a $\varphi : \mathbf{T}_1 \rightarrow \mathbf{T}_2$ \mathbf{T} -izomor smus. Buď $f \in \mathbf{T}_1[x]$ 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} -izomor smus $\psi : \mathbf{T}_1(a) \rightarrow \mathbf{T}_2(b)$ takový, ěe $\psi(a) = b$ a $\psi|_{\mathbf{T}_1} = \varphi$.

Dśkaz. Podle tvrzení 6.4 je $\mathbf{T}_1(a) = \mathbf{T}_1[a] = \mathbf{T}_1[g(a) : g \in \mathbf{T}_1[x]]$ a $\mathbf{T}_2(b) = \mathbf{T}_2[b] = \mathbf{T}_2[g(b) : g \in \mathbf{T}_2[x]]$. Uvaěujme tedy zobrazení

$$\psi : \mathbf{T}_1(a) \rightarrow \mathbf{T}_2(b), \quad g(a) \mapsto \varphi(g)(b).$$

Předně je třeba dokázat, ěe to je dobře definované zobrazení. Uvědomte si, ěe $f = m_{a, \mathbf{T}_1}$, protoěe f je ireducibilní polynom a a je jeho kořen, a zrovna tak $\varphi(f) = m_{b, \mathbf{T}_2}$, protoěe $\varphi(f)$ je ireducibilní polynom a b je jeho kořen. ěili

$$g(a) = h(a), \quad (g - h)(a) = 0, \quad f \mid (g - h)$$

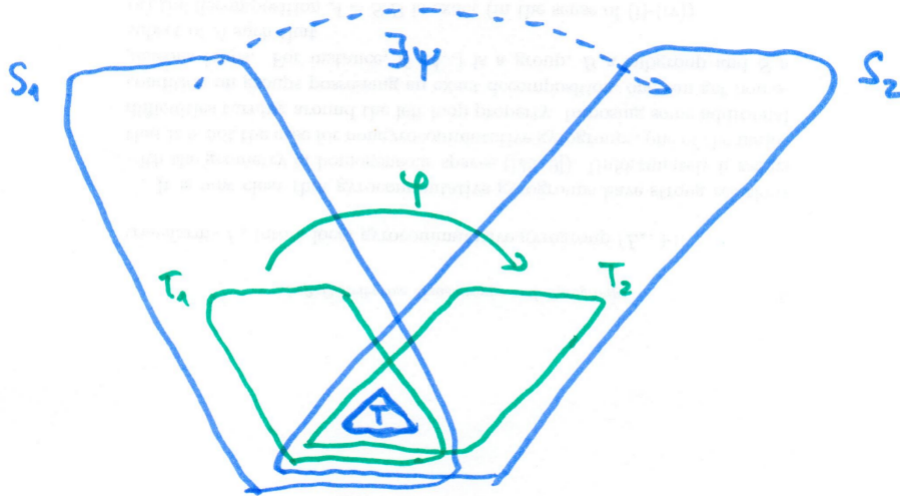
a analogicky

$$\varphi(g)(b) = \varphi(h)(b), \quad \varphi(g - h)(b) = 0, \quad \varphi(f) \mid \varphi(g - h).$$

Ekvivalence obou tvrzení na pravé straně plyne z pozorování výše. Dokázali jsme, ěe φ je dobře definované zobrazení a navíc prostě. Oěividně jde o bijekci a je snadné ovědit, ěe jde o okruhový homomor smus: pro každě $g, h \in \mathbf{T}_1[x]$ platě $\varphi(g(a) + h(a)) = \varphi((g + h)(a)) = \varphi(g + h)(b) = \varphi(g)(b) + \varphi(h)(b) = \varphi(g(a)) + \varphi(h(a))$ a analogicky pro násobení. Prvky tělesa \mathbf{T}_1 odpovídají volbě konstantního polynomu c , pro takový polynom platě $\varphi(c) = \varphi(c(a)) = \varphi(c)(b) = \varphi(c)$, ěili $\varphi|_{\mathbf{T}_1} = \varphi$. Volbou $g = x$ ověděme, ěe $\psi(a) = b$.

Lemma 8.3. Buď $\mathbf{T} = \mathbf{T}_1, \mathbf{T} = \mathbf{T}_2$ rozlišování těles a $\varphi : \mathbf{T}_1 \rightarrow \mathbf{T}_2$ \mathbf{T} -izomor smus. Buď $f \in \mathbf{T}_1[x]$ polynom stupně $n > 1$ a oznaěme \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} -izomor smus $\psi : \mathbf{S}_1 \rightarrow \mathbf{S}_2$ takový, ěe $\psi|_{\mathbf{T}_1} = \varphi$.

Dśkaz. Budeme postupovat indukcě 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 indukěním kroku uvaěujme ireducibilní dělitel g polynomu f a jeho kořen a v \mathbf{S}_1 . Pak $\varphi(g)$ je ireducibilní dělitel polynomu $\varphi(f)$ a uvaěujme jeho kořen b v \mathbf{S}_2 . Podle Lemmatu 8.2 existuje zobrazení $\rho : \mathbf{T}_1(a) \rightarrow \mathbf{T}_2(b)$



Obrázek 7. Ilustrace důkazu jednoznačnosti rozkladového nadřítlesa.

$\mathbf{T}_2(b)$ takové, že $\rho(a) = b$ a $\rho|_{\mathbf{T}_1} = \varphi$. Napišme $f = (x - a)h$ pro nějaký $h \in \mathbf{T}_1[x]$, čili také $\rho(f) = (x - b)\rho(h)$. Pak \mathbf{S}_1 je rozkladové nadřítleso polynomu h nad $\mathbf{T}_1(a)$ a \mathbf{S}_2 je rozkladové nadřítleso polynomu $\rho(h)$ nad $\mathbf{T}_2(b)$. Protože $\deg h < \deg f$, podle indukčního předpokladu existuje \mathbf{T} -izomorfismus $\psi : \mathbf{S}_1 \rightarrow \mathbf{S}_2$ takový, že $\psi|_{\mathbf{T}_1(a)} = \rho$, čili také $\psi|_{\mathbf{T}_1} = \varphi$.

Volbou $\mathbf{T}_1 = \mathbf{T}_2 = \mathbf{T}$ a $\varphi = id$ v obou lemmatech dostaneme Větu 8.1.

9. Klasifikace konečných řítles

9.1. Frobeniův endomorfismus.

Začneme příkladem homomorfismu, který hraje zásadní roli v okruzích nenulové charakteristiky.

Tvrzení 9.1 (Frobeniův endomorfismus). Buď \mathbf{R} komutativní okruh s jednotkou prvočíselné charakteristiky p a dejme zobrazení

$$\varphi_p : \mathbf{R} \rightarrow \mathbf{R}, \quad a \mapsto a^p.$$

- (1) Zobrazení φ_p je homomorfismus.
- (2) Je-li \mathbf{R} obor integrity, pak je φ_p prosté.
- (3) Je-li \mathbf{R} konečné řítleso, 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 směrem každého středokoláka, ale platí pouze za předpokladu prvočíselné charakteristiky p .

Důkaz. (1) Zřejmě $(a - b)^p = a^p - b^p$ a podle binomické věty

$$(a + b)^p = \sum_{i=0}^p \binom{p}{i} a^i b^{p-i} = a^p + b^p,$$

nebože p dělí všechny binomické koeficienty $\binom{p}{i}$, $i = 1, \dots, p-1$, protože všechny 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, ěli je prostý (Tvrzení 4.2).

(3) Prosté zobrazení na konečné množině je bijektivní.

9.2. Derivace a násobné kořeny.

Buď \mathbf{R} obor integrity. Buď $a \in \mathbf{R}$ a $f \in \mathbf{R}[x]$. Pŕipomeďme, že prvek a je kořenem f právě tehdy, když $x = a$ je kořen f . Kořen a nazveme vícenásobný, pokud $(x - a)^2 \mid f$.

Derivací polynomu $f = \sum_{i=0}^n a_i x^i$ rozumíme polynom $f' = \sum_{i=0}^{n-1} (i+1)a_{i+1}x^i$. Není třeba ověřit (cvičení!), že takto definovaná derivace splňuje všechny základní vzorečky, které znáte z kurzu analýzy. Obor \mathbf{R} může být jakýkoliv, žádné vlastnosti reálných ěísel se k tomu nepoužijí, pŕedstava derivace jako smírnicice teeny není potřeba.

Pro důkaz klasické lemmaty o násobných kořenech se nám bude hodit následující pozorování.

Lemma 9.2. Nechť polynom f má vícenásobný kořen a . Pak $f'(a) = 0$.

Důkaz. Napište $f = (x - a)^2 g$ a použijte součinný vzorec na výpočet derivace: $f' = 2(x - a)g + (x - a)^2 g'$, ěli $f'(a) = 0$.

9.3. Klasifikace konečných těles.

Aplikací vĕt o existenci a jednoznačnosti rozkladových nadtěles ukážeme, že pro každou mocninu prvočísla p^k existuje, až na izomorfiism, právě jedno těleso velikosti p^k . Princip důkazu je v tom, že těleso má pŕesně 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 9.3. 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 9.1 o Frobeniově endomorfismu říká, že zobrazení $\varphi : a \mapsto a^p$ je homomorfismus smem $\mathbf{T} \rightarrow \mathbf{T}$. Jeho k -násobné složení, φ^k , je také homomorfismus a zobrazuje $a \mapsto ((a^p)^p) \dots = a^{p^k} = a^q$, ěli

$$(a + b)^q = a^q + b^q \quad \text{a} \quad (a \cdot b)^q = a^q \cdot b^q$$

pro každé $a, b \in \mathbf{T}$. 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 stejnĕ tak $(a \cdot b)^q = a^q \cdot b^q = a \cdot b$, $(a^{-1})^q = a^{-q} = a^{-1}$ a $(a^{-1})^q = (a^q)^{-1} = a^{-1}$. ěli 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 pŕesně q prvků, stačí ověřit, že polynom f nemá vícenásobné kořeny. Kdyby byl prvek a vícenásobným kořenem, podle Lemmatu 9.2 by platilo $f'(a) = 0$. Ovšem $f' = qx^{q-1} - 1 = -1$, a tedy žádné kořeny nemá.

Lemma 9.4. Buď \mathbf{T} konečné těleso, $f(\mathbf{T}) = p^k$. Pak \mathbf{T} je rozkladovým nadtělesem polynomu $x^{p^k} - x$ nad tělesem \mathbb{Z}_p a v $\mathbf{T}[x]$ platí

$$x^{p^k} - x = \prod_{a \in \mathbf{T}} (x - a).$$

Důkaz. Označme $q = p^k$. Nejprve si všimnĕte, že každý prvek $a \in \mathbf{T}$ je kořenem polynomu $f = x^q - x$. Pro 0 to platí triviálně a pro nenulový prvek a využijeme Lagrangeovu vĕtu: $\text{ord}(a) \mid q - 1$, ěli $a^{q-1} = 1$ a $a^q = a$. Tedy $\prod_{a \in \mathbf{T}} (x - a) \mid f$, a z rovnosti stupňů i vedoucích koeficientů dostáváme rovnost těchto polynomů. Ukázali jsme, že f se v \mathbf{T} rozkládá na lineární ěinitele. Je to ale minimální rozložení? Ano je, neboť dle pŕedchozího lemmaty má rozkladové nadtěleso polynomu f právě q prvků, ěli \mathbf{T} je tímto tělesem.

Věta 9.5 (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) () plyne z Tvzení 5.3, () plyne z Lemmatu 9.3 a věty o existenci rozkladových nadtěles. (2) plyne z Lemmatu 9.4 a Věty 8.1 o jednoznačnosti rozkladových nadtěles.

V zimním semestru jsme představili konečná tělesa ve formě faktorokruhů $Z_p[\alpha]/(m)$. Lze každé konečné těleso tímto způsobem reprezentovat?

Věta 9.6 (reprezentace konečných těles). Pro každé prvočíslo p a přirozené číslo k existuje ireducibilní polynom $m \in Z_p[\alpha]$ stupně k a

$$F_{p^k} \cong Z_p[\alpha]/(m).$$

Důkaz. Podle Věty 9.5 existuje nějaké těleso \mathbf{T} Z_p velikosti p^k . V kapitole o cyklických grupách jsme si dokázali, že grupa \mathbf{T} je cyklická. Označme a nějaký generátor a uvažujme minimální polynom m_{a,Z_p} . Ten je jistě ireducibilní a jeho stupeň je

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

protože první rovnost plyne z Tvzení 6.5, druhá z faktu, že $\mathbf{T} = Z_p(a)$, 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 9.5 plyne, že $\mathbf{T} \cong Z_p[\alpha]/(m_{a,Z_p})$.

Všimněte si, jakým obratem jsme prokázali existenci ireducibilního polynomu stupně k v $Z_p[x]$: nejprve jsme prokázali existenci nějakého tělesa velikosti p^k , abychom mohli vzít generátor jeho multiplikativní grupy a jeho minimální polynom. První důkaz existence těchto polynomů je možný, ale mnohem techničtější a dává menší vhled do celé situace.

Algoritmy polynomiální aritmetiky

10. Modulární reprezentace

Cílem této sekce je pochopit princip modulární reprezentace a seznámit se s rychlým algoritmem, který ji počítá, tzv. rychlou Fourierovou transformací.

Principem modulární reprezentace je následující myšlenka: místo jednoho výpočtu ve velkém oboru (celá čísla, polynomy) provedeme několik výpočtů v menších oborech (čísla modulo m , obor koeficientů) a z výsledků zrekonstruujeme řešení původní úlohy. Chytré použití této myšlenky vede u některých úloh k překvapivě rychlým algoritmům.

Modulární reprezentací oboru \mathbf{R} , na kterém je definované dělení se zbytkem, rozumíme libovolný izomorfismus

$$\mathbf{R}/(m) \cong \mathbf{R}/(m_1) \times \dots \times \mathbf{R}/(m_n) \\ a \mapsto (a \bmod m_1, \dots, a \bmod m_n)$$

pro nějaká $m_1, \dots, m_n \in \mathbf{R}$ po dvou nesoudělná a $m = m_1 \dots m_n$. Fakt, že jde o izomorfismus, plyne ze zobecněné čínské věty o zbytcích. Tu jsme nedělali v plné obecnosti, nicméně dělali jsme dva speciální případy, které pokrývají všechny myslitelné aplikace: případ $\mathbf{R} = \mathbb{Z}$ a případ $\mathbf{R} = \mathbb{T}[x]$ pro nějaké těleso \mathbb{T} .

Vimniťte si, že modulární reprezentace nereprezentuje vůbec celý obor, ale pouze jeho část, danou prvkem m (tj. čísla $< m$, polynomy stupně menšího než $\deg m$). V praxi volíme tolik prvků m_1, \dots, m_n , kolik je potřeba k věrné reprezentaci celého výpočtu.

Algoritmus pro převod do modulární reprezentace je zřejmý, dělit se zbytkem je v principu snadné. Ale přesto, základní algoritmy mají kvadratickou složitost, což pro některé aplikace nemusí být dostačující.

Převod zpět znamená řešit soustavu lineárních kongruencí. Obecný algoritmus jsme měli v zimním semestru, příklady na cvičeních. Není třeba si spočítat, že jde také o algoritmus s kvadratickou složitostí.

Nešlo by to rychleji? Obecně ne, ale ve speciálních případech ano. Dál se budeme soustředít na obory $\mathbb{T}[x]$ a speciální případ polynomů $m_i = x - u_i$, kde u_i jsou po dvou různé prvky. V tom případě jde o dosazovací homomorfismus, $f \mapsto (f(u_1), \dots, f(u_n))$ a potřebovali bychom rychlý algoritmus jak na dosazování hodnot, tak na interpolaci. To je zdánlivě nemožné, vždy máme n hodnot a polynom má n členů, takže jak se dostat pod kvadratický čas? Řešením je zvolit chytře prvky u_i .

10.1. Diskrétní Fourierova transformace.

V celé sekci budeme uvažovat nějaké těleso \mathbb{T} . Źekneme, že prvek $\omega \in \mathbb{T}$ je primitivní n -tá odmocnina z jedné, jestliže je řád prvku ω v grupě \mathbb{T}^\times roven n . Jinými slovy, pokud platí

- (1) $\omega^n = 1$,
- (2) $\omega^i \neq 1$ pro všechna $i = 1, 2, \dots, n-1$.

