

NEŠTRUKTÚROVANÉ ROZPRÁVY O ŠTRUKTÚRACH

Kapitoly z matematiky pre informatikov

1

Peter Kostolányi

Bratislava 2025

Autor: Peter Kostolányi
Katedra informatiky FMFI UK v Bratislave

Názov: Neštruktúrované rozpravy o štruktúrach:
kapitoly z matematiky pre informatikov 1

Vydavateľ: Knížničné a edičné centrum FMFI UK v Bratislave

Recenzenti: doc. RNDr. Tatiana Jajcayová, PhD.
doc. RNDr. Martin Mačaj, PhD.

Rok vydania: 2025

Miesto vydania: Bratislava

Vydanie: prvé

Počet strán: 252

ISBN 978-80-8147-158-2

Dielo je vydané pod medzinárodnou licenciou Creative Commons CC BY-NC-ND 4.0 (vyžaduje sa: povinnosť uvádzať pôvodného autora diela; len nekomerčné použitie; žiadne odvodené diela).



Viac informácií o licencií a použití diela:
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Obsah

Úvod	1
Prehľad označení a konvencií	3
1 Grafy, matice a polokruhy	9
1.1 Matica susednosti ohodnoteného grafu	10
1.2 Násobenie matíc a niektoré úlohy na grafoch	12
1.3 Polokruhy	18
1.4 Úplné a spočítateľne úplné polokruhy	22
1.5 Výpočet iterácie matice nad spočítateľne úplným polokruhom	25
1.6 Konečné automaty nad polokruhmi	30
1.7 Normálny tvar konečných automatov nad polokruhmi	32
1.8 Racionálne prvky spočítateľne úplného polokruhu	33
Cvičenia	36
Poznámky a odkazy na samoštúdium	37
2 Formálne mocninové rady	39
2.1 Formálne mocninové rady o jednej premennej	40
2.2 Vytvárajúce funkcie a ich použitie	49
2.3 Rady o niekoľkých komutujúcich premenných	52
2.4 Rady o niekoľkých nekomutujúcich premenných	53
2.5 Matice formálnych mocninových radov	57
2.6 Iterácia súčtu a iterácia súčinu	59
Cvičenia	62
Poznámky a odkazy na samoštúdium	62
3 Automaty s váhami	65
3.1 Príklady na úvod	65
3.2 Konečné automaty s váhami nad polokruhom	68
3.3 Alternatívny pohľad na automaty s váhami	70
3.4 Normálne tvary automatov s váhami	73
3.5 Racionálne mocninové rady	75
3.6 Lineárne reprezentácie	81
3.7 Automaty s váhami nad poľom	84
3.8 Odbočka do lineárnej algebry	92
3.9 Minimalizácia automatov s váhami nad poľom	94
3.10 Racionálne rady nad poľom a unárnou abecedou	101
3.11 Rozhodovacie problémy pre automaty s váhami nad poľom	104
3.12 Využitie automatov s váhami na kódovanie obrazu	106
3.13 Conwayovské a čiastočné conwayovské polokruhy	110
Cvičenia	113
Poznámky a odkazy na samoštúdium	115

4	Trocha lineárnej algebry	117
4.1	Vlastné čísla, vlastné vektory a diagonalizácia matice	118
4.2	Spektrum, algebraická násobnosť a geometrická násobnosť	126
4.3	Ukážka aplikácií diagonalizácie	130
4.4	Mocniny lineárnych zobrazení	135
4.5	Zovšeobecnené vlastné vektory	136
4.6	Jordanov kanonický tvar matice	142
4.7	Mocniny matíc v Jordanovom kanonickom tvare	150
4.8	Spektrálne vlastnosti jednoduchých operácií na maticiach	151
4.9	Vlastné čísla a stopa matice	152
	Cvičenia	153
	Poznámky a odkazy na samoštúdium	154
5	Vlastné čísla a grafy I	155
5.1	Vandermondova a zovšeobecnená Vandermondova matica	155
5.2	Počet sledov dĺžky t v multigrafe	164
5.3	Počet slov dĺžky t v racionálnom jazyku	167
5.4	Perronova-Frobeniova teória nezáporných matíc	169
5.5	Primitívne matice a aperiodické grafy	181
5.6	Aplikácie Perronovej-Frobeniovej teórie	187
5.7	Maticovo-grafový slovník	191
	Cvičenia	192
	Poznámky a odkazy na samoštúdium	193
6	Diferenčné rovnice	195
6.1	Základy diferenčného počtu	195
6.2	Diferenčné rovnice a súvisiace pojmy	200
6.3	Systémy lineárnych diferenčných rovníc prvého rádu	204
6.4	Autonómne systémy	206
6.5	Lineárne diferenčné rovnice k -teho rádu	210
	Cvičenia	213
	Poznámky a odkazy na samoštúdium	214
7	Vlastné čísla a grafy II	215
7.1	Neorientované grafy	215
7.2	Spektrálne vlastnosti reálnych symetrických matíc	217
7.3	Rayleighove podiely a Courantova-Fischerova veta	222
7.4	Cauchyho veta o prepletaní vlastných čísel	226
7.5	Spektrálna charakterizácia regulárnych grafov	227
7.6	Kladne semidefinitné matice	230
7.7	Laplaceova matica grafu	232
7.8	Enumerácia kostier grafu	236
7.9	Algebraický stupeň súvislosti grafu	242
7.10	Fiedlerove vektory	244
	Cvičenia	247
	Poznámky a odkazy na samoštúdium	248
	Literatúra	249

Úvod

Táto učebnica vznikla ako vedľajší produkt rovnomenných prednášok pre študentov magisterského štúdia informatiky na FMFI UK v Bratislave. Jej ambíciou je aspoň v krátkosti predstaviť čitateľom niekoľko oblastí matematiky blízkyh informatike, s ktorými sa v súčasnosti nemožno stretnúť v rámci predmetov bežného informatického kurikula – a takisto upozorniť na niektoré dôležité prepojenia medzi rôznymi oblasťami. Matematika najrôznejších druhov je neoddeliteľnou súčasťou výskumu v mnohých oblastiach informatiky – často pritom ide o matematiku, s ktorou sa typický študent informatiky v rámci bakalárskeho štúdia nemá príležitosť zoznámiť. Tento text – spolu s prednáškami pre magisterských študentov, z ktorých vychádza – predstavuje autorov pokus o čiastočné zmiernenie tohto stavu.

Leitmotívom predkladaného textu je súvis grafov s maticami, ktorý sa tu objavuje v niekoľkých podobách – od kombinatorickej interpretácie operácie maticového násobenia nad polokruhmi v reči orientovaných grafov a jej aplikácií pri riešení niektorých algoritmických úloh na grafoch, cez metódy enumerácie sledov založené na lineárnej algebre, až po základy spektrálnej teórie grafov. Do tohto rámca prirodzene zapadajú niektoré ďalšie oblasti, ktorým sa text takisto venuje: teória automatov s váhami nad polokruhmi a formálnych mocninových radov o niekoľkých nekomutujúcich premenných, Perronova-Frobeniova teória nezáporných matic, či teória diferenčných rovníc. Približne v polovici učebnice je navyše zaradená kapitola o lineárnej algebre obsahujúca stručné zhrnutie teórie okolo vlastných čísel, vlastných vektorov a Jordanovho kanonického tvaru matice – túto kapitolu možno v závislosti od potrieb čitateľa považovať buď za opakovanie, alebo za stručný úvod do problematiky. Odkazy na súvisiacu literatúru možno nájsť predovšetkým v stručných zhrnutiach na konci jednotlivých kapitol slúžiacich špeciálne na tento účel.

Za opravy chýb a preklepov ďakujem Adriánovi Gogovi, Martinovi Pašenovi a Kláre Sládečkovej; ďakujem tiež všetkým ostatným študentom, ktorí v minulosti navštevovali prednášky pokryté touto učebnicou, za komentáre a cennú spätnú väzbu. Moja veľká vďaka napokon patrí obom recenzentom za ich nemalé vynaložené úsilie a množstvo príhodných postrehov.

V Bratislave, 24. apríla 2025

Autor

Prehľad označení a konvencií

Nasledujúci stručný prehľad predstavuje výber spomedzi notačných a niektorých ďalších konvencií používaných v tomto texte. Zameriava sa predovšetkým na označenia a konvencie, ktoré vzhľadom na odlišnosti v prístupnej literatúre nemožno považovať za úplne samozrejmé.

Číselné a ďalšie štandardné obory

$\mathbb{B}$	Množina booleovských hodnôt $\mathbb{B} = \{0, 1\}$.
$\mathbb{N}$	Množina všetkých prirodzených čísel <i>vrátane nuly</i> .
$\mathbb{Z}$	Množina všetkých celých čísel.
$\mathbb{Q}$	Množina všetkých racionálnych čísel.
$\mathbb{R}$	Množina všetkých reálnych čísel.
$\mathbb{R}_{\geq 0}$	Množina všetkých nezáporných reálnych čísel.
$\mathbb{R}_{> 0}$	Množina všetkých kladných reálnych čísel.
$[n]$	Množina $[n] = \{1, \dots, n\}$ pre $n \in \mathbb{N}$; pre $n = 0$ ide o prázdnu množinu.
$\mathbb{Z}_n$	Množina všetkých zvyškových tried celých čísel modulo n pre $n \in \mathbb{N} \setminus \{0\}$.
$[a, b], (a, b)$	Uzavretý resp. otvorený interval reálnych čísel.
$[a, b), (a, b]$	Zľava resp. sprava uzavretý interval reálnych čísel.
Δ_{n-1}	Štandardný $(n - 1)$ -simplex pre $n \in \mathbb{N} \setminus \{0\}$.