Dvě poznámky: díky Lagrangeovi víme stačí podmínku (2) testovat pouze pro $i \leq n$; z této podmínky také plyne, že $\omega^i \neq \omega^j$ pro všechna $i \neq j < n$.

Pro reprezentaci polynomů stupně $< n$ nad tělesem \mathbb{T} budeme uvažovat hodnoty v bodech $1, \omega, \omega^2, \dots, \omega^{n-1}$, kde ω je primitivní n -tá odmocnina z jedné v tělese \mathbb{T} .

Takové reprezentaci se říká diskretní Fourierova transformace a rychlému algoritmu na její výpočet se říká rychlá Fourierova transformace. Za jistých předpokladů dostaneme algoritmus s časovou složitostí $O(n \log n)$.

Příklad.

Těleso \mathbb{C} obsahuje primitivní n -tou odmocninu z jedné pro každé n , např.

$$\omega = e^{\frac{2\pi i}{n}} = \cos(2\pi/n) + i \sin(2\pi/n).$$

Toto číslo generuje cyklickou grupu, její prvky jsou reprezentovány vrcholy pravidelného n -úhelníka na jednotkové kružnici v komplexní rovině. Každý její generátor je primitivní n -tou odmocninou.

Těleso \mathbb{Q} obsahuje pouze primitivní druhou odmocninu z jedné, prvek -1 . Jiné primitivní odmocniny v něm nejsou.

Těleso \mathbb{F}_q obsahuje primitivní n -tou odmocninu z jedné právě tehdy, když $n \mid q-1$: v kapitole o cyklických grupách jsme si dokázali, že grupa \mathbb{F}_q je cyklická, má $q-1$ prvků, a tedy obsahuje prvek libovolného řádu $n \mid q-1$.

Definice. Diskretní Fourierovou transformací v bodě ω nazveme zobrazení $DFT_\omega : \mathbb{T}^n \rightarrow \mathbb{T}^n$ dle nově daného předpisem

$$DFT_\omega(a_0, \dots, a_{n-1}) = (f(\omega^0), f(\omega^1), \dots, f(\omega^{n-1})),$$

kde $f = \sum_{i=0}^{n-1} a_i x^i$.

Hodnotu polynomu $f = \sum_{i=0}^{n-1} a_i x^i$ v bodě α lze zapsat jako maticový součin

$$f(\alpha) = \sum_{i=0}^{n-1} a_i \alpha^i = (1, \alpha, \dots, \alpha^{n-1}) \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_{n-1} \end{pmatrix}.$$

Dosazujeme-li obecně n hodnot $\alpha_1, \dots, \alpha_n$, dostáváme vyjádření

$$\begin{pmatrix} f(\alpha_1) \\ f(\alpha_2) \\ \dots \\ f(\alpha_n) \end{pmatrix} = \begin{pmatrix} \alpha_1^0 & \alpha_1^1 & \alpha_1^2 & \dots & \alpha_1^{n-1} \\ \alpha_2^0 & \alpha_2^1 & \alpha_2^2 & \dots & \alpha_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ \alpha_n^0 & \alpha_n^1 & \alpha_n^2 & \dots & \alpha_n^{n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_{n-1} \end{pmatrix}.$$

Tedy modulární reprezentace v bodech $\alpha_1, \dots, \alpha_n$ je vlastně lineární zobrazení (endomorfismus) vektorového prostoru \mathbb{T}^n

$$\varphi : \mathbb{T}^n \rightarrow \mathbb{T}^n, \quad u \mapsto A u,$$

kde $u = (a_0, a_1, \dots, a_{n-1})^T$ je sloupcový vektor koeficientů a A je výše uvedená tzv. Vandermondova matice. Tato matice je regulární právě tehdy, když jsou prvky α_i po dvou různé. Diskretní Fourierova transformace v bodě ω je tedy lineárním zobrazením

$$DFT_\omega : \mathbb{T}^n \rightarrow \mathbb{T}^n, \quad u \mapsto A_\omega u,$$

kde A_ω je Vandermondova matice odpovídající prvkům $\omega^0, \omega^1, \dots, \omega^{n-1}$, tj.

$$A_\omega = (\omega^{ij})_{i,j=0}^{n-1}.$$

Protože jde o bijekci, můžeme uvažovat inverzní zobrazení DFT_ω^{-1} , kterému se říká inverzní DFT a značí se $IDFT_\omega$. Jde vlastně o interpolaci v uvedených bodech. Všimněte si, že

$$IDFT_\omega(u) = A_\omega^{-1} u.$$

Následující tvrzení ukazuje, že $IDFT$ je speciálním případem DFT . To nám umožní se soustředit jen na jeden algoritmus pro převod mezi reprezentacemi, který budeme používat pro cestu tam i zpět.

Tvrzení 10.1. Je-li ω primitivní n -tá odmocnina z jedné v \mathbb{T} , pak

$$(A_\omega)^{-1} = \frac{1}{n} A_{\omega^{-1}}.$$

Aby tvrzení vůbec dávalo smysl, musíme si uvědomit, že charakteristika tělesa \mathbb{T} nedělí n . To lze odvodit z faktu, že v \mathbb{T} existuje primitivní n -tá odmocnina z jedné, viz cvičení.

Důkaz. Stačí dokázat, že $A_\omega \left(\frac{1}{n} A_{\omega^{-1}}\right)$ je jednotková matice. Protože

$$A_\omega = (\omega^{ij})_{i,j=0}^{n-1} \quad \text{a} \quad A_{\omega^{-1}} = (\omega^{-ij})_{i,j=0}^{n-1},$$

podle vzorce pro součin matic platí

$$A_\omega A_{\omega^{-1}} = \left(\sum_{k=0}^{n-1} \omega^{ik} \omega^{-kj} \right)_{i,j=0}^{n-1}.$$

Pro $i = j$ dostáváme hodnotu

$$\sum_{k=0}^{n-1} \omega^{ik} \omega^{-ki} = \sum_{k=0}^{n-1} 1 = n$$

a pro $i \neq j$ hodnotu

$$\sum_{k=0}^{n-1} \omega^{ik} \omega^{-kj} = \sum_{k=0}^{n-1} \omega^{k(i-j)} = \sum_{k=0}^{n-1} (\omega^{i-j})^k,$$

což je součet geometrické řady o základu $\omega^{i-j} \neq 1$ (protože ω je primitivní n -tá odmocnina a $j \neq i$ ($j < n$)), a tedy

$$\sum_{k=0}^{n-1} (\omega^{i-j})^k = \frac{(\omega^{i-j})^n - 1}{\omega^{i-j} - 1} = \frac{1 - 1}{\omega^{i-j} - 1} = 0,$$

nebo $\omega^n = 1$. Dokázali jsme, že na diagonále jsou prvky n a mimo diagonálu nuly, tedy po přenásobení $\frac{1}{n}$ dostáváme jednotkovou matici.

Jinými slovy, dokázali jsme, že

$$\text{IDFT}_\omega = \frac{1}{n} \text{DFT}_{\omega^{-1}},$$

bude nám tedy stačit jediný algoritmus na výpočet diskretní Fourierovy transformace i jejího inverzu.

10.2. Rychlá Fourierova transformace.

Rychlou Fourierovou transformací (FFT) se rozumí rychlý algoritmus na výpočet DFT. Jeho principem je opět metoda prozdíl a panuj. Myšlenka je následující: dosazujeme-li hodnotu α do polynomu $f = \sum_{i=0}^{n-1} a_i x^i$ lichého stupně (tj. n sudé), můžeme psát

$$f(\alpha) = \underbrace{(a_0 + a_2 \alpha^2 + \dots + a_{n-2} \alpha^{n-2})}_{g(\alpha^2)} + \alpha \underbrace{(a_1 + a_3 \alpha^2 + a_5 \alpha^4 + \dots + a_{n-1} \alpha^{n-2})}_{h(\alpha^2)},$$

tj.

$$f(\alpha) = g(\alpha^2) + \alpha h(\alpha^2),$$

kde g, h jsou polynomy polovičního stupně de nované

$$g = \sum_{i=0}^{\frac{n}{2}-1} a_{2i} x^i \quad \text{a} \quad h = \sum_{i=0}^{\frac{n}{2}-1} a_{2i+1} x^i.$$

Tedy úlohu dosazení hodnoty α do polynomu s n koeficienty jsme rozdělili na dvě úlohy dosazení hodnoty α^2 do polynomů poloviční velikosti.

V DFT ale dosazujeme n -tici hodnot $\omega^0, \dots, \omega^{n-1}$. Druhá část triku spočívá v tom, že dosazovat stačí pouze hodnoty $\omega^0, \dots, \omega^{n/2-1}$. Pro primitivní n -tou odmocninu z jedné totiž platí

$$\omega^{n/2+i} = \omega^{n/2} \omega^i = (-1) \omega^i = -\omega^i,$$

takže v druhé mocnině jsou oba prvky stejné a dosazovat jich stačí jen polovinu. Původní úlohu jsme tak skutečně rozdělili na dvě poloviční: do polovičních polynomů dosazujeme polovinu hodnot.

Algoritmus 1 (rychlá Fourierova transformace, FFT).

vstup: $n = 2^k$, ω primitivní n -tá odmocnina z jedné, a_0, a_1, \dots, a_{n-1}

výstup: $DFT_\omega(a_0, a_1, \dots, a_{n-1})$

0. **if** $n = 1$ **then return** a_0

1. $(b_0, \dots, b_{\frac{n}{2}-1}) := FFT(n/2, \omega^2, a_0, a_2, \dots, a_{n-2})$

$(c_0, \dots, c_{\frac{n}{2}-1}) := FFT(n/2, \omega^2, a_1, a_3, \dots, a_{n-1})$

2. $d_i := b_i + \omega^i c_i$, $d_{\frac{n}{2}+i} := b_i - \omega^i c_i$ pro všechna $i = 0, \dots, \frac{n}{2}-1$

return (d_0, \dots, d_{n-1})

Tvrzení 10.2. Algoritmus 1 funguje.

Důkaz. Důkaz provedeme indukcí podle n . Pro $n = 1$ je DFT_ω dosazení do konstantního polynomu s koeficientem a_0 , tedy výsledek je a_0 . Provedeme indukční krok. Přední je třeba si uvědomit, že ω^2 je primitivní $\frac{n}{2}$ -tá odmocnina z jedné: zřejmě $(\omega^2)^{n/2} = \omega^n = 1$ a dále pro všechna $i = 1, 2, \dots, \frac{n}{2}-1$ platí $(\omega^2)^i = \omega^{2i} \neq 1$, protože $2i < n$ a ω je primitivní odmocnina.

Nechť tedy $f = \sum_{i=0}^{n-1} a_i x^i$ a dejme polynomy g, h jako výše. Podle indukčního předpokladu

$$(b_0, \dots, b_{\frac{n}{2}-1}) = (g(1), g(\omega^2), g(\omega^4), \dots, g(\omega^{n-2})),$$

$$(c_0, \dots, c_{\frac{n}{2}-1}) = (h(1), h(\omega^2), h(\omega^4), \dots, h(\omega^{n-2})).$$

Chceme dokázat, že pro $i = 0, 1, \dots, \frac{n}{2}-1$ platí

$$d_i = f(\omega^i) \quad \text{a} \quad d_{i+\frac{n}{2}} = f(\omega^{i+n/2}).$$

První vztah plyne přímo z výše odvozeného vzorce:

$$f(\omega^i) = g(\omega^{2i}) + \omega^i h(\omega^{2i}) = b_i + \omega^i c_i = d_i.$$

Druhý vztah dostaneme podobně

$$f(\omega^{i+n/2}) = g(\omega^{2i+n}) + \omega^{i+n/2} h(\omega^{2i+n}) = b_i - \omega^i c_i = d_{i+\frac{n}{2}}$$

využitím snadného pozorování, že $\omega^{2i+n} = \omega^{2i} \omega^n = \omega^{2i}$ a že $\omega^{i+n/2} = \omega^i \omega^{n/2} = \omega^i$. Zde $\omega^{n/2} = -1$, protože to je druhá odmocnina z jedné, a ty jsou pouze dvě: 1 (ta to není, nebo ω je primitivní) a -1 .

Tvrzení 10.3. Algoritmus 1 má časovou složitost $O(n \log n)$.

Důkaz. Budeme postupovat podle osvědčeného schématu pro algoritmy typu rozdíl a panuj. Označme $T(n)$ počet operací v tělese \mathbb{T} , které algoritmus provede na vstupu délky n . Pak $T(1) = 0$ a

$$T(n) = 2T(n/2) + O(n),$$

tedy $T(n) = 2T(n/2) + cn$ pro nějaké c a dosazením $n = 2^k$ dostaneme

$$\begin{aligned} T(2^k) &= 2T(2^{k-1}) + c2^k = 2(2T(2^{k-2}) + c2^{k-1}) + c2^k \\ &= 4T(2^{k-2}) + c(2^k + 2^k) \\ &\dots \\ &= 2^k T(2^{k-k}) + ck2^k = 2^k T(1) + ck2^k = O(k2^k). \end{aligned}$$

Tedy $T(n) = O(n \log n)$.

Příklad. Uvažujme $\mathbb{T} = \mathbb{Z}_{41}$, spočteme modulární reprezentaci polynomu

$$5x^3 + x + 1.$$

Nejprve najdeme primitivní čtvrtou odmocninu z jedné: např. $\omega = 9$ (vidíme, že $\omega^2 = -1$, $\omega^3 = 9$ a $\omega^4 = 1$). Dále počítáme $DFT_{-9}(1, 1, 0, 5)$. Úlohu rozdělíme na

$$DFT_{-9}(1, 0) = (1, 1) \quad \text{a} \quad DFT_{-9}(1, 5) = (6, -4).$$

Výsledek bude

$$(1 + (-9)^0 \cdot 6, 1 + (-9)^1 \cdot (-4), 1 + (-9)^0 \cdot 6, 1 + (-9)^1 \cdot (-4)) = (7, -4, 5, 6).$$

10.3. Primitivní odmocniny z jedné.

Pro FFT je klíčová existence primitivní odmocniny z jedné. Pokud neexistuje v tělese \mathbb{T} , budeme muset počítat v nějakém jiném tělese.

Pro konečný těleso F_q je situace jasná: primitivní n -tá odmocnina v F_q existuje právě tehdy, když $n \mid q-1$ a získáme ji jako mocninu $\omega = a^{(q-1)/n}$, kde a je generátor cyklické grupy F_q . Ten získáme náhodnou volbou a (pravděpodobnost úspěchu je $\varphi(q-1)/(q-1)$, což je typicky poměrně velký zlomek. (Viz zimní semestr.)

Rozebereme si podrobně případ tělesa \mathbb{Q} . Tam bohužel žádné primitivní odmocniny z jedné nejsou (kromě prvku 1). FFT se přesto používá, přičemž jsou v zásadě dva způsoby, jak to provést. V obou případech je dobré převést úlohu na celočíselnou: vstupní hodnoty násobíme dostatečně velkým číslem, aby jmenovatele zmizeli, výsledek pak tímto číslem vydělíme.

Jedna varianta je počítat v komplexních číslech s přibližnou hodnotou

$$\omega = e^{\frac{2\pi i}{n}} = \cos(2\pi/n) + i \sin(2\pi/n).$$

Jde tedy o numerickou metodu, použije se aritmetika v plovoucí čárce. Existují rigorózní odhady na to, s jakou přesností je třeba vzít ω , abychom dostali požadovanou přesnost výsledku. Pokud víme, že je výsledek celočíselný, lze dosáhnout absolutní přesnosti: stačí vzít ω tak, aby výsledná chyba byla menší než $\frac{1}{2}$ a výsledné hodnoty jednoduše zaokrouhlit. Oproti níže uvedenému modulárnímu metodě se však tato strategie ukazuje časově náročnější.