Nula na nultú. V celom texte pracujeme s konvenciou $0^0 = 1$; operácie umocňovania v algebrách definujeme v súlade s touto konvenciou.

Množiny a množinové operácie

V nasledujúcom označujú X, Y ľubovoľné množiny.

$ X $	Kardinalita množiny X (pri konečných množinách počet prvkov).
$X \setminus Y$	Rozdiel množín X a Y .
2^X	Potenčná množina – t. j. množina všetkých podmnožín – množiny X .
2_{fin}^X	Množina všetkých konečných podmnožín množiny X .
$\binom{X}{k}$	Množina všetkých k -prvkových podmnožín množiny X pre $k \in \mathbb{N}$.
Y^X	Množina všetkých zobrazení z X do Y .
$\subseteq, \supseteq$	Relácia <i>neostrej</i> inklúzie medzi množinami.
$\subsetneq, \supsetneq$	Relácia <i>ostrej</i> inklúzie medzi množinami.
$\not\subseteq, \not\supseteq$	Relácia komplementu neostrej inklúzie: $X \not\subseteq Y$ práve vtedy, keď nie je $X \subseteq Y$.

Zobrazenia

$f: X \rightarrow Y$	Zobrazenie f z množiny X do množiny Y .
$g \circ f$	Zloženie zobrazení $f: X \rightarrow Y$ a $g: Y \rightarrow Z$ – čiže zobrazenie $(g \circ f): X \rightarrow Z$ také, že pre všetky $x \in X$ je $(g \circ f)(x) = g(f(x))$.
$\text{supp}(\varphi)$	Nosič zobrazenia $\varphi: X \rightarrow (Y, 0)$ do množiny s nulou $(Y, 0)$.

Kroneckerova delta

$\delta_{i,j}$	Kroneckerova delta definovaná pre $i, j \in \mathbb{N}$ ako $\delta_{i,j} = \begin{cases} 1 & \text{ak } i = j, \\ 0 & \text{inak.} \end{cases}$
----------------	--

Elementárna teória čísel a kombinatorika

gcd	Najväčší spoločný deliteľ.
$a \equiv b \pmod{n}$	Čísla $a, b \in \mathbb{Z}$ sú kongruentné modulo $n \in \mathbb{N} \setminus \{0\}$ (dávajú rovnaký zvyšok po delení číslom n).
$n!$	Faktoriál čísla $n \in \mathbb{N}$.
α^k	k -ty klesajúci faktoriál čísla $\alpha \in \mathbb{C}$.
$\binom{\alpha}{k}$	Binomický koeficient pre $\alpha \in \mathbb{C}$ a $k \in \mathbb{N}$.

Komplexné čísla

V nasledujúcom je $z \in \mathbb{C}$ ľubovoľné komplexné číslo.

$ z $	Absolútna hodnota komplexného čísla z .
$\text{Re } z$	Reálna zložka komplexného čísla z .
$\text{Im } z$	Imaginárna zložka komplexného čísla z .
$\bar{z}$	Komplexne združené číslo k číslu z .

Postupnosti a systémy

$(a_t)_{t=k}^{\infty}$	Nekonečná postupnosť $(a_k, a_{k+1}, a_{k+2}, \dots)$ pre $k \in \mathbb{Z}$.
$(a_t)_{t \in \mathbb{N}}$	Nekonečná postupnosť $(a_0, a_1, a_2, \dots)$.
$(a_i \mid i \in I)$	Systém prvkov a_i pre všetky $i \in I$ (formálne ide o zobrazenie F s definičným oborom I také, že pre všetky $i \in I$ je $F(i) = a_i$).
$\mathcal{F}(X)$	Trieda všetkých systémov prvkov množiny X .
$\mathcal{F}_{\omega}(X)$	Trieda všetkých konečných alebo spočítateľne nekonečných systémov prvkov množiny X .

Matice, vektory a vektorové priestory

$\dim(\mathcal{V})$	Dimenzia konečnorozmerného vektorového priestoru $\mathcal{V}$.
$\mathcal{S}_1 + \dots + \mathcal{S}_n$	Lineárny súčet podpriestorov $\mathcal{S}_1, \dots, \mathcal{S}_n$ vektorového priestoru $\mathcal{V}$.
$\mathcal{S}_1 \oplus \dots \oplus \mathcal{S}_n$	Priamy súčet podpriestorov $\mathcal{S}_1, \dots, \mathcal{S}_n$ vektorového priestoru $\mathcal{V}$.
$\text{Id}_{\mathcal{V}}$	Identické lineárne zobrazenie $\text{Id}_{\mathcal{V}}: \mathcal{V} \rightarrow \mathcal{V}$ na vektorovom priestore $\mathcal{V}$. V prípade priestoru $\mathcal{V}$ zrejme z kontextu môže byť označované len Id .

$(a_{i,j})_{m \times n}$	Skrátený zápis matice $\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$ typu $m \times n$.
--------------------------	--

$S^{m \times n}$	Množina všetkých matíc typu $m \times n$ nad množinou S .
S^n	Množina všetkých usporiadaných n -tíc prvkov S , ktoré podľa potreby chápeme aj ako prvky $S^{1 \times n}$ alebo $S^{n \times 1}$ – t. j. ako riadkové alebo stĺpcové vektory.
$\mathbf{0}_{m \times n}$	Nulová matica $(0)_{m \times n}$ typu $m \times n$. V prípade rozmerov zrejmých z kontextu môže byť označovaná len $\mathbf{0}$.
$\mathbf{I}_n$	Jednotková matica $\mathbf{I}_n = (\delta_{i,j})_{n \times n}$ typu $n \times n$. V prípade rozmerov zrejmých z kontextu môže byť označovaná len $\mathbf{I}$.
A^T	Transpozícia matice A .
A^{-1}	Inverzná matica k nesingulárnej matici A .
$\mathbf{e}_t$	V kontexte vektorového priestoru $\mathbb{F}^n$ pre pole $\mathbb{F}$ a $n \in \mathbb{N}$ ide o označenie t -teho vektora štandardnej bázy $\mathbb{F}^n$ pre $t \in [n]$, podľa potreby chápaného ako stĺpcový alebo ako riadkový vektor.
$\mathcal{B}_{st}(\mathbb{F}^n)$	Štandardná báza vektorového priestoru $\mathbb{F}^n$ pre pole $\mathbb{F}$ a $n \in \mathbb{N}$.
$A \leq B$	Nerovnosť matíc $A, B \in \mathbb{R}^{m \times n}$ definovaná po zložkách; ak teda $A = (a_{i,j})_{m \times n}$ a $B = (b_{i,j})_{m \times n}$, je $A \leq B$ pokiaľ $a_{i,j} \leq b_{i,j}$ pre všetky $i \in [m]$ a $j \in [n]$.
$ A $	Absolútna hodnota matice $A \in \mathbb{C}^{m \times n}$ definovaná po zložkách; ak je teda $A = (a_{i,j})_{m \times n}$, tak $ A = (a_{i,j})_{m \times n}$.
$\text{Row}(A)$	Riadkový priestor matice A (vektorový priestor generovaný jej riadkami).
$\text{Col}(A)$	Stĺpcový priestor matice A (vektorový priestor generovaný jej stĺpcami).
$\text{Ker}(\varphi), \text{Ker}(A)$	Jadro lineárneho zobrazenia φ resp. jadro (nulový priestor) matice A .
$\text{Im}(\varphi), \text{Im}(A)$	Obraz lineárneho zobrazenia φ resp. obraz matice A (rovný stĺpcovému priestoru $\text{Col}(A)$ matice A). Ak je $Z = (z_{i,j})_{m \times n}$ matica nad $\mathbb{C}$, môže $\text{Im}(Z)$ zriedkavo označovať aj maticu imaginárnych zložiek $(\text{Im } z_{i,j})_{m \times n}$; použitie notácie v tomto význame je vždy zrejmé z kontextu.
$\text{sgn}(\pi)$	Znamienko permutácie $\pi: [n] \rightarrow [n]$.
S_n	Symetrická grupa n -prvkovej množiny pre $n \in \mathbb{N}$: grupa všetkých permutácií $\pi: [n] \rightarrow [n]$ s operáciou skladania.
$\det(A)$	Determinant štvorcovej matice $A = (a_{i,j})_{n \times n}$ nad poľom $\mathbb{F}$ definovaný ako

$$\det(A) = \sum_{\pi \in S_n} \text{sgn}(\pi) \prod_{k=1}^n a_{k, \pi(k)}.$$

Determinant matice typu 0×0 je rovný jednej, pretože S_0 obsahuje jedinú permutáciu $\pi_0: [0] \rightarrow [0]$, pre ktorú je $\text{sgn}(\pi_0) = 1$. Pre determinant matice $A = (a_{i,j})_{0 \times 0}$ tak dostávame

$$\det(A) = \sum_{\pi \in S_0} \text{sgn}(\pi) \prod_{k=1}^0 a_{k, \pi(k)} = \text{sgn}(\pi_0) \prod_{k=1}^0 a_{k, \pi_0(k)} = 1 \cdot 1 = 1.$$

$x[\mathcal{B}]$	Súradnice vektora $\mathbf{x} \in \mathcal{V}$ vzhľadom na bázu $\mathcal{B}$ konečnorozmerného vektorového priestoru $\mathcal{V}$.
φ_A	Lineárne zobrazenie prislúchajúce k matici A , určené násobením <i>stĺpcových</i> vektorov maticou A zľava.
$A_\varphi[\mathcal{B}]$	Matica lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ vzhľadom na bázu $\mathcal{B}$ priestoru $\mathcal{V}$.
A_φ	Matica lineárneho zobrazenia $\varphi: \mathbb{F}^n \rightarrow \mathbb{F}^n$ vzhľadom na štandardnú bázu priestoru $\mathbb{F}^n$ pre pole $\mathbb{F}$ a $n \in \mathbb{N}$.
$P[\mathcal{B}_2 \leftarrow \mathcal{B}_1]$	Matica prechodu od bázy $\mathcal{B}_1$ k báze $\mathcal{B}_2$.