Máme-li na vstupu celočíselný polynom, lze využít následující trik: místo v \mathbb{Z} budeme počítat v tělese \mathbb{Z}_p , kde p zvolíme tak velké, aby modulární aritmetika neovlivnila výsledek. Konkrétní realizace závisí na zadaném problému, metodu ilustrujeme v příloze sekci na příkladech násobení celočíselných polynomů. Na tomto místě poznamenejme, že nás budou zajímat především taková prvočísla p , pro která je $p-1$ dělitelné dostatečně velkou mocninou dvojky $\{$ tak velkou, aby \mathbb{Z}_p obsahovalo primitivní odmocninu vhodnou pro FFT. Takovým prvočíslem se někdy říká FFT-prvočísla, relativně velká mocnina dvojky dělí $p-1$ např. pro prvočísla 17, 41, 97. Kde je vzít? Existence FFT-prvočísel plyne z Dirichletovy věty: ta říká, že pro každá a, m nesoudílná existuje nekonečně mnoho prvočísel $p \equiv a \pmod{m}$, tedy speciálně nekonečně mnoho prvočísel $p \equiv 1 \pmod{2^k}$ pro libovolné k . Dirichletova věta však nehovoří nic o tom, jak jsou tato prvočísla rozložená, tj. jak dlouho budeme nějaké hledat. Na tuto otázku zatím není známa přesná odpověď (Riemannova hypotéza,

problém, za který je vypsaná odměna milion dolarů, implikuje, že dlouho hledat nebudeme).

11. Rychlé násobení a dělení polynomů

11.1. Rychlé násobení.

Snad nejtýpistiější ukázkou použití modulární metody je rychlý algoritmus na násobení polynomů. Tento algoritmus je používán prakticky ve všech aplikacích, v praxi je rychlejší než školské násobení už pro poměrně malé polynomy (trochu záleží, v kterém oboru pracujeme).

Připomeďme, že modulární reprezentace je okruhový homomorfismus, tedy zachovává základní operace sčítání a násobení. Konkrétně pro polynomy, při reprezentaci

$$\varphi: \mathbf{T}[x]/m \rightarrow \mathbf{T}^n$$

$$f \mapsto (f(\alpha_1), \dots, f(\alpha_n)),$$

kde $m = (x - \alpha_1) \dots (x - \alpha_n)$, platí

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

při číselném násobení v \mathbf{T}^n se rozumí po složkách. Při volbě

$$n > \deg(f \cdot g) = \deg f + \deg g$$

dostáváme $\varphi(f \cdot g) = \varphi(f) \cdot \varphi(g)$, tedy součin v modulární reprezentaci odpovídá přesně součinu ve standardní reprezentaci. Výhoda je, že k výpočtu součinu v modulární reprezentaci potřebujeme pouze n operací v tělese \mathbf{T} .

Na tomto principu je založen algoritmus na rychlé násobení v $\mathbf{T}[x]$. Pro dané polynomy nejprve zvolíme vhodnou modulární reprezentaci { kromě podmínky $n > \deg f + \deg g$ je třeba zvolit takovou reprezentaci, pro kterou máme rychlý převodní algoritmus, jako např. FFT. Poté si spočteme modulární reprezentace \bar{a}, \bar{b} polynomů f, g , spočteme jejich součin $\bar{c} = \bar{a} \cdot \bar{b}$ (po složkách) a pomocí interpolace zjistíme z vektoru \bar{c} součin $f \cdot g$.

Algoritmus 2 (modulární násobení polynomů).

vstup: $f, g \in \mathbf{T}[x]$

výstup: $f \cdot g$

0. zvol $n > \deg f + \deg g$ a vhodné body $\alpha_1, \dots, \alpha_n \in \mathbf{T}$

1. $\bar{a} = (f(\alpha_1), \dots, f(\alpha_n))$

$\bar{b} = (g(\alpha_1), \dots, g(\alpha_n))$

2. $\bar{c} = \bar{a} \cdot \bar{b}$

3. **return** polynom h stupně $< n$ takový, že $(h(\alpha_1), \dots, h(\alpha_n)) = \bar{c}$

Příklad. Spočítáme součin polynomů $f = \frac{1}{2}x^2 + \frac{1}{2}x + 1$ a $g = x^3 - \frac{1}{3}x$:

0. zvolíme např. body $\alpha_1 = 0, \alpha_2 = 1, \alpha_3 = -1, \alpha_4 = 2, \alpha_5 = -2, \alpha_6 = 3,$

1. $\bar{a} = (1, 2, 1, 4, 2, 7), \bar{b} = \frac{1}{3} (0, 2, -2, 22, -22, 78),$

2. $\bar{c} = \bar{a} \cdot \bar{b} = \frac{1}{3} (0, 4, -2, 88, -44, 546),$

3. interpolací zjistíme, že $h = f \cdot g = \frac{1}{2}x^5 + \frac{1}{2}x^4 + \frac{5}{6}x^3 - \frac{1}{6}x^2 - \frac{1}{3}x.$

Ponecháme-li stranou volbu modulární reprezentace v kroku 0., algoritmus obnáší dva převody do modulární reprezentace polynomů stupně $< n$, jeden z této modulární reprezentace a dále n násobení v tělese \mathbf{T} . Při nahodilé volbě $\alpha_1, \dots, \alpha_n$, použití prostého dosazení hodnot a standardních algoritmů interpolace dostáváme časovou složitost $3 \cdot O(n^2) + O(n) = O(n^2)$, což je na nic. Při chytré volbě DFT reprezentace a použití algoritmu FFT dostáváme časovou složitost

$$3 \cdot O(n \log n) + O(n) = O(n \log n).$$

(Formálně vzato, jde o složitost vyjádřenou vzhledem k n , které volíme v kroku 0. Toto n však typicky závisí lineárně na součtu stupňů daných polynomů { např. v případě FFT volíme $n = 2^k > \deg f + \deg g$, tedy $n \geq 2(\deg f + \deg g)$, takže složitost má stejné asymptotické vyjádření i vzhledem ke stupni zadaných polynomů.)

Ve zbytku sekce budeme diskutovat konkrétní realizaci kroku 0. Pokud v \mathbb{T} existuje primitivní n -tá odmocnina z jedné, můžeme přímočaše aplikovat FFT. Je-li $\mathbb{T} = \mathbb{Z}_p$ a příslušná odmocnina neexistuje, problém můžeme převést na výpočet v \mathbb{Z} a výsledek zredukovat modulo p . Násobení racionálních polynomů lze také snadno převést na násobení nad \mathbb{Z} : zadané polynomy vynásobíme tak velkými čísly u, v , aby jmenovatele v koeficientech zmizeli, provedeme součin celočíselných polynomů a výsledek vydělíme součinem uv . Budeme tedy řešit problém násobení v $\mathbb{Z}[x]$.

Jak jsme uvedli v předchozí sekci, jedna možnost je numerický výpočet s komplexní odmocninou $\omega = e^{2\pi i/n}$. Protože je výsledek celočíselný, stačí zvolit takovou přesnost, aby byla výsledná chyba $< \frac{1}{2}$ (konkrétní odhady lze najít v literatuře).

Příklad. Spočítáme součin polynomů $f = \frac{1}{2}x^2 + \frac{1}{2}x + 1$ a $g = x^3 - \frac{1}{3}x$.

Ve skutečnosti budeme počítat s polynomy $2f = x^2 + x + 2$ a $3g = 3x^3 - x$.

0. Zvolíme $n = 8 = 2^3$ a reprezentaci DFT_ω pro $\omega = e^{2\pi i/8} = 0,71 + 0,71i$.
Budeme počítat s přesností na dvě desetinná místa.

1. $\bar{a} = DFT_\omega(2, 1, 1, 0, 0, 0, 0, 0) =$

$$(4, 2,71 + 1,72i, 0,98 + 1,01i, 1,28 - 0,31i, 2,02, 1,28 + 0,32i, 0,95 - 1,02i, 2,73 - 1,79i),$$

$$\bar{b} = DFT_\omega(0, -1, 0, 3, 0, 0, 0, 0) =$$

$$(2, -2,86 + 1,44i, -4,08i, 2,92 + 1,48i, -2,13, 2,98 - 1,53i, 4,25i, -3,04 - 1,58i),$$

2. $\bar{c} = \bar{a} \cdot \bar{b} =$

$$(8, -10,21 - 1,01i, 4,12 - 4,02i, 4,20 + 1,01i, -4,30, 4,30 - 1,01i, 4,36 + 4,04i, -11,12 + 1,11i),$$

3. $\frac{1}{8}DFT_{\omega^{-1}}(\bar{c}) = (0,03, -2,00, -1,00, 4,98, 3,00, 2,98, 0,00, 0,00)$. Po zaokrouhlení dostáváme (správný) výsledek

$$h = 2f \cdot 3g = 3x^5 + 3x^4 + 5x^3 - x^2 - 2x.$$

Tedy $f \cdot g = \frac{1}{6}h = \frac{1}{2}x^5 + \frac{1}{2}x^4 + \frac{5}{6}x^3 - \frac{1}{6}x^2 - \frac{1}{3}x$.

V praxi se ukazuje efektivnější modulární metoda (tentokrát provedená na okruh koeficientů). Místo v $\mathbb{Z}[x]$ budeme počítat součin v $\mathbb{Z}_p[x]$, kde p zvolíme tak velké, aby

$$f \cdot g = (f \cdot g) \bmod p,$$

tj. tak velké, aby žádný koeficient součinu v absolutní hodnotě nepřekročil $p/2$. Prvky \mathbb{Z}_p zde interpretujeme jako $dp/2e, \dots, -1, 0, 1, \dots, bp/2c$. Jak velké p potřebujeme? Připomeňme vzorec

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

(předpokládejme $n \leq m$). Každý koeficient polynomu $f \cdot g$ je součtem nejvýše $n+1$ součinů $a_j \cdot b_{i-j}$. Označíme-li $r = \max_j |a_j|$ a $s = \max_j |b_j|$, pak je každý koeficient $f \cdot g$ v absolutní hodnotě shora omezen hodnotou $(n+1)rs$. Za p tedy zvolíme prvočíslo větší než

$$2(n+1)rs,$$

které navíc splňuje FFT podmínku $2^k \geq p - 1$, kde $2^k > m + n$.

Příklad. Spočítáme součin polynomů $f = \frac{1}{2}x^2 + \frac{1}{2}x + 1$ a $g = x^3 - \frac{1}{3}x$.
Ve skutečnosti budeme počítat s polynomy $2f = x^2 + x + 2$ a $3g = 3x^3 - x$.

0. Vidíme, že $r = 2$, $s = 3$, takže potřebujeme prvočíslo $p > 2 \cdot 4 \cdot 2 \cdot 3 = 48$ splňující $n = 2^3 \cdot j \cdot p - 1$. Budeme tedy počítat v Z_{97} a zvolíme reprezentaci DFT_ω např. pro $\omega = 50$.
1. $\bar{a} = DFT_\omega(2, 1, 1, 0, 0, 0, 0) = (4, 30, 76, 88, 2, 27, 23, 57)$,
 $\bar{b} = DFT_\omega(0, -1, 0, 3, 0, 0, 0) = (2, 45, 88, 86, 95, 52, 9, 11)$,
2. $\bar{c} = \bar{a} \cdot \bar{b} = (8, 89, 92, 2, 93, 46, 13, 45)$,
3. $\frac{1}{8} DFT_\omega^{-1}(\bar{c}) = (0, -2, -1, 5, 3, 3, 0, 0)$, tedy $h = 3x^5 + 3x^4 + 5x^3 - x^2 - 2x$.

Tedy $f \cdot g = \frac{1}{6} h = \frac{1}{2}x^5 + \frac{1}{2}x^4 + \frac{5}{6}x^3 - \frac{1}{6}x^2 - \frac{1}{3}x$.

V praxi se používají různé varianty výše uvedeného principu. Např. v knihovně NTL jsou implementovány následující dvě vylepšení:

(Ěínská věta o zbytcích) Místo v jednom $Z_p[x]$ se počítá v několika $Z_{p_1}[x], \dots, Z_{p_N}[x]$, přičemž výsledek se zrekonstruuje pomocí ěínské věty o zbytcích provedené na každém koeficient výsledného polynomu zvlášť. Všechna p_i samozřejmě musí být FFT prvočísla, ve smyslu $2^k \cdot j \cdot p_i - 1$. Pokud $p_1 \dots p_N > 2(n+1)rs$, zaručeně dostaneme správný výsledek. Výhoda této metody je, že všechna p_i se zpravidla vejdou do jednoho strojového slova, takže operace v Z_{p_i} jsou velmi rychlé.

(Schönhage-Strassenův trik) Místo prvočísla p se volí hodnota $M = 2^{2^k - 1}u + 1$, kde u je tak velké, aby $M > 2(n+1)rs$. Okruh Z_M sice není těleso, ale to u FFT nevadí (viz cvičení k minulé sekci). Důležité je, že prvek 2^u je 2^k -tá primitivní odmocnina z jedné (ověřte!). Výhoda těchto okruhů je v tom, že počítání modulo ěíslo tvaru $M = 2^e + 1$ je velmi rychlé (lineární, oproti klasickému kvadratickému): protože $2^e \equiv -1 \pmod{M}$, k redukci ěísla a modulo M stačí $\log_M ac$ operací sčítání a odčítání modulo M ; násobení a dělení mocninou dvojky se realizuje jako bitový posun, ěili nestojí (skoro) nic.

11.2. Rychlé dělení polynomů.

K dělení polynomů bohužel nelze přímočaše využít modulární reprezentaci, protože okruhový homomorfismus dělení nezachovává. Ukážeme si so stikovanější algoritmus: dělení polynomů převedeme na násobení a počítání inverzního prvku v oboru formálních mocninných řad. Rychlé násobení už umíme. A na rychlém výpočtu inverzní řady si předvedeme jednu z variant Newtonovy metody.

Zopakujme stručně pojem formální mocninné řady nad komutativním okruhem R . Jde o formální výrazy tvaru $\sum_{i=0}^{\infty} a_i x^i$, kde a_i jsou koeficienty z R . Na těchto výrazech jsou definovány okruhové operace podobně jako pro polynomy:

$$\left(\sum_{i=0}^{\infty} a_i x^i \right) \left(\sum_{i=0}^{\infty} b_i x^i \right) = \sum_{i=0}^{\infty} (a_i + b_i) x^i,$$

$$\left(\sum_{i=0}^{\infty} a_i x^i \right) \left(\sum_{i=0}^{\infty} b_i x^i \right) = \sum_{i=0}^{\infty} \left(\sum_{j=0}^i a_j b_{i-j} \right) x^i.$$

Okruh formálních mocninných řad nad R značíme $R[[x]]$. Je-li R obor integrity, pak je $R[[x]]$ také obor integrity. Polynomy nad R tvoří jeho podobor. V této sekci budeme potřebovat následující důležitou vlastnost.

Tvrzení 11.1. Buď R obor integrity a $f = \sum a_i x^i \in R[[x]]$. Pak f je invertibilní v $R[[x]]$ právě tehdy, když a_0 je invertibilní v R .

Důkaz tohoto tvrzení není těžký a plyne z úvah v části 11.3, kde se budeme věnovat efektivnímu výpočtu inverzní mocninné řady. Zatím se podívejme, jak se dá na tento problém převést dělení polynomů.

Pro účely této sekce zavedeme technické označení: pro polynom f definujeme

$$f = x^{\deg f} f(x^{-1}).$$

Jinými slovy, f je polynom, který vznikne z f , když napíšeme jeho koeficienty v opačném pořadí. Např. pro $f = 3x^3 + 2x^2 - 1$ je $f = x^3 (3x^{-3} + 2x^{-2} - 1) = 3 + 2x - x^3$.

Buď T těleso a uvažujme polynomy $f, g \in T[x]$, $g \neq 0$. Označme $n = \deg f$, $m = \deg g$ a předpokládejme $n \geq m$. Chceme spočítat podíl a zbytek, tj. hledáme polynomy $q, r \in T[x]$ splňující

$$f = gq + r, \quad \deg r < m.$$

Musí tedy platit také

$$f(x^{-1}) = g(x^{-1})q(x^{-1}) + r(x^{-1})$$

a po vynásobení x^n dostáváme podmínku

$$f = gq + x^{n-\deg r} r.$$

Dále pracujeme v oboru mocninných řad $T[[x]]$. Protože g má zaručeně nenulový absolutní člen, existuje inverzní řada $(g)^{-1}$ a po násobení touto řadou dostáváme vyjádření

$$q = f (g)^{-1} - x^{n-\deg r} r (g)^{-1}.$$

Levá strana rovnosti je polynom stupně nejvýše $n - m$, takže mocninná řada na pravé straně rovnosti má všechny členy počínaje x^{n-m+1} nulové. Jenže řada $x^{n-\deg r} r (g)^{-1}$ má prvních $n - \deg r > n - m$ členů nulových, takže q je ve skutečnosti rovno prvním $n - m + 1$ členům mocninné řady $f (g)^{-1}$!