$\text{ch}_A(z), \text{ch}_\varphi(z)$	Charakteristický polynóm matice A resp. lineárneho zobrazenia φ .
$E(\varphi, \lambda), E(A, \lambda)$	Vlastný priestor lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ resp. matice A prislúchajúci k vlastnému číslu λ .
$G(\varphi, \lambda), G(A, \lambda)$	Zovšeobecnený vlastný priestor lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ resp. matice A prislúchajúci k vlastnému číslu λ .
$\sigma(\varphi), \sigma(A)$	Spektrum lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ resp. matice A .
$\alpha(\varphi, \lambda), \alpha(A, \lambda)$	Algebraická násobnosť vlastného čísla λ lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ resp. matice A .
$\gamma(\varphi, \lambda), \gamma(A, \lambda)$	Geometrická násobnosť vlastného čísla λ lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ resp. matice A .
$\varrho(A)$	Spektrálny polomer matice A .
$\text{tr}(A)$	Stopa matice A .
A_p	Sprievodná (pridružená) matica polynómu $p(z)$.
$J_n(\lambda)$	Jordanova bunka rádu $n \in \mathbb{N} \setminus \{0\}$ prislúchajúca k číslu λ .
$\mathcal{V}^\perp$	Ortogonalný doplnok podpriestoru $\mathcal{V}$ vektorového priestoru $\mathbb{R}^n$ pre $n \in \mathbb{N}$.
$\ \mathbf{x}\ _2$	Euklidovská norma (dĺžka) vektora $\mathbf{x} \in \mathbb{R}^n$ pre $n \in \mathbb{N}$.
$\ \mathbf{x}\ _1$	Manhattanská norma vektora $\mathbf{x} \in \mathbb{R}^n$ pre $n \in \mathbb{N}$.
$V(\alpha_1, \dots, \alpha_n)$	Vandermondova matica.
Γ_A	Collatzova-Wielandtova funkcia prislúchajúca k matici A .
$R_A(\mathbf{x})$	Rayleighov podiel pre symetrickú maticu $A \in \mathbb{R}^{n \times n}$ a vektor $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.
$p(A)$	Index primitívnosti ireducibilnej matice $A \in \mathbb{R}_{\geq 0}^{n \times n}$ pre $n \in \mathbb{N} \setminus \{0\}$ s $\varrho(A) > 0$.

Bázy vektorových priestorov. Oproti častej definícii báz konečnorozmerných vektorových priestorov ako špeciálnych *množín* vektorov chápeme bázy ako *konečné systémy* alebo – predovšetkým – ako *usporiadané n -tice* vektorov. Pri pohľade na bázu ako na systém vektorov je napríklad zaručené, že priestor s bázou $(\mathbf{v}_k \mid k \in [n])$ má vždy dimenziu n ; báza chápaná ako množina $\{\mathbf{v}_k \mid k \in [n]\}$ túto vlastnosť vo všeobecnosti mať nemusí. Pohľad na bázy ako na usporiadané n -tice vektorov $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ navyše umožňuje jednoznačnú definíciu súradníc vektorov vzhľadom na uvažovanú bázu.

Diferenčný počet

$\mathcal{F}$	Vektorový priestor $\mathbb{C}^{\mathbb{N}}$ všetkých funkcií z $\mathbb{N}$ do $\mathbb{C}$.
Δ	Operátor konečnej diferencie na $\mathcal{F}$.
E	Operátor posunu na $\mathcal{F}$.
Id	Identický operátor na $\mathcal{F}$.
$\Delta^{-1}f(t), \sum_t f(t)$	Neurčitá suma funkcie $f: \mathbb{N} \rightarrow \mathbb{C}$.
$\mathcal{L}(\mathcal{F})$	Okruh lineárnych operátorov na $\mathcal{F}$.

Grafy

$V(\mathcal{G})$	Množina vrcholov (orientovaného alebo neorientovaného) grafu $\mathcal{G}$.
$E(\mathcal{G})$	Množina hrán (orientovaného alebo neorientovaného) grafu $\mathcal{G}$.
$\deg_{\mathcal{G}}(v)$	Stupeň vrcholu $v \in V(\mathcal{G})$ neorientovaného grafu $\mathcal{G}$.
$\mathbf{A}(\mathcal{G})$	Matica susednosti (orientovaného alebo neorientovaného) grafu $\mathcal{G}$.
$\text{Gr}(A)$	Graf štvorcovej matice A .

$p(\mathcal{G})$	Periódá silne súvislého ohodnoteného orientovaného grafu nad $\mathbb{R}_{\geq 0}$.
$L(\mathcal{G})$	Laplaceova matica neorientovaného grafu $\mathcal{G}$.
$B(\mathcal{G})$	Incidenčná matica orientovaného grafu bez slučiek $\mathcal{G}$.
$Q(\mathcal{G})$	Kladná Laplaceova matica neorientovaného grafu $\mathcal{G}$.

Formálne mocninové rady a polynómy

$S[[z]]$	Množina všetkých formálnych mocninových radov o jednej premennej z s koeficientmi v množine S .
$S[z]$	Množina všetkých polynómov o jednej premennej z s koeficientmi v S .
$[z^t]r(z)$	Koeficient pri z^t pre $t \in \mathbb{N}$ vo formálnom mocninovom rade $r(z)$.
$\text{codeg } r(z)$	Kostupeň formálneho mocninového radu $r(z)$.
$D^{(k)}$	Hasseho derivácia rádu $k \in \mathbb{N}$.
$S[[z_1, \dots, z_m]]$	Množina všetkých formálnych mocninových radov o $m \in \mathbb{N} \setminus \{0\}$ komutujúcich premenných $z_1, \dots, z_m$ s koeficientmi v množine S .
$S\langle\langle \Sigma^* \rangle\rangle$	Množina všetkých formálnych mocninových radov o niekoľkých nekomutujúcich premenných z abecedy Σ s koeficientmi v množine S .
$S\langle \Sigma^* \rangle$	Množina všetkých polynómov o niekoľkých nekomutujúcich premenných z abecedy Σ s koeficientmi v množine S .
(r, w)	Koeficient slova $w \in \Sigma^*$ vo formálnom mocninovom rade $r \in S\langle\langle \Sigma^* \rangle\rangle$.
$\text{supp}(r)$	Nosič formálneho mocninového radu $r \in S\langle\langle \Sigma^* \rangle\rangle$.
$S\langle\langle L \rangle\rangle$	Pre $L \subseteq \Sigma^*$ množina všetkých formálnych mocninových radov $r \in S\langle\langle \Sigma^* \rangle\rangle$ takých, že $\text{supp}(r) \subseteq L$.
$S\langle L \rangle$	Pre $L \subseteq \Sigma^*$ množina všetkých polynómov $r \in S\langle \Sigma^* \rangle$ takých, že $\text{supp}(r) \subseteq L$.
$x^{-1}r$	Ľavý kvocient formálneho mocninového radu $r \in S\langle\langle \Sigma^* \rangle\rangle$ podľa slova $x \in \Sigma^*$.
$\mathcal{Q}(r)$	Priestor kvocientov radu $r \in \mathbb{F}\langle\langle \Sigma^* \rangle\rangle$ nad poľom $\mathbb{F}$ a abecedou Σ .

Automaty s váhami a racionálne rady

$\ \mathcal{A}\ $	Formálny mocninový rad realizovaný konečným automatom s váhami $\mathcal{A}$.
$\mathcal{R}(\mathcal{A})$	Množina všetkých behov automatu $\mathcal{A}$.
$S\text{-RatE}(\Sigma^*)$	Množina všetkých racionálnych výrazov s váhami nad polokruhom S a abecedou Σ .
$\ \mathbf{E}\ $	Formálny mocninový rad realizovaný racionálnym výrazom s váhami $\mathbf{E}$.
$S\text{-Rat}(\Sigma^*)$	Množina všetkých radov $r \in S\langle\langle \Sigma^* \rangle\rangle$ racionálnych nad polokruhom S .
$\mathcal{P}_{\mathcal{A}}$	Lineárna reprezentácia zodpovedajúca automatu $\mathcal{A}$.

Kapitola 1

Grafy, matice a polokruhy

Orientované *grafy* s hranami ohodnotenými prvkami nejakej množiny S a štvorcové *matice* nad S sú, ako čoskoro uvidíme, dvoma stranami tej istej mince – ide o dva rôzne pohľady na ten istý objekt. V rámci tejto kapitoly preskúmame niektoré z dôsledkov tohto prepojenia.

Ukážeme napríklad, že už relatívne slabá algebraická štruktúra na množine S – túto štruktúru nazveme *polokruhom* – umožňuje definovať násobenie matíc nad S , ktoré možno často interpretovať ako riešenie prirodzenej úlohy na ohodnotených grafoch. Za určitých ďalších predpokladov o algebraickej štruktúre na S – nimi dospejeme k pojmu *úplného polokruhu* – možno definovať aj iteráciu matice, opäť s prirodzenou interpretáciou z perspektívy teórie ohodnotených grafov.

Pozorovanie súvisu medzi úlohami na ohodnotených grafoch a operáciami na maticiach má dvojaký význam. Na jednej strane možno maticové násobenie a algoritmus na výpočet iterácie matice využiť ako základ pre grafové algoritmy riešiace rôzne prakticky užitočné problémy. Grafová interpretácia maticových operácií na strane druhej značne napomáha ich intuitívnemu pochopeniu a metódy teórie grafov sa tak stávajú užitočným nástrojom pre štúdium niektorých aspektov maticového počtu.

V súvislosti so skúmaním ohodnotených grafov nad úplnými polokruhmi sa tiež prirodzene ponúka možnosť aplikovať niektoré myšlienky známe z teórie konečných automatov vo všeobecnejšom kontexte; to je nevelmi prekvapivým dôsledkom pozorovania, že konečné automaty v princípe zodpovedajú ohodnoteným grafom nad jedným konkrétnym úplným polokruhom spolu s trochou dodatočnej informácie navyše. Význam úvah tohto typu plne doceníme až v nasledujúcich kapitolách, kde napríklad ukážeme, že jazyky rozpoznávané konečnými automatmi a racionálne funkcie sú v skutočnosti rovnakými objektmi nad rôznymi polokruhmi.¹

Už v tejto kapitole ale napríklad uvidíme, že problém hľadania cien najlacnejších sledov v grafe a problém konštrukcie racionálneho výrazu ekvivalentného danému konečnému automatu sú v skutočnosti takmer rovnakými problémami nad rôznymi polokruhmi. A okrem iného tiež uvidíme, že dobre známe algoritmy na riešenie týchto problémov – konkrétne Floydov-Warshallov algoritmus hľadania cien najlacnejších ciest v grafe a McNaughtonov-Jamadov algoritmus konštrukcie racionálneho výrazu ku konečnému automatu² – možno chápať ako špeciálne prípady jediného všeobecného algoritmu na výpočet iterácie matice nad úplným polokruhom.

V tejto kapitole tak načrieme do oblastí na pomedzí matematiky a informatiky, v ktorých sú predmetom skúmania vzájomné prepojenia medzi teóriou ohodnotených grafov, teóriou matíc, teóriou automatov a jazykov a do určitej miery aj algebrou. V nasledujúcej kapitole ešte k týmto štyrom oblastiam pridáme piatu: kombinatoriku. Čitateľ sa snáď sám presvedčí o skutočnosti, že skúmať ktorúkoľvek z týchto oblastí oddelene od ostatných by znamenalo značné ochudobnenie.

¹V tomto texte preto upustíme od zaužívaných, avšak nevelmi výstižných, termínov *regulárny výraz* a *regulárny jazyk* – náhodnosť prívlastku „regulárny“ si v skutočnosti uvedomoval už sám jeho pôvodca S. C. Kleene [64, str. 46]. Namiesto toho budeme spoločne s mnohými matematickejšie orientovanými autormi [10, 11, 33, 94] hovoriť o *racionálnych výrazoch* a o *racionálnych jazykoch*.

²Pravdepodobne najznámejší algoritmus pre túto úlohu, ktorý v literatúre vystupuje pod rôznymi názvami. Pomenovanie „McNaughtonov-Jamadov algoritmus“, asi najbližšie historickej skutočnosti [83], preberáme od Sakarovitcha [94].

1.1 Matica susednosti ohodnoteného grafu

Pri obvyklej definícii *orientovaného grafu* – v našom ponímaní vždy *bez násobných hrán* medzi vrcholmi, avšak *s možnosťou existencie orientovaných slučiek* – sa hrany stotožňujú s usporiadanými dvojicami vrcholov. Naša definícia orientovaného grafu bude mierne netradičnou formalizáciou tohto tradičného prístupu. Namiesto množiny hrán bude pre každý graf dané zobrazenie φ priradujúce každej dvojici vrcholov prvok množiny booleovských hodnôt $\mathbb{B} = \{0, 1\}$. Z vrcholu u do vrcholu v potom vedie orientovaná hrana práve vtedy, keď $\varphi(u, v) = 1$. Táto formálna okľuka nie je samoúčelná; zakrátko uvidíme, že umožňuje priamočiare zovšeobecnenie definície na prípad ohodnotených grafov.

Definícia 1.1.1. *Orientovaný graf* je dvojica $\mathcal{G} = (V, \varphi)$, kde V je konečná množina a $\varphi: V^2 \rightarrow \mathbb{B}$ je zobrazenie.

Pre množiny vrcholov a hrán orientovaného grafu $\mathcal{G}$ budeme používať štandardné označenie $V(\mathcal{G})$ resp. $E(\mathcal{G})$. Definíciu množiny hrán $E(\mathcal{G})$ pritom sformulujeme s využitím dvoch pomocných pojmov. Pod *množinou s nulou* budeme rozumieť usporiadanú dvojicu $(Y, 0)$, kde Y je množina a 0 je prvok množiny Y . Pre každú množinu X , každú množinu s nulou $(Y, 0)$ a každé zobrazenie $\varphi: X \rightarrow Y$ nazveme *nosičom* zobrazenia φ množinu $\text{supp}(\varphi) = \{x \in X \mid \varphi(x) \neq 0\}$.³ Množinu $\mathbb{B}$ budeme interpretovať aj ako množinu s nulou $(\mathbb{B}, 0)$.

Definícia 1.1.2. Nech $\mathcal{G} = (V, \varphi)$ je orientovaný graf. Potom píšeme $V(\mathcal{G}) := V$ a $E(\mathcal{G}) := \text{supp}(\varphi)$. Prvky množiny $V(\mathcal{G})$ nazývame *vrcholy* grafu $\mathcal{G}$ a prvky množiny $E(\mathcal{G})$ nazývame *hrany* grafu $\mathcal{G}$.

Orientovaný graf je teda daný množinou vrcholov V a funkciou φ , ktorá pre každú dvojicu vrcholov $u, v \in V$ vráti hodnotu z množiny $\mathbb{B}$. Ak $\varphi(u, v) = 0$, z vrcholu u do vrcholu v nevedie žiadna orientovaná hrana; ak $\varphi(u, v) = 1$, vedie z u do v práve jedna orientovaná hrana (u, v) .

Ohodnotený graf je v intuitívnom ponímaní orientovaný graf, v ktorom každej hrane zodpovedá ohodnotenie z nejakej množiny H . Budeme navyše predpokladať, že H je množina s nulou 0 – v takom prípade totiž môžeme ohodnotenie nulou chápať ako známku neexistencie hrany. Definíciu ohodnoteného grafu tak získame jednoduchým nahradením zobrazenia $\varphi: V^2 \rightarrow \mathbb{B}$ v definícii orientovaného grafu zobrazením $\varphi: V^2 \rightarrow H$. Ak pre dvojicu vrcholov u, v platí $\varphi(u, v) = 0$, nevedie z u do v žiadna hrana; ak $\varphi(u, v) = h$ pre nejaké $h \in H \setminus \{0\}$, vedie z u do v hrana (u, v) ohodnotená prvkom h .

Definícia 1.1.3. Nech $(H, 0)$ je množina s nulou. *Ohodnotený graf* nad $(H, 0)$ je dvojica $\mathcal{G} = (V, \varphi)$, kde V je konečná množina a $\varphi: V^2 \rightarrow H$ je zobrazenie. Prvky množiny $V := V(\mathcal{G})$ nazývame *vrcholy* grafu $\mathcal{G}$ a prvky $\text{supp}(\varphi) := E(\mathcal{G})$ nazývame *hrany* grafu $\mathcal{G}$.

Vďaka evidentnej korešpondencii našich definícií orientovaných a ohodnotených grafov s ich obvyklými definíciami si v nasledujúcom môžeme dovoliť používať bežnú grafovú terminológiu.

Pripomeňme si teraz definíciu *matice*. Predovšetkým je nutné zdôrazniť, že maticu je v princípe možné definovať nad *ľubovoľnou množinou* S – ide jednoducho o „obdĺžnikovú tabuľku“ pozostávajúcu z prvkov S . Nutnosť požadovať algebraickú štruktúru na S vyvstane až v súvislosti s niektorými operáciami na maticiach. Uvidíme napríklad, že na „rozumnú“ definíciu sčítovania, násobenia a umocňovania matíc nad S je nutné a postačujúce, aby množina S tvorila s nejakými dvoma operáciami polokruh (oddiel 1.3). Podobne inverzné matice možno „rozumne“ skúmať kedykoľvek je S pole.⁴ Na zavedenie

³Pojem *nosiča* funkcie (angl. *support*) je v matematike štandardne používaný, zvyčajne však v špeciálnejších kontextoch, než sú zobrazenia do množín s nulou. Zanedlho k nim koniec koncov prejdeme aj my.

⁴V úvodnej literatúre o lineárnej algebre sa vďaka tejto a podobným vlastnostiam matíc nad poľom už samotný pojem matice nad S občas definuje len za predpokladu, že S je pole. To môže zvädzať k predstave o inherentnej spätosti matíc a polí, ktorá je však mylná. Takáto nešťastná interpretácia by nás – ako uvidíme neskôr v tejto kapitole – obrala nielen o pochopenie skutočného významu pojmu matice, ale aj o množstvo praktických aplikácií maticového počtu nad algebrami všeobecnejšími, než sú polia. (Samotnú lineárnu algebru navyše tiež možno – za cenu určitých strát – vybudovať aj o čosi všeobecnejšie.)

samotného pojmu matice ale žiadna takáto štruktúra potrebná nie je. Čitateľa, ktorý istotne nejdennokrát napísal program využívajúci dvojrozmerné polia, snáď o užitočnosti matíc nad množinami dlhšie presvedčať netreba. Navyše existujú aj *operácie* na maticiach, ktoré možno definovať už pre matice nad množinou – príkladom môže byť napríklad transpozícia.

Definícia 1.1.4. Nech S je množina a $m, n \in \mathbb{N}$. *Matica* typu $m \times n$ nad S je zobrazenie

$$A: [m] \times [n] \rightarrow S.$$

Ak pre $i = 1, \dots, m$ a $j = 1, \dots, n$ je $A(i, j) = a_{i,j}$, zapisujeme maticu A ako

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix},$$

alebo skrátené ako $A = (a_{i,j})_{m \times n}$. Matica A je *štvorcová*, ak $m = n$. Množinu všetkých matíc typu $m \times n$ nad S označujeme $S^{m \times n}$.