Algoritmus 3 (rychlé dělení se zbytkem).

vstup: $f, g \in T[x]$, $g \neq 0$

výstup: $f \operatorname{div} g, f \operatorname{mod} g$

0. $n := \deg f$, $m := \deg g$, **if** $n < m$ **then return** $0, f$
1. $h :=$ prvních $n - m + 1$ členů mocninné řady $(g)^{-1}$
2. $w := f - h \operatorname{mod} x^{n-m+1}$
3. buď q polynom stupně $n - m$, pro který $q = w$
4. **return** $q, f - gq$

V kroku 3. nemůžeme jednoduše položit $q = w$, nebo nemusí sedět stupně: např. pro $f = x^3$, $g = x$ je $w = 1$, ale $q = x^2 \neq w$.

Správnost algoritmu jsme odvodili v předchozím textu. Co se týče složitosti, dělení se zbytkem jsme převedli na dvě násobení a jeden výpočet inverzní řady. Konkrétně, pro vstup stupně n, m provádíme dvě násobení polynomů stupně $< n$ a hledáme $n - m + 1$ členů mocninné řady $(g)^{-1}$. Násobení umíme řešit s časovou složitostí $O(n \log n)$. Pokud bychom uměli se stejnou složitostí hledat i prvních n členů inverzní mocninné řady, dostali bychom časovou složitost dělení

$$O(n \log n).$$

(V praxi je dělení celočíselných polynomů zhruba pětikrát pomalejší než násobení, viz cvičení.)

Příklad. Spočítáme podíl a zbytek pro polynomy

$$f = x^4 + x^3 + x^2 + x + 1, \quad g = x^2 + x - 1.$$

1. Pro $g = x^2 + x + 1$ je $(g)^{-1} = 1 - x + 2x^2 + \dots$, tedy $h = 2x^2 - x + 1$.
2. $w = 1 + 2x^2 + 2x^3 + \dots \pmod{x^3} = 2x^2 + 1$.
3. $q = w^{-1} = x^2 + 2$.
4. Odpovídí je $f \operatorname{div} g = x^2 + 2$ a $f \operatorname{mod} g = -x + 3$.

11.3. Výpočet inverzní mocninné řady.

K dokončení rychlého algoritmu na dělení polynomů zbývá vyřešit problém výpočtu koeficientů členů inverzní mocninné řady daného polynomu. Uvažujme tedy mocninnou řadu $f = \sum a_i x^i \in T[[x]]$, budeme hledat řadu $g = \sum b_i x^i \in T[[x]]$ takovou, že $f \cdot g = 1$.

Připomeňme tvrzení 11.1. Pokud $a_0 = 0$, pak $x \nmid f$, a tedy nemůže existovat žádná taková řada g . Předpokládejme tedy $a_0 \neq 0$. Z vzorce pro násobení mocninných řad plynou následující podmínky:

$$\begin{aligned} a_0 b_0 &= 1 \\ a_0 b_1 + a_1 b_0 &= 0 \\ a_0 b_2 + a_1 b_1 + a_2 b_0 &= 0 \\ a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0 &= 0 \\ &\dots \end{aligned}$$

Můžeme tedy vyjádřit

$$\begin{aligned} b_0 &= a_0^{-1} \\ b_1 &= -a_0^{-1} a_1 b_0 \\ b_2 &= -a_0^{-1} (a_1 b_1 + a_2 b_0) \\ b_3 &= -a_0^{-1} (a_1 b_2 + a_2 b_1 + a_3 b_0) \\ &\dots \end{aligned}$$

a vidíme, že řada $g = f^{-1}$ existuje (a je jednoznačně určena) a zároveň jsme dostali algoritmus, který ji spočte.

K získání prvních n členů b_0, \dots, b_{n-1} potřebujeme 1 výpočet inverzu, $2 + 3 + 4 + \dots + n$ operací násobení a $1 + 2 + 3 + \dots + (n-1)$ operací sčítání v tělese \mathbb{T} , tedy celkově

$$O(n^2)$$

operací. To je moc.

K získání n členů je potřeba n kroků, každý z nich s lineární složitostí; v jednom kroku přibude jeden nový člen. My bychom potřebovali složitost $O(n \log n)$, chceme tedy postup, který by umožnil získat n členů v pouhých $\log n$ lineárních krocích. Jinými slovy, v jednom kroku by bylo dobré počít členů zdvojnásobit. Ukážeme si postup inspirovaný tzv. Newtonovou metodou na hledání kořenů obecných rovnic (viz nějaký kurz numerické matematiky).

Výpočet prvních n členů řady f^{-1} se dá interpretovat následujícím způsobem: k dané řadě f hledáme polynom g stupně $< n$ takový, že

$$f \cdot g = 1 + 0x + 0x^2 + \dots + 0x^{n-1} + x^n = h,$$

kde h je libovolná řada, nebo jinými slovy

$$f \cdot g \equiv 1 \pmod{x^n}.$$

(Ve výsledku tedy záleží pouze na prvních n členech řady f .)

Přepíšme tento vztah jako $x^n \nmid f \cdot g - 1$. Pak $x^{2n} \nmid (f \cdot g - 1)^2 = f^2 g^2 - 2fg + 1$ a dostáváme

$$f \cdot g \equiv (2 - fg) \pmod{x^{2n}}.$$

Vidíme, že $g = (2 - fg) \bmod x^{2n}$ je prvních $2n$ členů inverzní mocninné řady f^{-1} a dostáváme tak slibovaný způsob, jak v jednom kroku počet členů zdvojnásobit.

Algoritmus 4.

- vstup:** $n, f = \sum a_i x^i$ splývající $a_0 \neq 0$
výstup: prvních n členů mocninné řady f^{-1}
0. $g_0 := a_0^{-1}$
 1. **for** $i = 1, \dots, \lceil \log_2 n \rceil$ **do**
 $g_i := g_{i-1} (2 - fg_{i-1}) \bmod x^{2^i}$
 2. **return** $g_{\lceil \log_2 n \rceil} \bmod x^n$

Správnost algoritmu plyne z výše uvedených úvah: každé g_i obsahuje právě 2^i členů mocninné řady f^{-1} . Co se týče časové složitosti, v i -tém kroku cyklu vstupuje do hry polynom $f \bmod x^{2^i}$ stupně $< 2^i$ a polynom g_{i-1} stupně $< 2^{i-1}$. Složitost jednoho kroku (jeden rozdíl a dvě násobení) tak bude při použití rychlého násobení $O(2^i \log 2^i) = O(i2^i)$ a celkovou časovou složitost můžeme vyjádřit

$$O\left(\sum_{i=1}^{\lceil \log_2 n \rceil} i2^i\right) = O\left(\log n \sum_{i=1}^{\lceil \log_2 n \rceil} 2^i\right) = O\left(\log n (2^{\lceil \log_2 n \rceil + 1} - 1)\right) = O(n \log n)$$

(využíváme odhadu $i \leq \lceil \log_2 n \rceil$ a souětu geometrické řady).

Příklad. Spočítáme první 4 členy mocninné řady f^{-1} , kde

$$f = 1 - 2x + 3x^2 + x^4 - x^5.$$

Klasická metoda vede na vyjádření

$$\begin{aligned} b_0 &= a_0^{-1} = 1 \\ b_1 &= a_0^{-1}(a_1 b_0) = (-2 - 1) = -3 \\ b_2 &= a_0^{-1}(a_1 b_1 + a_2 b_0) = ((-2)(-3) + 3 - 1) = 2 \\ b_3 &= a_0^{-1}(a_1 b_2 + a_2 b_1 + a_3 b_0) = ((-2)(2) + 3(-3) + 0 - 1) = -14 \end{aligned}$$

zatímco Newtonův algoritmus dává

$$\begin{aligned} g_0 &= 1 \\ g_1 &= 1 - (2 - f) \bmod x^2 = 1 + 2x \\ g_2 &= (1 + 2x)(2 - f \bmod (1 + 2x)) \bmod x^4 = 1 + 2x + x^2 - 4x^3 \end{aligned}$$

(Vidíme, že celý výpočet závisí výhradně na $f \bmod x^4$.)

12. Rozklady polynomů nad konečnými tělesy

Vitina algoritmů na faktorizaci předpokládá, že je vstupní polynom f tzv. bezčtvercový, tedy že v jeho ireducibilním rozkladu jsou všechny exponenty $k_1, \dots, k_m = 1$. Nejdříve probereme algoritmus, který daný polynom rozloží na součin bezčtvercových (ne nutně ireducibilních), a to s časovou složitostí $O(n^3)$. V druhé části pak ukážeme Berlekampův algoritmus na faktorizaci v $\mathbb{F}_q[x]$.

12.1. Bezčtvercová faktorizace.

Definice. Polynom f se nazývá bezčtvercový, pokud neexistuje nekonstantní polynom g takový, že $g^2 \mid f$. Bezčtvercovým rozkladem polynomu f rozumíme po dvou nesoudělné bezčtvercové polynomy h_1, \dots, h_k splývající

$$f = h_1 h_2^2 h_3^3 \dots h_k^k$$

(tedy h_i obsahuje právě ty ireducibilní činitele, které se v f vyskytují v i -té mocnini).

Příklad. Bezètercovým rozkladem polynomu

$$f = x^8 + x^7 - x^6 - x^5 - x^4 - x^3 + x^2 + x = (x^3 + x)(x - 1)^2(x + 1)^3$$

v $\mathbb{Z}[x]$ jsou polynomy $h_1 = x^3 + x$, $h_2 = x - 1$, $h_3 = x + 1$.

Pøipomeòme, že charakteristikou daného oboru rozumíme nejmenší pøirozené k splòující $k \cdot 1 = 0$, pokud takové k existuje, resp. 0 v opaèném pøípadì. Koneènè těleso \mathbb{F}_q , kde $q = p^n$, má charakteristiku p .

Základní verze algoritmu na bezètercovou faktorizaci funguje pro polynomy nad libovolným gaussovským oborem charakteristiky 0 (rekurzivně tedy i pro polynomy více promìnných). V pøípadì nenulové charakteristiky nastává menší zàdrhel, jehož øešení pøedvedeme pro polynomy jedné promìnné nad koneèným tělesem.

Pro obory charakteristiky 0 platí známé tvrzení z kurzu analýzy: pokud $f' = 0$, pak je polynom f konstantní. Obecnì to pravda není: pøi derivaci vznikají celoèíselné konstanty, které se mohou v nenulové charakteristice interpretovat jako nuly. Pro koneèná tělesa naštítí mùžeme takovou situaci snadno popsat.

Lemma 12.1. Bui f polynom z $\mathbb{F}_q[x]$, $q = p^n$, takový, že $f' = 0$. Pak $f = g^p$ pro nìjaký polynom $g \in \mathbb{F}_q[x]$.

Polynom g budeme znaèit $\sqrt[p]{f}$.

Dùkaz. Pokud $f' = 0$, pak všechny nenulové èleny v f musí mít exponent dìlitelný p , tj. mùžeme psát $f = \sum a_i x^{ip}$. Dejme $g = \sum b_i x^i$ pro b_i taková, že $b_i^p = a_i$ (mùžeme volit $b_i = a_i^{p^{-1}}$, protože $b_i^p = a_i^{p^{-1} \cdot p} = a_i$ podle tvrzení 9.3). Užitím tvrzení 9.1 dostaneme $g^p = \sum b_i^p x^{ip} = f$.

Princip algoritmu na bezètercovou faktorizaci popisuje následující vèta.

Vèta 12.2. Bui \mathbf{R} gaussovský obor charakteristiky 0 nebo $\mathbf{R} = \mathbb{F}_q$, a bui f primitivní polynom z $\mathbf{R}[x]$.

- (1) f je bezètercový právi tehdy, když $\text{NSD}(f, f') = 1$.
- (2) Bui $f = \prod_{i=1}^k h_i^{i-1}$ bezètercový rozklad. Pak
 - (a) pokud $\text{char}(\mathbf{R}) = 0$, pak $\text{NSD}(f, f') = \prod_{i=1}^k h_i^{i-1}$;
 - (b) pokud $\mathbf{R} = \mathbb{F}_q$, $q = p^n$, pak $\text{NSD}(f, f') = \prod_{p \nmid i} h_i^{i-1} \prod_{p \mid i} h_i^{i-1}$.

Dùkaz. (1) Pøedpokládejme, že f není bezètercový. Pak $f = g^2 \cdot h$ pro nìjaké netriviální g, h a po zderivování dostáváme $f' = 2gg'h + g^2h'$. Tedy g je společný dìlitel f i f' , spor.

(2) Bui $f = \prod_{i=1}^m g_i$ ireducibilní rozklad (tj. g_1, \dots, g_m jsou po dvou neasociované ireducibilní polynomy). Pak

$$f' = g_1' g_2 \dots g_m + g_1 g_2' g_3 \dots g_m + \dots + g_1 \dots g_{m-1} g_m'$$

Pøedpokládejme, že f a f' mají nìjakého netriviálního společného dìlitele. Pak existuje nìjaký ireducibilní společný dìlitel, a ten je roven nìkterému g_i . Protože g_i dílí f' a pøitom se vyskytuje ve všech èlenech výše uvedeného souètu kromì i -tého, musí g_i dilit i i -tý èlen, tj. $g_i \mid g_1 g_2 \dots g_{i-1} g_{i+1} \dots g_m$. Protože $g_i \nmid g_j$ pro žádné $j \neq i$, musí g_i dilit g_i' . Jenže $\deg g_i' < \deg g_i$, a tedy $g_i' = 0$. V pøípadì charakteristiky 0 to znamená, že je g_i konstantní, spor. V pøípadì koneèných těles mùžeme ještì nastat možnost, že $g_i = g^p$ pro nìjaký (nekonstantní) polynom g , to je ale ve sporu s ireducibilitou g_i .

(2a) Derivací bezètercového rozkladu f dostáváme

$$f' = \sum_{j=1}^k (h_j^j)' \prod_{i \neq j} h_i^i = \sum_{j=1}^k j h_j^{j-1} h_j' \prod_{i \neq j} h_i^i.$$

Vidíme, že $\prod_{i=1}^k h_i^{i-1}$ je společným dělitelem polynomů f a f^0 . Dokážeme, že to je největší společný dělitel. Předpokládejme, že ne, tedy že existuje nekonztantní polynom g , který dělí oba polynomy

$$\frac{f}{\prod_{i=1}^k h_i^{i-1}} = \prod_{i=1}^k h_i, \quad \frac{f^0}{\prod_{i=1}^k h_i^{i-1}} = \sum_{j=1}^k j h_j^0 \prod_{i \neq j} h_i.$$

Opět můžeme předpokládat, že je polynom g ireducibilní a tedy že $g \mid h_m$ pro nějaké m . Podobně jako v bodě (1), protože se h_m vyskytuje ve všech členech součtu $\sum_{i=1}^k j h_j^0 \prod_{i \neq j} h_i$ kromě m -tého, musí g dělit mh_m^0 . Jenže polynom h_m je bezčtercový, tedy podle (1) je nesoudlný s h_m^0 , spor.

(2b) Derivace má stejné vyjádření jako v předchozím případě; protože v charakteristice p vypadnou všechny členy, kde $p \mid j$, dostáváme

$$f^0 = \sum_{j=1}^k j h_j^{j-1} h_j^0 \prod_{i \neq j} h_i = \prod_{p \nmid i} h_i^i \left(\sum_{p \nmid j} j h_j^{j-1} h_j^0 \prod_{p \nmid i \neq j} h_i^i \right)$$

a vidíme, že

$$\prod_{p \nmid i} h_i^i \prod_{p \mid i} h_i^{i-1}$$

je společným dělitelem polynomů f, f^0 . Podobně jako v předchozím případě se dokáže, že jde o největšího společného dělitele.