Definícia 1.1.5. Nech S je množina, m, n sú prirodzené čísla a $A = (a_{i,j})_{m \times n}$ je matica typu $m \times n$ nad S . *Transpozícia* matice A je matica A^T typu $n \times m$ nad S daná ako $A^T = (a_{j,i})_{n \times m}$.

Nech $(H, 0)$ je množina s nulou a n je prirodzené číslo. Z vyššie uvedeného je už potom zrejmé, že *ohodnotené grafy nad H s množinou vrcholov $[n]$ a štvorcové matice typu $n \times n$ nad H sú dvoma odlišnými pohľadmi na ten istý objekt*. (Je pritom zrejmé, že prípadným premenovaním vrcholov grafu na prirodzené čísla $1, \dots, n$ nestrácame žiadnu podstatnú informáciu.)

Toto takmer až triviálne pozorovanie vedie k nasledujúcim dvom definíciám – ak sa na graf pozrieme ako na maticu, budeme hovoriť o *matici susednosti* grafu; ak sa naopak pozrieme na štvorcovú maticu ako na graf, budeme hovoriť o *grafe matice*.

Definícia 1.1.6. Nech $(H, 0)$ je množina s nulou, $n \in \mathbb{N}$ a $\mathcal{G} = ([n], \varphi)$ je ohodnotený graf nad $(H, 0)$. *Maticou susednosti* grafu $\mathcal{G}$ nazveme štvorcovú maticu $\mathbf{A}(\mathcal{G}) \in H^{n \times n}$ danú ako $\mathbf{A}(\mathcal{G}) = (\varphi(i, j))_{n \times n}$.

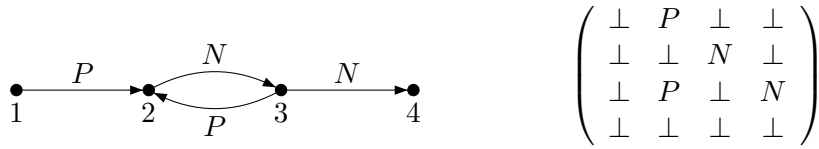
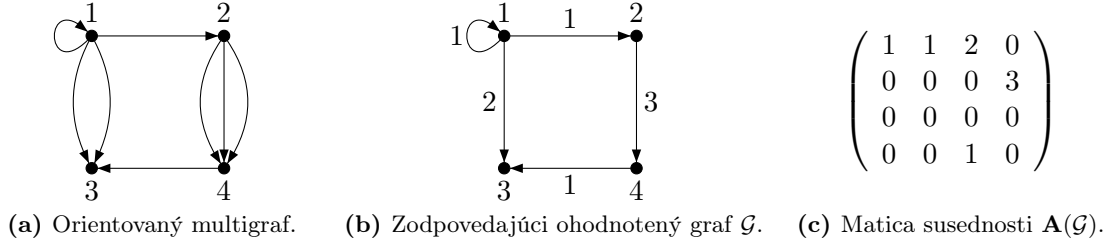
Definícia 1.1.7. Nech $(H, 0)$ je množina s nulou, $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad $(H, 0)$. *Grafom matice* A nazveme ohodnotený graf $\text{Gr}(A)$ nad množinou $(H, 0)$ daný ako $\text{Gr}(A) = ([n], \varphi)$, kde pre $i, j = 1, \dots, n$ je $\varphi(i, j) = a_{i,j}$.

Je očividné, že operátory $\mathbf{A}$ a Gr zavedené v predošlých dvoch definíciách sú – pre pevne danú množinu s nulou $(H, 0)$ a prirodzené číslo n – vzájomne inverzné bijekcie.

Práve opísaný súvis medzi grafmi a maticami má ďalekosiahle implikácie. So zlomkom tých najvýznamnejších sa postupne zoznámime v tejto a aj v nasledujúcich kapitolách; niektoré ďalšie budú stručne spomenuté v rámci odkazov na samoštúdium.

Príklad 1.1.8. Uvažujme orientovaný graf zjednodušenej cestnej siete, kde vrcholy zodpovedajú križovatkám a hrana z u do v zodpovedá cestnému úseku *priamo* spájajúcemu križovátku u s križovatkou v . Predpokladajme, že do takéhoto grafu potrebujeme zakomponovať informáciu o výlukách na niektorých úsekoch, a to bez modifikácie samotnej množiny hrán. To možno urobiť napríklad s použitím ohodnotených grafov nad množinou s nulou $(\{P, N, \perp\}, \perp)$, kde P zodpovedá prejazdnému úseku, N zodpovedá neprejazdnému úseku a $\perp$ zodpovedá neexistujúcej hrane. Príklad ohodnoteného grafu nad takouto množinou a jeho matice susednosti je na obrázku 1.1. Zdôraznime pritom, že na množine $\{P, N, \perp\}$ nie sú *a priori* dané žiadne operácie, a teda ani žiadna algebraická štruktúra.

Príklad 1.1.9. Každý orientovaný *multigraf* – čiže orientovaný graf s možnosťou existencie násobných hrán medzi vrcholmi – možno reprezentovať ako ohodnotený graf nad množinou $\mathbb{N}$ so štandardne danou nulou. Ak v pôvodnom multigrafe vedie z vrcholu u do vrcholu v práve k orientovaných hrán, v príslušnom ohodnotenom grafe vedie z vrcholu u do vrcholu v orientovaná hrana ohodnotená prirodzeným číslom k . Príklad orientovaného multigrafu, k nemu prislúchajúceho ohodnoteného grafu a jeho matice susednosti je znázornený na obrázku 1.2.

(a) Ohodnotený graf $\mathcal{G}$.(b) Matica susednosti $\mathbf{A}(\mathcal{G})$.**Obr. 1.1:** Ohodnotený graf nad množinou $\{P, N, \perp\}$ s nulou $\perp$ a jeho matica susednosti.

(a) Orientovaný multigraf.

(b) Zodpovedajúci ohodnotený graf $\mathcal{G}$.(c) Matica susednosti $\mathbf{A}(\mathcal{G})$.**Obr. 1.2:** Multigraf, jeho reprezentácia pomocou ohodnoteného grafu $\mathcal{G}$ nad $\mathbb{N}$ a matica susednosti $\mathbf{A}(\mathcal{G})$.

1.2 Násobenie matíc a niektoré úlohy na grafoch

Prostredníctvom pojmu matice susednosti grafu a „inverzného“ pojmu grafu matice sme poukázali na súvis medzi ohodnotenými grafmi a štvorcovými maticami. Jediným predpokladom o univerze možných ohodnotení hrán (či prvkov matice) tu bola existencia špeciálneho prvku – ten sme nazvali nulou – zodpovedajúceho neexistujúcim hranám.

K tomuto predpokladu teraz pridáme algebraickú štruktúru: preskúmame niekoľko konkrétnych príkladov množín, na ktorých sú definované operácie sčítania a násobenia, $\oplus$ a $\otimes$, a ktoré obsahujú neutrálne prvky 0 resp. 1 vzhľadom na tieto operácie. To nám spravidla umožní definovať *maticové násobenie* nad takto danou algebrou a občas aj *iteráciu matice* – čiže maticu $A^* = A^0 \oplus A^1 \oplus A^2 \oplus \dots$ – pričom uvidíme, že obidve tieto operácie možno často využiť na riešenie pomerne prirodzených a užitočných úloh na grafoch. Viaceré zdanlivo nesúvisiace problémy sa tak v skutočnosti ukážu byť rovnakými úlohami nad rôznymi algebrami.⁵

Nateraz sa teda obmedzíme na niekoľko konkrétnych algebr typu $(S, \oplus, \otimes, 0, 1)$, ktoré budeme zakaždým chápať aj ako množinu s nulou $(S, 0)$. Dôraz pritom budeme klásť na aplikácie násobenia matíc nad týmito algebrami pri riešení grafových úloh. Všeobecne platnými podmienkami, ktoré musí algebra $(S, \oplus, \otimes, 0, 1)$ za účelom definície „rozumného“ maticového násobenia spĺňať, sa zatiaľ zaoberať nebudeme – ich identifikácia je náplňou oddielu 1.3, kde zdefinujeme spoločnú abstrakciu konkrétnych algebr z príkladov nižšie. Takýto postup koniec koncov zohľadňuje zákonitosti typického procesu matematickej abstrakcie: definície abstraktných matematických štruktúr zvyčajne nie sú produktom svojvôle matematikov, ale bývajú spravidla motivované potrebou jednotného skúmania väčšieho množstva konkrétnych prípadov. Najlepšou cestou k pochopeniu významu definície abstraktnej štruktúry je teda získanie citu pre jednotlivé konkrétne prípady a porozumenia tomu, ktoré vlastnosti sú v týchto konkrétnych prípadoch pre dané účely kľúčové. Samotná abstraktná definícia by už následne mala byť len málo prekvapivým „destilátom“ týchto kľúčových vlastností.

Príklad 1.2.1. Orientované grafy, chápané v zmysle definície 1.1.1, zodpovedajú ohodnoteným grafom nad množinou booleovských hodnôt $\mathbb{B} = \{0, 1\}$; možno ich teda reprezentovať booleovskou maticou. Uvažujme teraz na množine $\mathbb{B}$ ako aditívnu operáciu logickú disjunkciu $\vee$ a ako multiplikatívnu operáciu logickú konjunkciu $\wedge$. Aditívnym neutrálnym prvkom je potom 0 a multiplikatívnym neutrálnym prvkom je 1. Vo výsledku teda dostávame algebru $(\mathbb{B}, \vee, \wedge, 0, 1)$.

⁵Pod *algebrou* rozumieme ľubovoľnú štruktúru danú nosnou množinou S a systémom operácií nejakej arity na S ; nulárne operácie pritom zodpovedajú konštantám z S .

Násobenie matíc nad $(\mathbb{B}, \vee, \wedge, 0, 1)$ možno definovať prirodzene: ak $A = (a_{i,j})_{m \times n}$ a $B = (b_{i,j})_{n \times r}$ sú matice nad $\mathbb{B}$, je ich súčin $A \cdot B = (c_{i,j})_{m \times r}$ daný pre $i = 1, \dots, m$ a $j = 1, \dots, r$ predpisom

$$c_{i,j} = \bigvee_{k=1}^n a_{i,k} \wedge b_{k,j}. \quad (1.1)$$

Mocninu štvorcovej matice A typu $n \times n$ potom možno definovať rekurzívne: $A^0 = \mathbf{I}_n$ a $A^{t+1} = A^t \cdot A$ pre všetky $t \in \mathbb{N}$, kde $\mathbf{I}_n$ je jednotková matica typu $n \times n$ daná ako

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Čitateľ iste ľahko overí, že pre všetky $B \in \mathbb{B}^{n \times n}$ je $\mathbf{I}_n \cdot B = B \cdot \mathbf{I}_n = B$.

Nech teraz $A = \mathbf{A}(\mathcal{G}) = (a_{i,j})_{n \times n}$ je matica susednosti orientovaného grafu $\mathcal{G}$ na množine vrcholov $V(\mathcal{G}) = [n]$, nech $t \in \mathbb{N}$ a $A^t = (a_{i,j}[t])_{n \times n}$. Pokúsme sa nájsť grafovú interpretáciu prvkov matice A^t .

Keďže pre $i, j = 1, \dots, n$ je $a_{i,j}[1] = a_{i,j}$, z definície matice susednosti pre $t = 1$ bezprostredne vyplýva, že $a_{i,j}[1] = 1$ práve vtedy, keď z vrcholu i do vrcholu j vedie v grafe $\mathcal{G}$ hrana. Zamerajme sa teraz na prípad $t = 2$. Tu vďaka vzťahu $A^2 = A^{1+1} = A^1 \cdot A$ podľa (1.1) dostávame

$$a_{i,j}[2] = \bigvee_{k=1}^n a_{i,k}[1] \wedge a_{k,j} = \bigvee_{k=1}^n a_{i,k} \wedge a_{k,j}.$$

Vidíme teda, že $a_{i,j}[2] = 1$ práve vtedy, keď existuje vrchol k taký, že v grafe $\mathcal{G}$ vedie hrana ako z i do k , tak aj z k do j ; čiže práve vtedy, keď v grafe $\mathcal{G}$ z vrcholu i do vrcholu j vedie orientovaný sled dĺžky práve 2. Podobne pre $t = 3$ prichádzame k vzťahu

$$a_{i,j}[3] = \bigvee_{k=1}^n a_{i,k}[2] \wedge a_{k,j},$$

vďaka ktorému $a_{i,j}[3] = 1$ práve vtedy, keď existuje vrchol k taký, že v grafe $\mathcal{G}$ vedie orientovaný sled z i do k dĺžky práve 2 a súčasne hrana z k do j ; to nastane práve vtedy, keď z i do j vedie orientovaný sled dĺžky práve 3.

Ak sa teraz vrátíme k prípadu $t = 1$, zisťujeme, že podmienka existencie hrany je ekvivalentná existencii sledu dĺžky práve 1 – preto $a_{i,j}[1] = 1$ práve vtedy, keď z i do j vedie v grafe $\mathcal{G}$ orientovaný sled dĺžky práve 1. V prípade $t = 0$ napokon v dôsledku vzťahu $A^0 = \mathbf{I}_n$ vidíme, že $a_{i,j}[0] = 1$ práve vtedy, keď $i = j$, čiže keď v grafe $\mathcal{G}$ vedie z i do j orientovaný sled dĺžky 0.

Nie je teda ťažké prísť k nasledujúcej hypotéze: *pre všetky $t \in \mathbb{N}$ a $i, j \in V(\mathcal{G})$ je $a_{i,j}[t] = 1$ práve vtedy, keď z vrcholu i do vrcholu j vedie v grafe $\mathcal{G}$ orientovaný sled dĺžky práve t* . Uvedme pre úplnosť aj dôkaz matematickou indukciou vzhľadom na t . Ako sme pozorovali vyššie, tvrdenie platí pre $t = 0$. Predpokladajme teraz jeho platnosť pre $t = s$ a uvažujme $t = s + 1$. Platí $A^{s+1} = A^s \cdot A$, pričom podľa indukčného predpokladu je $a_{i,k}[s] = 1$ práve vtedy, keď v $\mathcal{G}$ existuje orientovaný sled z i do k dĺžky s . Podľa (1.1) ďalej

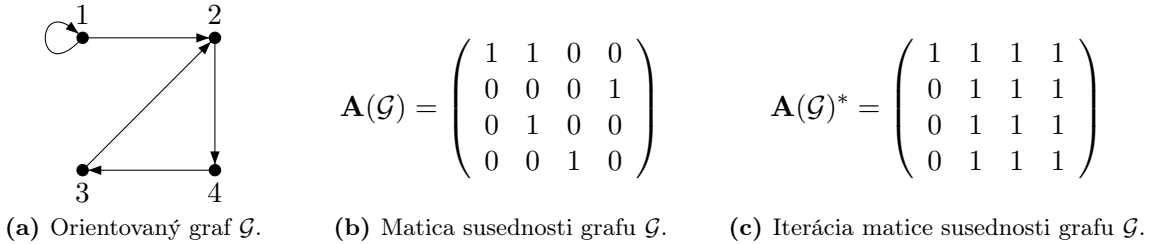
$$a_{i,j}[s+1] = \bigvee_{k=1}^n a_{i,k}[s] \wedge a_{k,j}.$$

Preto $a_{i,j}[s+1] = 1$ práve vtedy, keď existuje vrchol $k \in V(\mathcal{G})$ taký, že z i do k vedie orientovaný sled dĺžky s a z k do j vedie hrana. Ak ale takéto k existuje, spojením uvedených dvoch sledov získame sled dĺžky $s+1$ vedúci z i do j ; ak naopak existuje sled dĺžky $s+1$ vedúci z i do j , jeho predposledný vrchol je hľadané k . To dokazuje naše tvrdenie.

Pre ľubovoľnú štvorcovú maticu A nad $\mathbb{B}$ ďalej možno definovať jej iteráciu ako

$$A^* = \bigvee_{t=0}^{\infty} A^t,$$

kde disjunkciu matíc chápeme po zložkách. Ak je $A = \mathbf{A}(\mathcal{G})$ matica susednosti orientovaného grafu $\mathcal{G}$ a $A^* = (a_{i,j}[*])_{n \times n}$, je $a_{i,j}[*] = 1$ práve vtedy, keď v $\mathcal{G}$ existuje orientovaný sled z vrcholu i do vrcholu j – čiže ak je vrchol j dosiahnuteľný z vrcholu i . Matica A^* tak reprezentuje reflexívno-tranzitívny uzáver grafu $\mathcal{G}$, z ktorého možno ľahko vyčítať aj jednotlivé silne súvislé komponenty tohto grafu. Na obrázku 1.3 je znázornený orientovaný graf $\mathcal{G}$ spoločne s maticami $\mathbf{A}(\mathcal{G})$ a $\mathbf{A}(\mathcal{G})^*$.



Obr. 1.3: Orientovaný graf $\mathcal{G}$, jeho matica susednosti $\mathbf{A}(\mathcal{G})$ a iterácia $\mathbf{A}(\mathcal{G})^*$ matice $\mathbf{A}(\mathcal{G})$.

V nasledujúcich príkladoch už tvrdenia o „grafovej interpretácii“ t -tej mocniny matice susednosti dokazovať nebudeme. Vždy tak ale možno urobiť – rovnako ako v príklade 1.2.1 – jednoduchou matematickou indukciou vzhľadom na t . Čitateľovi – obzvlášť takému, ktorému nižšie uvedené tvrdenia nebudú úplne zrejmé – môže táto drobná neporiadnosť poslúžiť ako príležitosť na užitočné cvičenie (návodom pritom môže byť argumentácia v príklade 1.2.1).

Príklad 1.2.2. *Nedeterministický konečný automat* nad abecedou Σ (a s prechodmi na prázdne slovo) možno definovať ako štvoricu $\mathcal{A} = (Q, I, T, F)$, kde Q je konečná množina stavov, $I \subseteq Q$ je množina počiatočných stavov, $T \subseteq Q \times (\Sigma \cup \{\varepsilon\}) \times Q$ je množina prechodov a $F \subseteq Q$ je množina koncových stavov. Pre dvojicu stavov $p, q \in Q$ a slovo $w \in \Sigma^*$ potom píšeme $p \xrightarrow{w}_{\mathcal{A}} q$ práve vtedy, keď existuje $t \in \mathbb{N}$, stavy $q_0, \dots, q_t \in Q$ a slová $z_1, \dots, z_t \in \Sigma \cup \{\varepsilon\}$ také, že $w = z_1 \dots z_t$, pričom $q_0 = p$, $q_t = q$ a pre $i = 1, \dots, t$ je $(q_{i-1}, z_i, q_i) \in T$. Jazyk $\|\mathcal{A}\|$ rozoznávaný automatom $\mathcal{A}$ napokon definujeme ako $\|\mathcal{A}\| = \{w \in \Sigma^* \mid \exists p \in I \exists q \in F : p \xrightarrow{w}_{\mathcal{A}} q\}$.