12.1.1. Bezčtercová faktorizace v charakteristice 0.

Mějme dán polynom f , položme

$$f_1 = \text{NSD}(f, f^0), \quad g_1 = f/f_1$$

a dále definujeme induktivně

$$g_{j+1} = \text{NSD}(f_j, g_j), \quad f_{j+1} = f_j/g_{j+1}.$$

V charakteristice 0 vycházejí pro $f = \prod_{i=1}^k h_i^i$ následující hodnoty:

j	f_j	g_j
1	$\prod_{i=2}^k h_i^{i-1}$	$\prod_{i=1}^k h_i$
2	$\prod_{i=3}^k h_i^{i-2}$	$\prod_{i=2}^k h_i$
3	$\prod_{i=4}^k h_i^{i-3}$	$\prod_{i=3}^k h_i$
...
$k-1$	h_k	$h_{k-1} h_k$
k	1	h_k
$k+1$	1	1

Vidíme, že délku bezčtercového rozkladu, tj. hodnotu k , poznáme podle toho, kdy g_{k+1} vyjde konstantní. Z tabulky však lze vyčíst mnohem víc: především to, že podíl g_j/g_{j+1} je roven hledanému polynomu h_j .

Vzhledem k tomu, že NSD jsou definovány až na asociovanost, všechny políčka v tabulce jsou určena až na asociovanost; speciálně, tímto postupem nemusíme nutně získat bezčtercový rozklad polynomu f , ale nějakého polynomu asociovaného s f . Tuto technickou obtíž je samozřejmě jednoduché vyřešit.

Na základě uvedených myšlenek můžeme zformulovat algoritmus.

Algoritmus 5 (bezřetvercová faktorizace v charakteristice 0).

vstup: $f \in R[x]$ nekonzstantní primitivní

výstup: bezřetvercový rozklad h_1, \dots, h_k nějakého polynomu asociovaného s f

1. $f_1 := \text{NSD}(f, f^0)$, $g_1 := f/f_1$, $j := 1$
2. **while** $\deg g_j > 0$ **do**
 $g_{j+1} := \text{NSD}(f_j, g_j)$, $f_{j+1} := f_j/g_{j+1}$, $h_j := g_j/g_{j+1}$
 $j := j + 1$
3. **return** h_1, \dots, h_{j-1}

Tvrzení 12.3. Algoritmus 5 funguje.

Důkaz. Mějme na vstupu polynom $f = \prod_{i=1}^k h_i^i$. Stačí formálně ověřit indukci, že $f_j \mid \prod_{i=j+1}^k h_i^i$ a $g_j \mid \prod_{i=j}^k h_i$. Z toho plyne, že skutečně $h_j \mid g_j/g_{j+1}$, a také že se algoritmus zastaví pro $j = k + 1$.

Pro $j = 1$ obě tvrzení plynou z věty 12.2. V indukčním kroku pak

$$g_{j+1} \mid \text{NSD}(f_j, g_j) = \text{NSD}\left(\prod_{i=j+1}^k h_i^i, \prod_{i=j}^k h_i\right) = \prod_{i=j+1}^k h_i,$$

a tudíž $f_{j+1} = f_j/g_{j+1} \mid \prod_{i=j+1}^k h_i^i / \prod_{i=j+1}^k h_i = \prod_{i=j+2}^k h_i^{i-j-1}$.

Tvrzení 12.4. Ěasová složitost algoritmu 5 je $O(nN(n))$, kde $n = \deg f$ a funkce N značí složitost výpočtu NSD v oboru $R[x]$.

Důkaz. Algoritmus počítá $k + 1$ hodnot f_j, g_j a k hodnot h_j . Výpočet každé z nich obnáší NSD nebo dělení, přičemž dělení je méně náročné. Celkovou složitost tedy můžeme odhadnout jako $(3k + 2)O(N(n)) = O(nN(n))$. (Odhad $k = n$ je nejlepší možný: rovnost nastane v případě, že $h_1 = \dots = h_{n-1} = 1$ a $\deg h_n = 1$.)

Příklad. Uvažujme polynom

$$f = x^7 + x^6 - x^5 - x^4 - x^3 - x^2 + x + 1.$$

Algoritmus 5 v oboru $Z[x]$ probíhá následovně:

j	f_j	g_j	h_{j-1}
1	$x^3 + x^2$	$x - 1$	$x^4 - 1$
2	$x + 1$	$x^2 - 1$	$x^2 + 1$
3	1	$x + 1$	$x - 1$
4	1	1	$x + 1$

Odpovědí je $h_1 = x^2 + 1$, $h_2 = x - 1$, $h_3 = x + 1$, èli

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

Příklad. Uvažujme polynom

$$f = (x + 1)^6.$$

Algoritmus 5 v oboru $Z[x]$ probíhá následovně:

j	f_j	g_j	h_{j-1}
1	$(x + 1)^5$	$x + 1$	
2	$(x + 1)^4$	$x + 1$	1
3	$(x + 1)^3$	$x + 1$	1
4	$(x + 1)^2$	$x + 1$	1
5	$x + 1$	$x + 1$	1
6	1	$x + 1$	1
7	1	1	$x + 1$

Odpovědí je $h_1 = \dots = h_5 = 1$ a $h_6 = x + 1$.

12.1.2. Bezètercová faktorizace nad koneènými tělesy.

Uvaŕujme nyní stejný postup nad tělesem \mathbb{F}_q . Podle vĕty 12.2 vycházejí pro polynom $f = \prod_{i=1}^k h_i^i$ následující hodnoty:

j	f_j	g_j
1	$\prod_{p^j} h_i^i$	$\prod_{p-i} h_i^i$
2	$\prod_{p^j} h_i^i$	$\prod_{p-i} h_i^i$
3	$\prod_{p^j} h_i^i$	$\prod_{p-i} h_i^i$
	\dots	\dots
k	$\prod_{p^j} h_i^i$	$\prod_{p-i} h_i^i$
$k+1$	$\prod_{p^j} h_i^i$	1

Délku bezètercového rozkladu, tj. hodnotu k , opět poznáme podle toho, ŕe g_{k+1} vyjde konstantní. Rozdíly jsou dva: v f_k nám zbyde souèin tĕch bezètercových faktorù, které jsou v mocninì dĕlitelné p . Podíl g_j/g_{j+1} vyjde roven h_j pouze v pøípadì, ŕe $p-j$, v opaèném pøípadì vyjde 1. Postupem tedy zjistíme vĕchny bezètercové faktory kromì p -tého, $2p$ -tého atd. Na konci tedy staèí vzít p -tou odmocninu polynomu f_k a postup opakovat.

Algoritmus 6 (bezètercová faktorizace nad koneènými tělesy).

vstup: $f \in \mathbb{F}_q[x]$ nekonstantní

výstup: bezètercový rozklad h_1, \dots, h_k nĕjakého polynomu asociovaného s f

0. **if** $f^0 = 0$ **then goto** 3.

1. $f_1 := \text{NSD}(f, f^0)$, $g_1 := f/f_1$, $j := 1$

2. **while** $\deg g_j > 0$ **do**

$g_{j+1} := \text{NSD}(f_j, g_j)$, $f_{j+1} := f_j/g_{j+1}$, $h_j := g_j/g_{j+1}$

$j := j + 1$

$f := f_j$

3. **if** $\deg f = 0$ **then return** h_1, \dots, h_{j-1}

else spoèti bezètercovou faktorizaci $h_p, h_{2p}, \dots, h_{lp}$ polynomu $\prod_p f$,

return $h_1, h_2, \dots, h_{\max(j-1, lp)}$

Dùkaz správnosti je analogický jako pro algoritmus 5.

Pøíklad. Uvaŕujme polynom

$$f = x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x + 1.$$

Algoritmus 6 v oboru $\mathbb{Z}_3[x]$ probĕhne následovně:

j	f_j	g_j	h_{j-1}
1	$x^4 + x^3 + x + 1$	$x^3 + x^2 + x + 1$	1
2	$x^3 + 1$	$x + 1$	$x^2 + 1$
3	$x^3 + 1$	1	$x + 1$

Dostáváme $h_1 = x^2 + 1$, $h_2 = x + 1$ a zůstává nám polynom $f = x^3 + 1$. Tøetí odmocnina je $x + 1$, provedeme tedy nový výpoèet s tímto polynomem a výsledek uloŕíme do h_3 (event. h_6, h_9, \dots , kdyby tyto vyšly netriviální). Dostáváme

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

Pøíklad. Uvaŕujme polynom

$$f = x^6 + x^4 + x^2 + 1.$$

Algoritmus 6 v oboru $\mathbb{Z}_2[x]$ probĕhne následovně: protože $f^0 = 0$, budeme rovnou uvaŕovat druhou odmocninu, tedy polynom $x^3 + x^2 + x + 1$.

j	f_j	g_j	h_{j-1}
1	$x^2 + 1$	$x + 1$	1
2	$x + 1$	$x + 1$	1
3	1	$x + 1$	1
4	1	1	$x + 1$

Spočetali jsme, že $P_{\bar{f}} = (x + 1)^3$, tedy výsledek je $f = (x + 1)^6$.

12.2. Berlekampův algoritmus.

Nejjednodušším faktorizačním algoritmem v $F_q[x]$ je Berlekampův algoritmus. Jeho složitost vzhledem ke stupni daného polynomu je kubická, nevýhodou jeho základní verze je exponenciální složitost vzhledem k $l(q)$, kde $l(q)$ značí počet cifer čísla q (tedy $l(q) = \Theta(\log q)$). Vstupem je monický bezčtevcový polynom. Principem Berlekampova algoritmu je následující tvrzení.

Tvrzení 12.5. Buď f monický bezčtevcový polynom z $F_q[x]$ a uvažujme nekonstantní polynom $h \in F_q[x]$ splňující

$$h^q \equiv h \pmod{f}.$$

Pak

$$f = \prod_{a \in F_q} \text{NSD}(f, h - a).$$

Důkaz. Polynomy $h - a$ jsou po dvou nesouditelné (protože $\text{NSD}(h - a_1, h - a_2) = \text{NSD}(h - a_1, a_1 - a_2) = 1$), tedy po dvou nesouditelné jsou i polynomy $\text{NSD}(f, h - a)$. Protože každý z nich dělí polynom f , tak i $\prod_{a \in F_q} \text{NSD}(f, h - a)$ dělí f . Zbývá dokázat opak.

Buď $f = g_1 g_2 \dots g_m$ rozklad f na ireducibilní činitele. Z předpokladu plyne, že $f \equiv h^q \pmod{h}$, a tedy podle tvrzení 9.4 dosazením polynomu h za proměnnou x dostaneme $f \equiv \prod_{a \in F_q} (h - a)$. Protože jsou g_i ireducibilní a $h - a$ po dvou nesouditelné, pro každé i existuje právě jeden prvek $a \in F_q$ takový, že $g_i \equiv h - a$, a tedy také $g_i \equiv \text{NSD}(f, h - a)$. Díky bezčtevcovosti jsou polynomy g_i po dvou nesouditelné, takže $f = g_1 \dots g_m \equiv \prod_{a \in F_q} \text{NSD}(f, h - a)$.

Zjistili jsme, že polynomy f a $\prod_{a \in F_q} \text{NSD}(f, h - a)$ se navzájem dělí. Protože jsou oba monické, jsou stejné.

Tvrzení poskytuje netriviální rozklad polynomu f , kdykoliv do něj dosadíme (nekonstantní) polynom h stupně menšího než $\deg f$. Otázka je, kde takový h vzít. Označme

$$W = \{h \in F_q[x] : \deg h < \deg f, h^q \equiv h \pmod{f}\}.$$

Další tvrzení vypovídá o struktuře této množiny.

Tvrzení 12.6. Nechť f je bezčtevcový polynom z $F_q[x]$ s ireducibilním rozkladem $f = g_1 g_2 \dots g_m$. Pak

- (1) pro každý polynom $h \in W$ a každé i je

$$h \pmod{g_i} \in F_q,$$

- (2) množina W tvoří vektorový prostor nad F_q dimenze m a zobrazení

$$\varphi : W \rightarrow (F_q)^m, \quad h \mapsto (h \pmod{g_1}, \dots, h \pmod{g_m})$$

je izomorfní smys vektorových prostorů.

Důkaz. (1) V důkazu předchozího tvrzení jsme viděli, že pro každý nekonstantní polynom $h \in W$ a každé i existuje prvek $a \in F_q$ takový, že $g_i \equiv h - a$. ělí $h \pmod{g_i} = a \in F_q$. Pro konstantní polynomy je tvrzení triviální.

(2) Uvedené zobrazení zřejmě zachovává sčítání a skalární násobení. Pokud ověříme, že jde o bijekci, W bude vektorovým prostorem nad \mathbb{F}_q a φ izomorfismus. Zvolme libovolný vektor $(a_1, \dots, a_m) \in \mathbb{F}_q^m$ a uvažujme soustavu kongruencí $h \equiv a_i \pmod{g_i}$, $i = 1, \dots, m$. Protože jsou g_i po dvou nesoudělné (z bezčtyvercovosti), čínská věta o zbytcích zaručuje právě jedno řešení h modulo $g_1 \cdots g_m = f$ stupně $< \deg f$. Přitom podle tvrzení 9.3

$$h^q \equiv a_i^q = a_i \pmod{g_i},$$

a tedy díky nesoudělnosti platí $h^q \equiv h \pmod{g_1 \cdots g_m = f}$. Zobrazení φ je tedy bijekce, jediným vzorem vektoru (a_1, \dots, a_m) je uvedený polynom h .

Zbývá vyřešit otázku, jak nějaké netriviální prvky W nalézt. Označme $n = \deg f$. Uvažujme polynom

$$h = a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in \mathbb{F}_q[x].$$

Zajímá nás, kdy $h^q \equiv h \pmod{f}$, neboli kdy $h^q \bmod f = h$. Z tvrzení 9.1 a 9.3 vidíme, že

$$h^q = \left(\sum_{i=0}^{n-1} a_i x^i \right)^q = \sum_{i=0}^{n-1} (a_i x^i)^q = \sum_{i=0}^{n-1} a_i^q x^{iq} = \sum_{i=0}^{n-1} a_i x^{iq}.$$

Označíme-li $q_{i,j}$ koeficienty polynomu $x^{jq} \bmod f$, tj.

$$\begin{aligned} 1 = x^0 \bmod f &= q_{0,0} + q_{1,0}x + \dots + q_{n-1,0}x^{n-1} \\ x^q \bmod f &= q_{0,1} + q_{1,1}x + \dots + q_{n-1,1}x^{n-1} \\ &\dots \\ x^{(n-1)q} \bmod f &= q_{0,n-1} + q_{1,n-1}x + \dots + q_{n-1,n-1}x^{n-1}, \end{aligned}$$

pak

$$\begin{aligned} h^q \bmod f &= \left(\sum_{j=0}^{n-1} a_j x^{jq} \right) \bmod f = \sum_{j=0}^{n-1} a_j (x^{jq} \bmod f) = \\ &= \sum_{j=0}^{n-1} \left(a_j \sum_{i=0}^{n-1} q_{i,j} x^i \right) = \sum_{i=0}^{n-1} \left(\sum_{j=0}^{n-1} q_{i,j} a_j \right) x^i. \end{aligned}$$

Označíme-li koeficienty $h^q \bmod f = b_0 + b_1x + \dots + b_{n-1}x^{n-1}$, pak lze odvozený vztah zapsat maticově jako

$$\begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{pmatrix} = Q \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{pmatrix}, \text{ kde } Q = \begin{pmatrix} q_{0,0} & q_{0,1} & \dots & q_{0,n-1} \\ q_{1,0} & q_{1,1} & \dots & q_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n-1,0} & q_{n-1,1} & \dots & q_{n-1,n-1} \end{pmatrix}$$

Rovnost $h^q \bmod f = h$ tedy nastane právě tehdy, když

$$Q (a_0, \dots, a_{n-1})^T = (a_0, \dots, a_{n-1})^T,$$

tedy právě tehdy, když

$$(Q - E) (a_0, \dots, a_{n-1})^T = (0, 0, \dots, 0)^T$$

(zde E značí jednotkovou matici). Odvodili jsme následující tvrzení.

Tvrzení 12.7. Buď Q matice $n \times n$ jejíž sloupce tvoří koeficienty polynomů

$$1, x^q \bmod f, x^{2q} \bmod f, \dots, x^{(n-1)q} \bmod f.$$

Pak polynom $h = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ leží ve W právě tehdy, když (a_0, \dots, a_{n-1}) je řešením homogenní soustavy rovnic s maticí $Q - E$.

Podle tvrzení 12.6 je dimenze prostoru W rovna počtu ireducibilních faktorů polynomu f , a to je rovno dimenzi prostoru řešení uvedené soustavy, tj. hodnotě $n - \text{h}(Q - E)$, kde h značí hodnotu matice. Všimněte si, že první sloupec matice $Q - E$ je nulový, takže vektory $(a, 0, 0, \dots, 0)$ jsou vždy řešením; tyto vektory odpovídají konstantním polynomům ve W , které nás nezajímají, protože neposkytují netriviální rozklad.

Kostra Berlekampova algoritmu je následující:

- (1) Gaussovou eliminací vyřešíme soustavu rovnic s maticí $Q - E$. Je-li dimenze prostoru řešení 1, polynom f je ireducibilní. V opačném případě vezmeme libovolné řešení dané soustavy různé od $(a, 0, 0, \dots, 0)$ a označíme příslušný polynom h .
- (2) Ze vzorce z tvrzení 12.5 získáme netriviální rozklad $f = g_1 g_2 \dots g_l$.
- (3) Pokud je l rovno dimenzi prostoru řešení, jsme hotovi. V opačném případě pokračujeme rekurzivně pro každý polynom g_1, g_2, \dots, g_l .

Uvedený postup lze zdatelně optimalizovat: místo náhodné volby h si spočteme bázi prostoru řešení $h_1 = 1, h_2, \dots, h_m$ (polynom $h_1 = 1$ a jeho násobky odpovídají nezajímavým konstantním polynomům). Použijeme h_2 na nalezení netriviálního rozkladu f , jednotlivé faktory se pokusíme dále rozložit polynomem h_3 , atd. (Zde aplikujeme tvrzení 12.5 na jednotlivé faktory, předpoklad $h^q \equiv h \pmod{f^0}$ platí pro libovolné $f^0 \mid f$.) Takto postupujeme, dokud nenajdeme m netriviálních faktorů. Je samozřejmě třeba ukázat, že uvedený postup vede k cíli.

V popisu algoritmu ztotožňujeme polynomy stupně $< n$ s prvky $(\mathbb{F}_q)^n$. Proměnná F obsahuje v každém kroku rozklad polynomu f , který je v kroku 4. dále zjemňován.

Algoritmus 7 (Berlekampův).

vstup: $f \in \mathbb{F}_q[x]$ bezčtvercový monický polynom stupně n

výstup: ireducibilní rozklad g_1, \dots, g_m polynomu $f \in \mathbb{F}_q[x]$

1. $Q :=$ matice se sloupci $x^0 \pmod{f}, x^q \pmod{f}, \dots, x^{(n-1)q} \pmod{f}$
2. spočítáme bázi $h_1 = 1, h_2, \dots, h_m$ prostoru řešení soustavy rovnic $(Q - E)h = 0$
3. $i := 2, F := f$
4. **while** $\text{deg } F < m$ **do**
 nahraňme každé $g \in F$ netriviálními faktory z rozkladu

$$g = \prod_{a \in \mathbb{F}_q} \text{NSD}(g, h_i - a)$$

 $i := i + 1$
5. **return** F

Tvrzení 12.8. Algoritmus 7 funguje.

Důkaz. K důkazu správnosti zbývá ukázat, že každé dva různé ireducibilní faktory polynomu f budou odděleny pomocí nějakého polynomu h_k , tedy že pro každé i, j existuje k a různé prvky $a, b \in \mathbb{F}_q$ takové, že $g_i \mid \text{NSD}(f, h_k - a)$ a $g_j \nmid \text{NSD}(f, h_k - b)$. Protože $g_i, g_j \mid f$, stačí hledat k, a, b taková, že $g_i \mid h_k - a$ a $g_j \nmid h_k - b$, tj. taková, že $h_k \pmod{g_i} = a$ a $h_k \pmod{g_j} = b$. Ještě jinými slovy, pro každé i, j hledáme k takové, že $h_k \pmod{g_i} \neq h_k \pmod{g_j}$. Existence takového k plyne snadno z tvrzení 12.6: vektory $\varphi(h_1), \dots, \varphi(h_m)$ tvoří bázi prostoru $(\mathbb{F}_q)^m$, není tedy možné, aby měly všechny stejnou i -tou a j -tou složku { přepíšeme, že $\varphi(h) = (\dots, h \pmod{g_i}, \dots, h \pmod{g_j}, \dots)$.

Příklad. Uvažujme polynom

$$f = x^4 + 1 \in \mathbb{Z}_3[x].$$

Protože $\text{NSD}(f, f^0) = 1$, jde o bezetvercový polynom, takže můžeme použít Berlekampův algoritmus. Nejprve spočteme matici Q . Platí

$$\begin{aligned}x^0 \bmod f &= 1, \\x^3 \bmod f &= x^3, \\x^6 \bmod f &= 2x^2, \\x^9 \bmod f &= x,\end{aligned}$$

takže

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Gaussovou eliminací převedeme $Q \quad E$ do odstupňovaného tvaru

$$Q \quad E = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Báze prostoru $\mathbb{F}_q[x]$ je například $(1, 0, 0, 0)$, $(0, 1, 0, 1)$, což odpovídá polynomům $h_1 = 1$, $h_2 = x + x^3$. Polynom f se tedy rozkládá na 2 ireducibilní činitele. Spočteme

$$\begin{aligned}\text{NSD}(f, h_2 \quad 0) &= \text{NSD}(x^4 + 1, x^3 + x) = 1, \\ \text{NSD}(f, h_2 \quad 1) &= \text{NSD}(x^4 + 1, x^3 + x + 2) = x^2 + 2x + 2, \\ \text{NSD}(f, h_2 \quad 2) &= \text{NSD}(x^4 + 1, x^3 + x + 1) = x^2 + x + 2.\end{aligned}$$

Hledaný rozklad je tedy

$$x^4 + 1 = (x^2 + 2x + 2)(x^2 + x + 2).$$

Tvrzení 12.9. Časová složitost algoritmu 7 v tělese \mathbb{F}_q je $O(n^3ql(q)^2)$, kde $l(q)$ značí počet cifer čísla q .

Důkaz. Nejprve spočteme mocninu $x^q \bmod f$ binárním algoritmem se složitostí $O(n^2l(q)^3)$. V kroku 1. při výpočtu matice Q využijeme vztahu

$$x^{iq} \bmod f = (x^{(i-1)q} \bmod f) \cdot x^q \bmod f,$$

takže potřebujeme n dělení se zbytkem v $\mathbb{F}_q[x]$, přičemž dělíme x^q -násobek předchozího polynomu (stupně $< 2n$) polynomem f stupně n . Krok 1. tedy má složitost $O(n^3l(q)^3)$. Časová složitost Gaussovy eliminace matice $Q \quad E$ velikosti $n \times n$ nad tělesem \mathbb{F}_q je $O(n^3l(q)^2)$. Cyklus v kroku 4. projdeme nejvýše m -krát, $m \leq n$. Při daném průchodu pro každý prvek $g \in \mathbb{F}_q$ počítáme $\text{NSD}(g, h_i - a)$ pro každé $a \in \mathbb{F}_q$, tedy celkem q výpočtů NSD s každým g . Celková složitost kroku 4. tedy bude $m \cdot q \cdot \sum_{g \in \mathbb{F}_q} O(n \deg gl(q)^2) = O(n^2ql(q)^2(\sum_{g \in \mathbb{F}_q} \deg g)) = O(n^3ql(q)^2)$.

Vidíme, že Berlekampův algoritmus má exponenciální složitost vzhledem k délce čísla q , protože v kroku 4. procházíme všechny prvky a tělesa \mathbb{F}_q . Přitom ostatní kroky mají složitost pouze $O(n^3l(q)^3)$, jde tedy o úzké hrdlo tohoto postupu. Dosud není znám žádný deterministický algoritmus polynomiální vzhledem k n i $l(q)$, nicméně existuje řada pravděpodobnostních algoritmů (např. Cantor-Zassenhausův nebo Kaltofen-Shoupův) s polynomiální střední složitostí. První takový algoritmus objevil sám Berlekamp.

Další třídy algebraických struktur

13. Obecné algebraické struktury

13.1. Algebraické struktury.

Dosud se čtenář seznámil se základy tří klasických algebraických teorií: lineární algebry, komutativní algebry a teorie grup. Objektem studia každé z těchto disciplín je speciální typ algebraické struktury (vektorový prostor, komutativní okruh, grupa), popsáný jako množina, na níž jsou definovány nějaké operace, přičemž každá teorie na tyto objekty klade nějaké podmínky, tzv. axiomy, které vycházejí z vlastností, které sdílejí stejné příklady (např. v komutativní algebře obory polynomů a číselné obory). Tyto společné znaky jsou abstrahovány v definici obecného pojmu algebraické struktury.

Definice. Jazykem rozumíme množinu Σ spolu se zobrazením $ar : \Sigma \rightarrow \mathbb{N} \setminus \{0\}$. Význam této definice je následující: Σ je množina operačních symbolů, které budeme v dané teorii používat, a zobrazení ar udává aritu každého symbolu. Říkáme, že symbol $\sigma \in \Sigma$ je $ar(\sigma)$ -ární. Místo 1-ární říkáme unární, místo 2-ární říkáme binární. Pro binární symboly se zpravidla používají infixové znaky $+$, \cdot , \cdot apod., pro unární symboly se nikdy používají postfixové znaky 0 , 1 apod.

Buď A množina. n -ární operací na A rozumíme zobrazení z kartézské mocniny $A^n = A \times \dots \times A$ do A . Speciálně, 0-ární operace je zobrazení z jednoprvkové množiny do A , lze ji tedy interpretovat jako konstantu.

Algebraická struktura v jazyce Σ je dvojice $\mathbf{A} = (A, \Phi)$, kde A je neprázdňá množina, zvaná nosná množina, a Φ je zobrazení z množiny Σ do množiny všech operací na A přiřazující symbolu σ nějakou $ar(\sigma)$ -ární operaci $\sigma^{\mathbf{A}}$.

Algebraické struktury budeme značit tučným písmenem, jejich nosné množiny kurzívou, s výjimkou standardních značení, která jsme potkali dříve (číselné či polynomiální obory \mathbb{Q} , $\mathbb{Q}[x]$ apod.). Nebude-li výslovně uvedeno jinak, strukturu a její nosnou množinu značíme stejným písmenem, různým písmem. Struktury v konečném jazyce $f\sigma_1, \dots, \sigma_n g$ budeme zapisovat ve tvaru $\mathbf{A} = (A, \sigma_1^{\mathbf{A}}, \dots, \sigma_n^{\mathbf{A}})$.

Příklad. Grupy jsou algebraické struktury $\mathbf{G} = (G, \Phi)$ v jazyce $f, \cdot, e g$, kde

$$ar(\cdot) = 2, ar(e) = 1, ar(f) = 0,$$

splňující následující podmínky pro všechny prvky $a, b, c \in G$:

$$a \cdot^{\mathbf{G}} (b \cdot^{\mathbf{G}} c) = (a \cdot^{\mathbf{G}} b) \cdot^{\mathbf{G}} c,$$

$$a \cdot^{\mathbf{G}} e^{\mathbf{G}} = e^{\mathbf{G}} \cdot^{\mathbf{G}} a = a,$$

$$a \cdot^{\mathbf{G}} a^{\theta \mathbf{G}} = a^{\theta \mathbf{G}} \cdot^{\mathbf{G}} a = e^{\mathbf{G}}.$$

Příklad. Okruhy s jednotkou jsou algebraické struktury $\mathbf{R} = (R, \Phi)$ v jazyce $f+, \cdot, 0, 1 g$, kde $ar(+)$ = $ar(\cdot)$ = 2, $ar(0)$ = $ar(1)$ = 0, takové, že $(R, +^{\mathbf{R}}, \cdot^{\mathbf{R}}, 0^{\mathbf{R}})$ je abelovská grupa, operace $\cdot^{\mathbf{R}}$ je asociativní, pro operace $+^{\mathbf{R}}, \cdot^{\mathbf{R}}$ platí levý a pravý distributivní zákon a platí $a \cdot^{\mathbf{R}} 1^{\mathbf{R}} = 1^{\mathbf{R}} \cdot^{\mathbf{R}} a = a$ pro všechna $a \in R$.

Je-li z kontextu zřejmé, zda mluvíme o symbolu nebo příslušné operaci, budeme pro přehlednost vynechávat horní index.

Příklad. Latinský čtverec $(a_{i,j})_{i,j \in X}$ nad množinou X lze považovat za algebraickou strukturu (X, \cdot) s jednou binární operací, kde $u \cdot v = a_{u,v}$. Struktury vzniklé z latinských čtverců se nazývají kvazigrupy.

Následující dva příklady ukazují jisté zádrhele v definici struktury.

Příklad. Tříska lze považovat za algebraické struktury v jazyce okruhu s jednotkou. Můžeme postulovat, že pro každé $0 \neq a \in T$ existuje právě jedno $b \in T$ takové, že $a \cdot b = 1$ a toto b značí a^{-1} , avšak nejde o operaci ve výše uvedeném smyslu, neboť není definovaná pro nulu.

Příklad. Vektorové prostory nad tělesem T lze považovat za algebraické struktury v jazyce $(+, \cdot, 0, f_\alpha : \alpha \in T)$, kde $ar(+)=2$, $ar(\cdot)=1$, $ar(0)=0$ a $ar(f_\alpha)=1$ pro všechna $\alpha \in T$, takové, že $(V, +, \cdot, 0)$ je abelovská grupa a pro všechna $a, b \in V$, $\alpha, \beta \in T$, platí

$$\begin{aligned} f_{(\alpha+\tau\beta)}(a) &= f_\alpha(a) + f_\beta(a), & f_{\alpha\tau\beta}(a) &= f_\alpha(f_\beta(a)), \\ f_\alpha(a+b) &= f_\alpha(a) + f_\alpha(b), & f_1(a) &= a. \end{aligned}$$

Symboly f_α interpretujeme jako skalární násobení prvkem α , tj. $f_\alpha(v) = \alpha \cdot v$. Skalární součet není binární operace ve výše uvedeném smyslu, neboť jde o zobrazení $T \times V \rightarrow V$.

Další příklady zajímavých tříd algebraických struktur uvidíme v sekci 14 o svazech a Booleových algebrách.

Ve všech výše uvedených teoriích se opakují některé základní algebraické pojmy: podstruktury, direktního součiny, homomorfismy a faktorstruktury. Ve zbytku kapitoly dáme těmto konceptům společný rámec.

13.2. Podstruktury.

Definice. Buď f n -ární operace na množině A a $B \subseteq A$. Řekneme, že podmnožina B je uzavřena na operaci f , pokud pro všechna $b_1, \dots, b_n \in B$

$$f(b_1, \dots, b_n) \in B.$$

Buď $\mathbf{A} = (A, \Phi)$ struktura v jazyce Σ . Podstrukturou struktury \mathbf{A} rozumíme strukturu $\mathbf{B} = (B, \Psi)$ ve stejném jazyce, kde $B \subseteq A$ je uzavřena na všechny operace z Φ a kde Ψ obsahuje restrikce operací z Φ na množinu B , tj. $\sigma^{\mathbf{B}} = \sigma^{\mathbf{A}}|_B$ pro všechna $\sigma \in \Sigma$. Značíme $\mathbf{B} \subseteq \mathbf{A}$.

Uvedená definice je kompatibilní s definicemi podstruktur, které již znáte: podstruktura vektorového prostoru je totéž co podprostor, podstruktura okruhu je podokruh, podstruktura grupy je podgrupa.

Příklad. Na operacích dané struktury závisí: podmnožina \mathbb{N} tvoří podstrukturu struktury $(\mathbb{Z}, +)$, ale nikoliv (\mathbb{Z}, \cdot) .

Tvrzení 13.1 (průnik podstruktur). Buď \mathbf{A} struktura a \mathbf{B}_i , $i \in I$, její podstruktury. Pak $\bigcap_{i \in I} \mathbf{B}_i$ je buď prázdná množina, nebo tvoří podstrukturu struktury \mathbf{A} .

Jde-li o podstrukturu, budeme ji značit $\bigcap_{i \in I} \mathbf{B}_i$.