Čitateľovi je iste známy pojem *diagramu* konečného automatu. Ide v podstate o ohodnotený graf nad $\Sigma \cup \{\varepsilon\}$ s množinou vrcholov Q ; môže ale obsahovať násobné hrany a niektoré jeho vrcholy môžu byť označené ako zodpovedajúce počiatočným alebo koncovým stavom (šípkou dnu alebo von zo stavu). Potrebu uvažovať násobné hrany možno eliminovať nahradením viacerých hrán vedúcich z p do q jedinou hranou, ktorá je ohodnotená *jazykom* všetkých z takých, že $(p, z, q) \in T$. Navyše možno bez ujmy na všeobecnosti predpokladať, že $Q = [n]$ pre nejaké $n \in \mathbb{N}$.

Nedeterministický konečný automat nad Σ teda možno zadať pomocou prirodzeného čísla n , dvojice množín $I, F \subseteq [n]$ a ohodnoteného grafu $\mathcal{G} = ([n], \varphi)$ nad množinou 2^{Σ^*} s nulou $\emptyset$.⁶ Graf $\mathcal{G}$ navyše možno zadať maticou nad 2^{Σ^*} .

Uvažujme teraz na 2^{Σ^*} ako aditívnu operáciu zjednotenie a ako multiplikatívnu operáciu zretáženie; aditívnym neutrálnym prvkom je v takom prípade $\emptyset$ a multiplikatívnym neutrálnym prvkom je $\{\varepsilon\}$. Pohybujeme sa teda nad algebrou $(2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\})$.

Maticové násobenie nad $(2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\})$ možno prirodzene definovať tak, že pre $A = (A_{i,j})_{m \times n}$ a $B = (B_{i,j})_{n \times r}$ je matica $A \cdot B = (C_{i,j})_{m \times r}$ daná predpisom

$$C_{i,j} = \bigcup_{k=1}^n A_{i,k} \cdot B_{k,j}$$

⁶Ohodnotenia hrán sú v skutočnosti prvkami množiny $2^{\Sigma \cup \{\varepsilon\}}$; širšie univerzum 2^{Σ^*} sa nám však zide pri našich neskorších úvahách (najmä kvôli uzavretosti na zretáženie).

pre $i = 1, \dots, m$ a $j = 1, \dots, r$. Mocninu štvorcovej matice A typu $n \times n$ definujeme rekurzívne: $A^0 = \mathbf{I}_n$ a $A^{t+1} = A^t \cdot A$ pre všetky $t \in \mathbb{N}$, kde $\mathbf{I}_n$ je jednotková matica typu $n \times n$ daná ako

$$\mathbf{I}_n = \begin{pmatrix} \{\varepsilon\} & \emptyset & \dots & \emptyset \\ \emptyset & \{\varepsilon\} & \dots & \emptyset \\ \vdots & \vdots & \ddots & \vdots \\ \emptyset & \emptyset & \dots & \{\varepsilon\} \end{pmatrix}.$$

Vezmime teraz konečný automat $\mathcal{A}$ daný prirodzeným číslom n , množinami $I, F \subseteq [n]$ a ohodnoteným grafom $\mathcal{G} = ([n], \varphi)$ nad 2^{Σ^*} . Nech $A = \mathbf{A}(\mathcal{G}) = (A_{i,j})_{n \times n}$ je matica susednosti grafu $\mathcal{G}$ – budeme hovoriť, že A je *prechodová matica* automatu $\mathcal{A}$. Nech ďalej t je prirodzené číslo, $A^t = (A_{i,j}[t])_{n \times n}$ a $i, j \in [n]$. Ľahko potom možno nahliadnuť, že *jazyk* $A_{i,j}[t]$ *pozostáva z práve všetkých slov* $w \in \Sigma^*$, *ktoré automat* $\mathcal{A}$ *prečíta počas nejakého behu dĺžky* t *zo stavu* i *do stavu* j .⁷

Ak navyše definujeme iteráciu $A^* = (A_{i,j}[*])_{n \times n}$ matice A ako

$$A^* = \bigcup_{t=0}^{\infty} A^t,$$

vyplýva z uvedeného, že jazyk $A_{i,j}[*]$ pozostáva z práve všetkých slov $w \in \Sigma^*$, ktoré automat $\mathcal{A}$ prečíta počas nejakého svojho behu zo stavu i do stavu j .

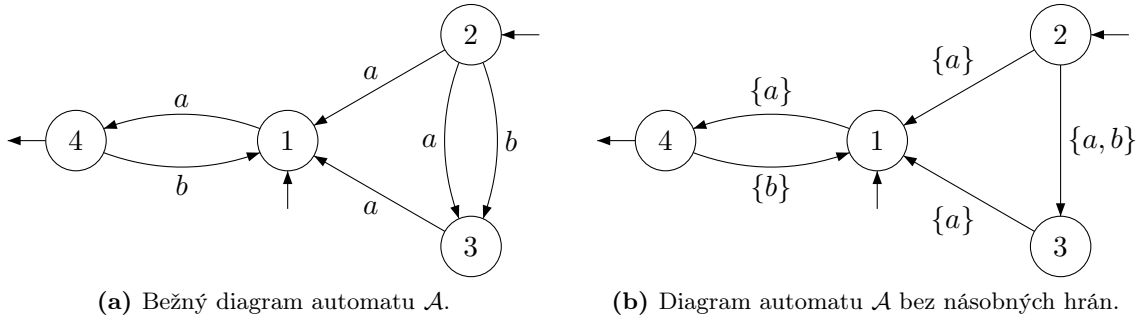
Množiny I, F ďalej možno „zakódovať“ ako vektory⁸: množinu I možno zadať ako riadkový vektor $\mathbf{i} = (\iota_1, \dots, \iota_n)$ a množinu F možno zadať ako stĺpcový vektor $\mathbf{f} = (\tau_1, \dots, \tau_n)^T$ tak, že pre $k = 1, \dots, n$ je

$$\iota_k = \begin{cases} \{\varepsilon\} & \text{ak } k \in I, \\ \emptyset & \text{ak } k \notin I, \end{cases} \quad \tau_k = \begin{cases} \{\varepsilon\} & \text{ak } k \in F, \\ \emptyset & \text{ak } k \notin F. \end{cases}$$

Zjavne potom

$$\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f}. \quad (1.2)$$

Mohli by sme teda definovať konečný automat $\mathcal{A}$ aj ako štvoricu $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ a jazyk $\|\mathcal{A}\|$ rozoznávajúci automat $\mathcal{A}$ definovať vzťahom (1.2).



Obr. 1.4: Konečný automat $\mathcal{A}$.

Pre ilustráciu uvažujme konečný automat $\mathcal{A}$ s diagramom znázorneným na obrázku 1.4a, alebo – ak namiesto násobných hrán uvažujeme hrany ohodnotené konečnými jazykmi – na obrázku 1.4b.

⁷Formalizáciu definícií pojmov ako „beh“, „beh dĺžky t “, alebo „slovo prečítané počas behu“ prenechávame čitateľovi ako jednoduché cvičenie, ktoré ale má viac než jedno správne riešenie. Uvedené tvrdenie by však malo zostať pravdivé bez ohľadu na zvolený formalizačný prístup.

⁸Vektory môžeme stotožniť, podľa toho či sú riadkové alebo stĺpcové, s maticami o jedinom riadku alebo stĺpci. Nemusíme tak nanovo definovať operáciu násobenia vektora s maticou, ktorú môžeme chápať ako špeciálny prípad maticového násobenia.

Prechodovou maticou tohto automatu je

$$A = \begin{pmatrix} \emptyset & \emptyset & \emptyset & \{a\} \\ \{a\} & \emptyset & \{a, b\} & \emptyset \\ \{a\} & \emptyset & \emptyset & \emptyset \\ \{b\} & \emptyset & \emptyset & \emptyset \end{pmatrix}$$

a vektory $\mathbf{i}$ a $\mathbf{f}$ zodpovedajúce počiatočným a koncovým stavom sú dané ako

$$\mathbf{i} = (\{\varepsilon\}, \{\varepsilon\}, \emptyset, \emptyset) \quad \text{a} \quad \mathbf{f} = (\emptyset, \emptyset, \emptyset, \{\varepsilon\})^T.$$

Pri vynaložení trochy mechanického úsilia možno dokázať, že

$$A^* = \begin{pmatrix} (ab)^* & \emptyset & \emptyset & (ab)^*a \\ a(ab)^* \cup \{a, b\}a(ab)^* & \{\varepsilon\} & \{a, b\} & a(ab)^*a \cup \{a, b\}a(ab)^*a \\ a(ab)^* & \emptyset & \{\varepsilon\} & a(ab)^*a \\ (ba)^*b & \emptyset & \emptyset & (ba)^* \end{pmatrix},$$

z čoho pre jazyk $\|\mathcal{A}\|$ rozoznávaný automatom $\mathcal{A}$ dostávame

$$\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f} = (ab)^*a \cup (a(ab)^*a \cup \{a, b\}a(ab)^*a) = \{\varepsilon, a, aa, ba\}(ab)^*a.$$

Príklad 1.2.3. Uvažujme teraz ohodnotené grafy nad množinou prirodzených čísel $\mathbb{N}$. Už v rámci príkladu 1.1.9 sme učinili pozorovanie, že takéto ohodnotené grafy možno považovať za reprezentáciu orientovaných *multigrafov*, čiže orientovaných grafov s násobnými hranami medzi vrcholmi. Hrana z vrcholu i do vrcholu j ohodnotená číslom k znamená, že v zodpovedajúcom multigrafe vedie z vrcholu i do vrcholu j práve k orientovaných hrán.