Důkaz. Označme $B = \bigcap_{i \in I} B_i$ a předpokládejme $B \neq \emptyset$. Buď σ symbol arity n a $b_1, \dots, b_n \in B$. Pak $b_1, \dots, b_n \in B_i$ pro všechna $i \in I$, tedy $\sigma^{\mathbf{A}}(b_1, \dots, b_n) \in B_i$ pro všechna $i \in I$, neboť každá množina B_i je na tuto operaci uzavřena, a tedy $\sigma^{\mathbf{A}}(b_1, \dots, b_n) \in \bigcap_{i \in I} B_i = B$.

Tvrzení 13.2 (sjednocení podstruktur). Buď \mathbf{A} struktura a $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \dots$ její podstruktury. Pak $\bigcup_{i \in \mathbb{N}} \mathbf{B}_i$ tvoří podstrukturu struktury \mathbf{A} .

Tuto podstrukturu budeme značit $\bigcup_{i \in \mathbb{N}} B_i$.

Důkaz. Označme $B = \bigcup_{i \in \mathbb{N}} B_i$. Buď σ symbol arity n a $b_1, \dots, b_n \in B$. Pak existuje $k \in \mathbb{N}$ takové, že $b_1, \dots, b_n \in B_k$: vzhledem k tomu, že každé $b_j \in B_{k_j}$ pro nějaké $k_j \in \mathbb{N}$, stačí zvolit $k = \max(k_1, \dots, k_n)$. Pro toto k pak platí $\sigma^A(b_1, \dots, b_n) \in B_k$, a tedy $\sigma^A(b_1, \dots, b_n) \in \bigcup_{i \in \mathbb{N}} B_i = B$.

Uvažujme podmnožinu $X \subseteq A$ v struktuře A . Podstrukturou generovanou množinou X rozumíme nejmenší podstrukturu (vzhledem k inkluzi) struktury A obsahující podmnožinu X , značíme ji $\langle X \rangle_A$. Taková podstruktura jistě existuje: stačí vzít průnik všech podstruktur obsahujících množinu X , tj.

$$\langle X \rangle_A = \bigcap_{X \subseteq B, B \text{ A}} B.$$

Podle Tvzení 13.1 je tento průnik skutečně podstrukturou, jistě obsahuje množinu X a mezi všemi takovými podstrukturami je nejmenší.

Prvky podstruktury $\langle X \rangle_A$ můžeme najít tak, že začneme s prvky množiny X a aplikováním operací struktury A získáváme postupně další prvky. Pokud již žádné nové prvky nevznikají, tedy když je výsledná podmnožina uzavřena na všechny operace struktury A , našli jsme podstrukturu.

Příklad. Na operacích dané struktury závisí: pro dané $n \in \mathbb{Z}$

$$\langle n \rangle_{(\mathbb{Z}, +)} = \{kn : k \in \mathbb{N}\},$$

$$\langle n \rangle_{(\mathbb{Z}, \cdot)} = \{n^k : k \in \mathbb{N}\}.$$

13.3. Homomorfismy a izomorfismy.

Definice. Buď A, B struktury ve stejném jazyce Σ . Zobrazení $\varphi : A \rightarrow B$ se nazývá homomorfismus struktur A, B , pokud

$$\varphi(\sigma^A(a_1, \dots, a_n)) = \sigma^B(\varphi(a_1), \dots, \varphi(a_n))$$

pro každý n -ární symbol $\sigma \in \Sigma$ a všechny $a_1, \dots, a_n \in A$. Říkáme, že φ zachovává operace těchto algeber. Píšíme $\varphi : A \rightarrow B$.

Pro speciální typy homomorfismů se používá následující terminologie:

vnoření je prostý homomorfismus (někdy se užívá značení $A \hookrightarrow B$),
 izomorfismus je homomorfismus, který je bijekcí (značení $A \cong B$),

a dále

endomorfismus struktury A rozumíme homomorfismus $A \rightarrow A$,
 automorfismus struktury A rozumíme izomorfismus $A \rightarrow A$.

Identické zobrazení $\text{id} : A \rightarrow A, x \mapsto x$, je vždy homomorfismus.

Obor hodnot daného homomorfismu $\varphi : A \rightarrow B$ se v algebře říká obraz a značí se $\text{Im}(\varphi) = \{\varphi(a) : a \in A\}$. Obraz vždy tvoří podstrukturu struktury B : je-li σ n -ární symbol a $b_1, \dots, b_n \in \text{Im}(\varphi)$, pak $b_1 = \varphi(a_1), \dots, b_n = \varphi(a_n)$ pro nějaká $a_1, \dots, a_n \in A$ a platí

$$\sigma^B(b_1, \dots, b_n) = \sigma^B(\varphi(a_1), \dots, \varphi(a_n)) = \varphi(\sigma^A(a_1, \dots, a_n)) \in \text{Im}(\varphi).$$

Homomorfismy jsou jednoznačně určeny svými hodnotami na generátorech. (Ale není pravda, že dané hodnoty na generátorech lze rozšířit do homomorfismu: to je speciálně u vektorových prostorů, nebo obecně tzv. volných algeber.) Tento princip lze využít k hledání všech homomorfismů mezi dvěma strukturami.

Úloha. Najděte všechny homomorfismy $(\mathbb{Z}, +) \rightarrow (\mathbb{Z}, +)$.

Ŕešení. Všimníte si, že $(\mathbb{Z}, +) = \langle 1 \rangle$, tedy z hodnoty 1 pŕijde spočítat všechny ostatní hodnoty. Uvažujme homomorfismus φ . Je-li $\varphi(1) = k$, dokážeme indukci, že $\varphi(a) = ka$ pro všechna a . Nejprve spočteme $\varphi(0) = \varphi(1 - 1) = \varphi(1) - \varphi(1) = k - k = 0$ a $\varphi(-1) = \varphi(0 - 1) = \varphi(0) - \varphi(1) = 0 - k = -k$. V indukčním kroku pak pro kladná a platí $\varphi(a) = \varphi((a-1) + 1) = \varphi(a-1) + \varphi(1) = k(a-1) + k = ka$ a podobně pro postupujeme pro záporná a . Zbývá ověřit, že jsme skutečně dostali homomorfismus: $\varphi(a+b) = k(a+b) = ka + kb = \varphi(a) + \varphi(b)$ pro všechna a, b .

Pokud nejsme schopni efektivně použít generující množinu, jako třeba v následující úloze, můžeme zkusit využít prvků se zvláštními vlastnostmi, které jsou zachovány každým homomorfismem (viz též diskuse invariantů v sekci 1.3).

Úloha. Najděte všechny homomorfismy $\text{sm}y (\mathbb{Z}, +) \rightarrow (\mathbb{Z}, +)$.

Ŕešení. Uvažujme homomorfismus φ . Z rovnosti $\varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0)$ plyne $\varphi(0) = 0$ a dostáváme $0 = \varphi(0) = \varphi(n+0) = \varphi(n) + \varphi(0) = \varphi(n)$ pro každé $n \in \mathbb{Z}$. Existuje tedy jediný homomorfismus $\text{sm}y \mathbb{Z} \rightarrow 0$.

Podobně jako pro grupy a okruhy se dokáží následující vlastnosti.

Tvrzení 13.3. Buď A, B, C struktury ve stejném jazyce a $\varphi: A \rightarrow B$ a $\psi: B \rightarrow C$ homomorfismy. Pak

- (1) složené zobrazení $\psi \circ \varphi$ je homomorfismus $A \rightarrow C$;
- (2) je-li φ izomorfismus, pak inverzní zobrazení φ^{-1} je izomorfismus $B \rightarrow A$.

Z Tvrzení 13.3 plyne, že automorfismy smy dané struktury A tvoří podgrupu symetrické grupy S_A , značíme $\text{Aut}(A)$.

Ŕekneme, že struktury A a B jsou izomorfní, značíme $A \cong B$, pokud existuje izomorfismus $A \rightarrow B$. Stejně jako pro grupy a okruhy, izomorfismus si lze představit jako kopírování operací z jedné nosné množiny na druhou, tj. dvě algebraické struktury jsou izomorfní, pokud se liší pouze přejmenováním prvků. Z Tvrzení 13.3 plyne, že izomorfismus dává ekvivalenci na třídě všech algeber v daném jazyce.

Připomeňme, že invariantem izomorfismu rozumíme vlastnost V takovou, že kdykoliv A má vlastnost V a $B \cong A$, pak B má také vlastnost V . O grupových invariantech jsme psali v sekci 1.3 a řada z nich má obecnou platnost (minimální počet generátorů, rovnosti, jisté typy význačných prvků). Obecně lze říci, že invariantem je jakákoliv vlastnost, kterou lze vyjádřit pomocí tzv. formulí prvního řádu v daném jazyce, tj. pomocí výrazů používajících kvantifikátory, proměnné, logické spojky, rovnítko a operace z daného jazyka. Přesnou formulaci a důkaz najdete ve většině učebnic matematické logiky.

13.4. Kongruence a faktorstruktury.

V úvodu sekce 2.2 jsme psali o tom, jak se napříč matematikou opakuje myšlenka konstrukce faktorobjektu, ztotožněním navzájem blízkých prvků. Nyní popíšeme, jak tato konstrukce funguje pro obecné algebraické struktury.

Buď $A = (A, \Phi)$ algebraická struktura v jazyce Σ . Uvažujme ekvivalenci \sim na množině A , která nám bude říkat, které prvky chceme ztotožnit. Operace faktorstruktury A/\sim bychom rádi definovali tak, že výsledek n -ární operace σ^A na blocích $[a_1], \dots, [a_n]$ by měl být roven bloku $[\sigma^A(a_1, \dots, a_n)]$. Ovšem aby byla taková operace dobře definovaná, ekvivalence \sim nemůže být ledajaká. Takové ekvivalence se nazývají kongruence.

Definice. Buď A struktura v jazyce Σ . Ekvivalence \sim na nosné množině A se nazývá kongruence struktury A , pokud pro každý n -ární symbol $\sigma \in \Sigma$ a všechna $a_1, \dots, a_n, b_1, \dots, b_n \in A$ platí

$$(\sigma^A(a_1, \dots, a_n) \sim \sigma^A(b_1, \dots, b_n)).$$

Pro unární symbol 0 podmínka říká, že pokud $a \sim b$, pak $a^0 \sim b^0$. Pro binární symbol \sim podmínka říká

$$(a \sim b, c \sim d) \Rightarrow (a \cdot c \sim b \cdot d),$$

což, jak si snadno lze odvodit, je ekvivalentní podmínce $(a \sim b) \Rightarrow (a \cdot c \sim b \cdot c)$ a $(c \sim a) \Rightarrow (c \cdot b \sim a \cdot b)$ pro všechna c . Konstanty nehrají u kongruencí žádnou roli, protože $c \sim c$ pro každé c .

Příklad. Uvažujme strukturu $(\mathbb{Z}, +, \cdot)$. V zimním semestru jsme si dokázali, že relace $\sim \pmod{n}$ je kongruencí této struktury.

Vzhledem k tomu, že obecné algebraické struktury nemusí obsahovat význačné prvky (jednotka v grupách, nula v okruzích, apod.), nelze očekávat, že by kongruence byly určeny nějakými podobjekty (jako jsou normální podgrupy či ideály).

Definice. Buď \mathbf{A} algebraická struktura a její kongruence. Uvažujme množinu $A/\sim = \{[a] : a \in A\}$ a definujme na ní operace odpovídajícím způsobem

$$\sigma^{\mathbf{A}}([a_1], \dots, [a_n]) = [\sigma^{\mathbf{A}}(a_1, \dots, a_n)]$$

pro každý n -ární symbol $\sigma \in \Sigma$ a všechna $a_1, \dots, a_n \in A$. Z definice kongruence vidíme, že jsou operace dobře definované: pokud označíme bloky jiným způsobem, tj. pokud $[a_i] = [b_i], \dots, [a_n] = [b_n]$, pak $[\sigma^{\mathbf{A}}(a_1, \dots, a_n)] = [\sigma^{\mathbf{A}}(b_1, \dots, b_n)]$, čímž výsledek operace je na označení nezávislý. Struktura

$$\mathbf{A}/\sim = (A/\sim, (\sigma^{\mathbf{A}} : \sigma \in \Sigma))$$

se nazývá faktorstruktura \mathbf{A} podle kongruence \sim .

Buď $\varphi : \mathbf{A} \rightarrow \mathbf{B}$ homomorfismus. Jeho jádrem rozumíme relaci na A definovanou

$$a \sim_{\varphi} b \iff \varphi(a) = \varphi(b).$$

Následující tvrzení říká, že jádro je kongruencí struktury \mathbf{A} , a že každá kongruence je jádrem nějakého homomorfismu.

Tvrzení 13.4 (jádra vs. kongruence). Buď \mathbf{A} struktura a relace na její nosné množině A . Pak \sim je kongruencí struktury \mathbf{A} právě tehdy, když je jádrem nějakého homomorfismu z \mathbf{A} do nějaké struktury \mathbf{B} .

Důkaz. (\Rightarrow) Uvažujme zobrazení

$$\varphi : A \rightarrow A/\sim, \quad a \mapsto [a].$$

Z definice operací faktorstruktury ihned plyne, že jde o homomorfismus $\mathbf{A} \rightarrow \mathbf{A}/\sim$. Jeho jádrem jsou právě ty dvojice (a, b) , pro které $[a] = [b]$, tedy $a \sim b$.

(\Leftarrow) Uvažujme nějaký homomorfismus $\varphi : \mathbf{A} \rightarrow \mathbf{B}$. Jeho jádro je zřejmě ekvivalencí, dokážeme, že jde o kongruenci. Uvažujme n -ární symbol σ . Buď $a_1, \dots, a_n, b_1, \dots, b_n \in A$ taková, že $a_i \sim_{\varphi} b_i$, tj. $\varphi(a_i) = \varphi(b_i)$, pro všechna i . Pak

$$\begin{aligned} \varphi(\sigma^{\mathbf{A}}(a_1, \dots, a_n)) &= \sigma^{\mathbf{B}}(\varphi(a_1), \dots, \varphi(a_n)) \\ &= \sigma^{\mathbf{B}}(\varphi(b_1), \dots, \varphi(b_n)) = \varphi(\sigma^{\mathbf{A}}(b_1, \dots, b_n)), \end{aligned}$$

a tedy $\sigma^{\mathbf{A}}(a_1, \dots, a_n) \sim_{\varphi} \sigma^{\mathbf{A}}(b_1, \dots, b_n)$.

Podobně jako pro grupy a okruhy platí věta o homomorfismu a 1. věta o izomorfismu.

Věta 13.5 (věta o homomorfismu). Buď $\varphi : \mathbf{A} \rightarrow \mathbf{B}$ homomorfismus algeber.

(1) Je-li kongruence struktury \mathbf{A} obsažená v jádru \sim_{φ} , pak je zobrazení

$$\psi : \mathbf{A}/\sim \rightarrow \mathbf{B}, \quad [a] \mapsto \varphi(a)$$

dobře definované a je to homomorfismus.

(2) (1. věta o izomorfismu) $\mathbf{A}/\varphi' \cong \mathbf{Im}(\varphi)$.

Důkaz. (1) Přední je třeba ověřit, že je zobrazení ψ dobře definované: pokud $[a] = [b]$, tj. pokud $a \sim b$, pak $a \sim_{\varphi} b$, a tedy $\varphi(a) = \varphi(b)$. Je to homomorfismus, protože pro n -ární symbol σ a každé $a_1, \dots, a_n \in A$ platí

$$\begin{aligned} \psi(\sigma^{\mathbf{A}}([a_1], \dots, [a_n])) &= \psi([\sigma^{\mathbf{A}}(a_1, \dots, a_n)]) = \varphi(\sigma^{\mathbf{A}}(a_1, \dots, a_n)) \\ &= \sigma^{\mathbf{B}}(\varphi(a_1), \dots, \varphi(a_n)) = \sigma^{\mathbf{B}}(\psi([a_1]), \dots, \psi([a_n])). \end{aligned}$$

(2) Vztáhneme část (1) na samu kongruenci \sim_{φ} : výsledný homomorfismus ψ je prostý, neboť $a \sim b, \varphi(a) = \varphi(b)$.

Poznámka. Konstrukce faktorgrup a faktorokruhů je speciálním případem konstrukce faktorstruktury.

Buď G grupa a N její normální podgrupa. Jak jsme si ukázali v sekci 2.2, relace dobře definovaná $a \sim b, ab^{-1} \in N$ je kongruencí grupy G . Struktury G/N a G/\sim jsou totožné.

Buď R okruh a I jeho ideál. Jak jsme si ukázali v sekci 4.3, relace dobře definovaná $a \sim b, a - b \in I$ je kongruencí okruhu R . Struktury R/I a R/\sim jsou totožné.

Naopak, žádné jiné faktorstruktury grup a okruhů nejsou: kongruence a normální podgrupy, resp. kongruence a ideály, si vzájemně jednoznačně odpovídají.

Je-li \sim je kongruencí grupy G , blok $[1]$ je normální podgrupou v G : pokud $a, b \in [1]$, tj. $a \sim b \sim 1$, pak $a \sim b \sim 1 \sim 1 = 1$ a $a^{-1} \sim 1^{-1} = 1$, èili jde o podgrupu. Pokud $a \sim 1$ a $g \in G$, pak $gag^{-1} \sim g1g^{-1} = 1$, což prokazuje normalitu.

Je-li \sim je kongruencí okruhu R , blok $[0]$ je ideálem v R : výše jsme dokázali, že jde o aditivní podgrupu, a pokud $a \sim 0$ a $r \in R$, pak $ar \sim 0r = 0$ a $ra \sim r0 = 0$, èili je to ideál.

14. Uspořádání a svazy

14.1. Uspořádané množiny.

V této části zopakujeme základní pojmy týkající se uspořádání. Většinu z nich by měl čtenář znát z předchozích kurzů.

Definice. Relaci \leq na množině A nazýváme částečné uspořádání, pokud je

- (1) reflexivní, tj. $a \leq a$ pro všechna $a \in A$,
- (2) tranzitivní, tj. $a \leq b$ a $b \leq c$ implikuje $a \leq c$,
- (3) antisymetrická, tj. $a \leq b$ a $b \leq a$ implikuje $a = b$.

Alternativně říkáme, že (A, \leq) je uspořádaná množina. Uspořádání se nazývá lineární, pokud navíc pro každé a, b nastane $a \leq b$ nebo $b \leq a$. Intervalem rozumíme množinu

$$[a, b] = \{u \in A : a \leq u \leq b\}.$$

Pokud $a \leq b$ a $a \neq b$, píšeme $a < b$.

Příklad. Na množině přirozených čísel jsou dvě přirozená uspořádání:

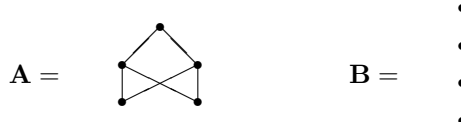
- (\mathbb{N}, \leq) dané standardním uspořádáním $1 < 2 < 3 < \dots$ (toto uspořádání je lineární);
- (\mathbb{N}, \preceq) dané dělitelností, tj. $a \preceq b$ je menší než b pokud $a \mid b$ (toto uspořádání není lineární).

Příklad. Na množině množin je přirozené brát uspořádání inkluzí:

- ($\mathcal{P}(X), \subseteq$): na množině $\mathcal{P}(X)$ všech podmnožin dané množiny X ťekneme, že $A \subseteq B$ je menší než B , pokud $A \subseteq B$;

$(Eq(X), \subseteq)$: na množině $Eq(X)$ všech ekvivalencí na množině X ůekneme, \mathbb{p} je menší než \mathbb{y} , pokud $\mathbb{p} \subseteq \mathbb{y}$, tj. pokud $a \sim b$ implikuje $a \sim \mathbb{y}$; analogicky lze uspořádat podstruktury dané struktury (inkluzí jejich nosných množin), či kongruence dané struktury.

Konečně uspořádané množiny lze popsat pomocí tzv. Hasseova diagramu. Jde o graf relace \leq , přičemž nekreslíme smyčky (reflexivita), vynecháváme všechny hrany, jejich existence je zaručena tranzitivitou, a místo šipek kreslíme neorientované hrany tak, aby větší prvky byly výše. Například:



Definice. ůekneme, \mathbb{a} prvek $a \in A$ je v (A, \leq)

- největší, pokud pro každé $b \in A$ platí $b \leq a$;
- nejmenší, pokud pro každé $b \in A$ platí $b \geq a$;
- maximální, pokud neexistuje žádné $b \in A$ takové, $\mathbb{a} b > a$;
- minimální, pokud neexistuje žádné $b \in A$ takové, $\mathbb{a} b < a$.

Příklady.

Uspořádaná množina A má jeden největší prvek, jeden maximální (ten samý), žádné nejmenší a dva minimální prvky.

Uspořádaná množina B má jeden největší (a zároveň maximální) a jeden nejmenší (a zároveň minimální) prvek. Je to lineární uspořádání.

Uspořádané množiny (\mathbb{N}, \leq) i (\mathbb{N}, \geq) mají nejmenší prvek 1, ale žádné maximální prvky.

Uspořádaná množina $(\mathbb{N} \setminus \{1\}, \leq)$ má za minimální prvky právě všechna prvočísla.

Definice. Nechť $B \subseteq A$. ůekneme, \mathbb{a} prvek $a \in A$ je v (A, \leq)

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

Příklady.

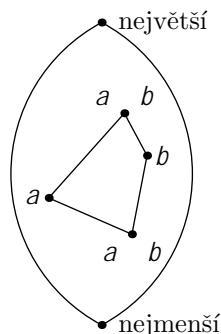
V uspořádané množině A podmnožina sestávající z obou minimálních prvků nemá supremum ani infimum. Infimum proto, \mathbb{a} nemá ani žádné dolní mez. Horní meze sice tato podmnožina má tři, avšak žádná z nich není nejmenší.

V každé lineární uspořádané množině má každá neprázdňá konečná podmnožina supremum i infimum, přičemž $\sup B = \max B$ a $\inf B = \min B$. Pro nekonečné to existovat nemusí, např. $\sup \mathbb{N}$ v (\mathbb{N}, \leq) .

V uspořádané množině $(P(X), \subseteq)$ má každá podmnožina infimum i supremum, přičemž $\inf B$ je rovno průniku všech množin z B a $\sup B$ je rovno sjednocení všech množin z B .

V uspořádané množině (\mathbb{N}, \geq) má každá konečná podmnožina infimum i supremum, přičemž $\inf B$ je rovno NSD všech čísel z B a $\sup B$ je rovno NSN všech čísel z B . Na druhou stranu, např. supremum množiny všech prvočísel neexistuje.

Definice. Uspořádanou množinu nazveme svazově uspořádanou, pokud v ní existují suprema a infima všech dvouprvkových podmnožin (indukcí je snadné dokázat,



Obrázek 8. Průsek a spojení ve svazově uspořádané množině

Je pak existují i suprema a in ma v'ech neprázdných konečných podmnožin). Nazveme ji úplně svazově uspořádanou, pokud existují suprema a in ma v'ech podmnožin. Ve svazově uspořádaných množinách obvykle značíme zkráceně

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

symboly $_$, \wedge čteme jako spojení a průsek.

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

Příklady.

Lineárně uspořádané množiny jsou vždy svazově uspořádané, $a _ b = \max(a, b)$, $a \wedge b = \min(a, b)$. Úplně svazově uspořádané být nemusí, příkladem je $(\mathbb{N}, _)$. $(P(X), _)$ je úplně svazově uspořádaná množina, pro $U \subseteq P(X)$ je $\sup U = \bigcup_{A \in U} A$, $\inf U = \bigcap_{A \in U} A$. $(\mathbb{N}, _)$ je svazově uspořádaná množina (ne úplně): $a _ b = \text{NSN}(a, b)$, $a \wedge b = \text{NSD}(a, b)$.

K ověření, zda je dané uspořádání úplně svazové, stačí ověřit pouze existenci in m, anebo pouze existenci suprem.

Tvrzení 14.1. Uspořádaná množina, ve které existují in ma v'ech podmnožin, je úplně svazově uspořádaná. Uspořádaná množina, ve které existují suprema v'ech podmnožin, je úplně svazově uspořádaná.

Důkaz. Označme danou uspořádanou množinu $(A, _)$. Zde nice suprema plyne, že

$$\sup B = \inf \{a \in A : a _ b \text{ pro každé } b \in B\},$$

tedy suprema lze de novat pomocí in m. Duálně, $\inf B = \sup \{a \in A : a _ b \text{ pro každé } b \in B\}$.

Pro některé algebraické konstrukce se hodí Zornovo lemma. Jde o jednu z forem tzv. axiomu výběru, jednoho ze základních axiomů teorie množin. Zornovo lemma se proto nedokazuje, nýbrž postuluje. Četizcem v částečně uspořádané množině rozumíme podmnožinu, která je lineárně uspořádaná.

Axiom 14.2 (Zornovo lemma). Buď $(X, _)$ neprázdná částečně uspořádaná množina. Předpokládejme, že každý četizec má horní mez. Pak $(X, _)$ obsahuje aspoď jeden maximální prvek.

Na závěr si vyjasníme, jak to je s izomorfiem uspořádaných množin. Buď (A, \leq) a (B, \leq) dvě uspořádané množiny. Zobrazení $\varphi: A \rightarrow B$ nazveme monotónní, pokud pro každé $a, b \in A$ platí

$$a \leq b \implies \varphi(a) \leq \varphi(b).$$

Zobrazení φ nazveme izomorfiem směrů těchto uspořádaných množin, je-li bijektivní a obě zobrazení φ, φ^{-1} jsou monotónní. Složení monotónních zobrazení je monotónní, ale pozor, inverz monotónní bijekce nemusí být monotónní: příkladem je identické zobrazení $x \mapsto x$ mezi uspořádanými množinami (\mathbb{N}, \leq) a (\mathbb{N}, \geq) , které je monotónní pouze v jednom směru ($a \leq b$ implikuje $a \geq b$, ale nikoliv naopak).

14.2. Svazy a Booleovy algebry.

Na svazově uspořádané množině lze pohledět dvěma způsoby: jako na uspořádané množině, kde má každá dvojice prvků supremum i infimum, ale také jako na algebraické struktury s dvěma binárními operacemi suprema a infima. Vlastností suprema a infima lze popsat abstraktně, pomocí jisté sady axiomů.

Definice. Algebraická struktura (A, \wedge, \vee) s dvěma binárními operacemi se nazývá svaz, pokud platí následující podmínky pro všechna $a, b, c \in A$:

$$\begin{aligned} (a \wedge b) \wedge c &= a \wedge (b \wedge c), & a \wedge b &= b \wedge a, & a \wedge a &= a, \\ (a \vee b) \vee c &= a \vee (b \vee c), & a \vee b &= b \vee a, & a \vee a &= a, \\ a \wedge (a \vee b) &= a, & a \vee (a \wedge b) &= a & \text{(tato podmínka se nazývá absorpce)}. \end{aligned}$$

Tvrzení 14.3 (svazy jako algebry vs. uspořádání).

- (1) Je-li (A, \leq) svazově uspořádaná množina, pak je struktura (A, \inf, \sup) svazem.
- (2) Je-li struktura (A, \wedge, \vee) svazem a dále máme-li $a \leq b$, $a \wedge b = a$, pak je (A, \leq) svazově uspořádanou množinou.

Důkaz. (1) Ověříme axiomy svazů. Je snadné (byť trochu pracně) ověřit, že $(a \wedge b) \wedge c = \inf \{a, b, c\}$, $a \wedge b = \inf \{a, b\}$ a $a \wedge a = \inf \{a\} = a$. Analogická tvrzení platí pro \vee a \sup . Dále $a \wedge (a \vee b) = \inf \{a, \sup \{a, b\}\} = a$, nebo $a \vee (a \wedge b) = \sup \{a, \inf \{a, b\}\} = a$, a analogicky ověříme druhou absorpční podmínku.

(2) Reexista plyne z toho, že $a \wedge a = a$. Transitivita: pokud $a \leq b \leq c$, tedy $a \wedge b = a$ a $b \wedge c = b$, pak $a \wedge c = (a \wedge b) \wedge c = a \wedge (b \wedge c) = a \wedge b = a$, či $a \leq c$. Antisymetrie: pokud $a \leq b$ a $b \leq a$, pak $a = a \wedge b = b \wedge a = b$.

Dokážeme existenci infima, a to tak, že $\inf \{a, b\} = a \wedge b$. Předně $a \wedge b$ je dolní mezí obou prvků, nebo $(a \wedge b) \wedge a = (a \wedge a) \wedge b = a \wedge b$, a analogicky pro b . Dále, je-li $c \leq a, b$ jiná dolní mez, pak $(a \wedge b) \wedge c = a \wedge (b \wedge c) = a \wedge c = c$, tedy $c \leq a \wedge b$.

Nyní si všimněte, že podmínku $a \leq b$ lze ekvivalentně vyjádřit jako $a \vee b = b$: pokud $a \wedge b = a$, z absorpce plyne $a \vee b = (a \wedge b) \vee b = b$; a naopak, pokud $a \vee b = b$, z absorpce plyne $a \wedge b = a \wedge (a \vee b) = a$.

Pomocí tohoto pozorování je snadné dokázat existenci suprema, a to tak, že $\sup \{a, b\} = a \vee b$: použijeme stejný argument jako pro infima s operací \vee místo \wedge .

Více o svazech najdete v libovolné učebnici univerzální algebry.

Jedním ze základních algebraických nástrojů matematické logiky jsou Booleovy algebry. Uvažujme logické hodnoty T, F . Pokud interpretujeme \wedge a \vee jako konjunktci a disjunktci, dostaneme svaz, a spolu s negací je to mnohem bohatší struktura.

Definice. Algebraická struktura $(A, \wedge, \vee, \neg, 0, 1)$ s dvěma binárními operacemi \wedge, \vee , unární operací \neg a konstantami $0, 1$ se nazývá Booleova algebra, pokud platí následující podmínky pro všechna $a, b, c \in A$:

$(A, \wedge, _)$ je svaz,
 $a \wedge (b _ c) = (a \wedge b) _ (a \wedge c)$ (distributivita),
 $a \wedge 0 = 0, a _ 1 = 1,$
 $a \wedge a^\theta = 0, a _ a^\theta = 1.$

Z axiomů je snadné odvodit další užitečné vlastnosti:

$a _ (b \wedge c) = (a _ b) \wedge (a _ c)$ (duální distributivita),
 $(a^\theta)^\theta = a, 0^\theta = 1, 1^\theta = 0,$
 $(a \wedge b)^\theta = a^\theta _ b^\theta$ a $(a _ b)^\theta = a^\theta \wedge b^\theta$ (de Morganovy zákony).

Důkazy si čtenář provede snadno sám.

Příklad. Základním příkladem Booleovy algebry je množinová algebra

$$(P(X), \cap, \setminus, \bar{}, \cup, X),$$

s operacemi průniku, sjednocení a doplňku množin ($\bar{A} = X \setminus A$).

Není třeba dokázat, že každá konečná Booleova algebra je izomorfní některé množinové algebře. Nekonečných Booleových algeber existuje spousta různých typů.

Příklad. Jak jsme již zmínili, pravdivostní hodnoty T, F spoly s operacemi konjunkce, disjunkce a negace tvoří Booleovu algebru $(\{T, F\}, \wedge, _, \neg, \vee, T, F)$. Tato algebra se váže k základní výrokové logice.

Uvažujme nyní obecnou teorii T v jazyce L (například teorii grup či Peanovu aritmetiku). Označme F_L množinu všech formulí v jazyce L . Dvě formule φ, ψ nazveme T -ekvivalentní, pokud lze v teorii T dokázat ekvivalenci $\varphi \sim \psi$. Uvažujme množinu F_T , do které dáme z každé třídy ekvivalence po jedné formuli. Tzv. Lindenbaumova algebra teorie T bude Booleova algebra $(F_T, \wedge, _, \neg, \vee, F, T)$.

Lindenbaumovy algebry měří neúplnost dané teorie: pokud je každá formule z axiomů dokazatelná či vyvratitelná, F_T bude mít pouze dva prvky, v opačném případě jich bude víc (jako třeba pro Peanovu aritmetiku, jak říká Gödelova věta o neúplnosti).

Více o teorii a použití Booleových algeber najdete ve většině učebnic logiky či teorie množin. Jiným myšlenkovým směrem jsou pak neklasické logiky, které vycházejí ze zobecnění Booleových algeber.