Uvažujme ďalej na množine $\mathbb{N}$ štandardné operácie sčítavania a násobenia, čím dostávame algebru $(\mathbb{N}, +, \cdot, 0, 1)$. Maticové násobenie nad touto algebrou je tiež definované štandardným spôsobom.

Nech $A = \mathbf{A}(\mathcal{G}) = (a_{i,j})_{n \times n}$ je matica susednosti ohodnoteného grafu $\mathcal{G}$ nad $\mathbb{N}$ s množinou vrcholov $V(\mathcal{G}) = [n]$ pre nejaké $n \in \mathbb{N}$, nech t je prirodzené číslo a $A^t = (a_{i,j}[t])_{n \times n}$. Ľahko potom nahliadnuť, že číslo $a_{i,j}[t]$ je rovné počtu orientovaných sledov dĺžky t vedúcich z vrcholu i do vrcholu j v multigrafe reprezentovanom ohodnoteným grafom $\mathcal{G}$.

Rozumne definovanú iteráciu matice A nad algebrou $(\mathbb{N}, +, \cdot, 0, 1)$ očakávať nemožno: napríklad pre maticu $A = (1)_{1 \times 1}$ by totiž muselo byť $A^* = (1 + 1 + 1 + \dots)_{1 \times 1}$; nekonečný súčet $1 + 1 + 1 + \dots$ ale nad $(\mathbb{N}, +, \cdot, 0, 1)$ neexistuje.

Príklad 1.2.4. Algebru $(\mathbb{N}, +, \cdot, 0, 1)$ z predchádzajúceho príkladu možno rozšíriť pridaním prvku ∞ na algebru $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$, kde súčet a súčin dvojíc prvkov z $\mathbb{N}$ je daný štandardným spôsobom a kde pre všetky $a \in \mathbb{N} \cup \{\infty\}$ je $a + \infty = \infty + a = \infty$ a

$$a \cdot \infty = \infty \cdot a = \begin{cases} 0 & \text{ak } a = 0, \\ \infty & \text{inak.} \end{cases}$$

Mocnina matice susednosti $A = (a_{i,j})_{n \times n}$ ohodnoteného grafu $\mathcal{G}$ nad $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$ má potom podobnú interpretáciu ako v predchádzajúcom príklade. Nad algebrou $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$ ale navyše možno definovať nekonečný súčet cez indexovú množinu I predpisom

$$\sum_{i \in I} a_i = \begin{cases} \sum_{i \in I_+} a_i & \text{ak je množina } I_+ = \{i \in I \mid a_i \neq 0\} \text{ konečná,} \\ \infty & \text{inak,} \end{cases}$$

pričom nekonečný súčet matíc možno definovať po zložkách. Iterácia $A^* = (a_{i,j}[*])_{n \times n}$ matice A je potom daná ako

$$A^* = A^0 + A^1 + A^2 + \dots = \sum_{t \in \mathbb{N}} A^t.$$

Ak v multigrafe zodpovedajúcom ohodnotenému grafu $\mathcal{G}$ existuje konečne veľa orientovaných sledov z vrcholu i do vrcholu j , číslo $a_{i,j}[*]$ zrejme udáva ich počet; v opačnom prípade je $a_{i,j}[*] = \infty$.

Príklad 1.2.5. Opäť uvažujme ohodnotené grafy nad množinou $\mathbb{N} \cup \{\infty\}$, avšak tentokrát interpretujme ohodnotenie ako *cenu hrany*. Hrany s cenou ∞ pritom budeme stotožňovať s neexistujúcimi hranami; nulou v našej množine ohodnotení je teda prvok ∞ .

Na množine $\mathbb{N} \cup \{\infty\}$ uvažujme ako aditívnu operáciu minimum a ako multiplikatívnu operáciu súčet. Aditívnym neutrálnym prvkom je ∞ a multiplikatívnym neutrálnym prvkom je 0 – dostávame teda algebru $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$.

Násobenie matíc nad $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ možno prirodzene definovať takto: ak $A = (a_{i,j})_{m \times n}$ a $B = (b_{i,j})_{n \times r}$ sú matice nad $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$, je súčin $A \cdot B = (c_{i,j})_{m \times r}$ daný pre $i = 1, \dots, m$ a $j = 1, \dots, r$ predpisom

$$c_{i,j} = \min_{k=1,\dots,n} (a_{i,k} + b_{k,j}).$$

Mocnina štvorcovej matice A typu $n \times n$ je potom daná rekurzívne: $A^0 = \mathbf{I}_n$ a $A^{t+1} = A^t \cdot A$ pre všetky $t \in \mathbb{N}$, kde

$$\mathbf{I}_n = \begin{pmatrix} 0 & \infty & \dots & \infty \\ \infty & 0 & \dots & \infty \\ \vdots & \vdots & \ddots & \vdots \\ \infty & \infty & \dots & 0 \end{pmatrix}.$$

Pre t -tu mocninu $A^t = (a_{i,j}[t])_{n \times n}$ matice susednosti $A = (a_{i,j})_{n \times n}$ ohodnoteného grafu $\mathcal{G}$ nad $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ možno ľahko dokázať, že *prvok $a_{i,j}[t]$ udáva cenu najlacnejšieho sledu dĺžky práve t vedúceho z vrcholu i do vrcholu j* (v prípade neexistencie takéhoto sledu chápanú ako ∞).

Iteráciu $A^* = (a_{i,j}[*])_{n \times n}$ matice A možno navyše definovať ako

$$A^* = \min_{t \in \mathbb{N}} A^t = \left(\min_{t \in \mathbb{N}} a_{i,j}[t] \right)_{n \times n}$$

(čiže minimum je definované po zložkách). Ľahko potom vidieť, že $a_{i,j}[*]$ udáva cenu najlacnejšieho sledu z vrcholu i do vrcholu j , alebo ekvivalentne cenu najlacnejšej cesty z vrcholu i do vrcholu j .

Príklad 1.2.6. Predpokladajme, že sieť kuriérnych spojení je daná grafom, v ktorom je každá hrana ohodnotená spoľahlivosťou príslušného spojenia. Tou je číslo r z intervalu $[0, 1]$ dané ako $r = (1 - p)$, kde p je pravdepodobnosť straty zásielky. Neexistujúce spojenia pritom budeme stotožňovať so spojeniami s nulovou spoľahlivosťou.

Uvažujme teraz na množine $[0, 1]$ ako aditívnu operáciu suprénum⁹ a ako multiplikatívnu operáciu štandardné násobenie reálnych čísel, čím dostávame algebru $([0, 1], \sup, \cdot, 0, 1)$.

Maticové násobenie nad $([0, 1], \sup, \cdot, 0, 1)$ možno definovať takto: ak $A = (a_{i,j})_{m \times n}$ a $B = (b_{i,j})_{n \times r}$ sú matice nad $([0, 1], \sup, \cdot, 0, 1)$, je ich súčin $A \cdot B = (c_{i,j})_{m \times r}$ daný pre $i = 1, \dots, m$ a $j = 1, \dots, r$ ako

$$c_{i,j} = \sup_{k=1,\dots,n} a_{i,k} \cdot b_{k,j}.$$

Štandardne potom možno definovať aj mocninu štvorcovej matice. Pre t -tu mocninu $A^t = (a_{i,j}[t])_{n \times n}$ matice susednosti $A = (a_{i,j})_{n \times n}$ ohodnoteného grafu $\mathcal{G}$ nad $([0, 1], \sup, \cdot, 0, 1)$ pritom možno dokázať, že *prvok $a_{i,j}[t]$ udáva maximálnu spoľahlivosť sledu dĺžky práve t vedúceho z vrcholu i do vrcholu j* .

Ak navyše definujeme iteráciu $A^* = (a_{i,j}[*])_{n \times n}$ matice A ako

$$A^* = \sup_{t \in \mathbb{N}} A^t = \left(\sup_{t \in \mathbb{N}} a_{i,j}[t] \right)_{n \times n},$$

udáva hodnota $a_{i,j}[*]$ maximálnu spoľahlivosť sledu vedúceho z vrcholu i do vrcholu j , alebo ekvivalentne maximálnu spoľahlivosť cesty z i do j .¹⁰

⁹Binárne suprénum nie je nič iné než binárne maximum. V skutočnosti by sme si v celom tomto príklade vystačili s maximami – miestami ale nebude *a priori* zaručené, že tieto maximá budú existovať.

¹⁰Nie je ťažké vidieť, že sa tu skutočne nadobúda maximum, hoci maximum nekonečnej množiny prvkov intervalu $[0, 1]$ nemusí existovať. Dôvodom je fakt, že spoľahlivosť ľubovoľného sledu z i do j možno zhora ohraničiť spoľahlivosťou nejakej cesty z i do j ; ciest z i do j je však vždy iba konečne veľa.

Príklad 1.2.7. Uvažujme teraz nasledujúcu situáciu: dopravca má k dispozícii množinu vozidiel M , ktoré vypravuje po cestnej sieti danej orientovaným grafom – hrany v ňom zodpovedajú jednotlivým cestným úsekom. Nie všetky vozidlá ale možno vypraviť na všetky úseky (či už kvôli ich hmotnosti, rozmerom, alebo chýbajúcim povoleniam). Každá hrana grafu je preto ohodnotená množinou tých vozidiel, ktoré po zodpovedajúcom cestnom úseku môžu prejsť – ide teda o ohodnotený graf nad množinou 2^M . Po neexistujúcich hranách nemôže prejsť žiadne vozidlo – za nulu množiny 2^M teda budeme považovať prázdnu množinu.

Na množine 2^M uvažujme ako aditívnu operáciu zjednotenie a ako multiplikatívnu operáciu prienik – aditívnym neutrálnym prvkom je teda $\emptyset$ a multiplikatívnym neutrálnym prvkom je M . Vo výsledku tak dostávame algebru $(2^M, \cup, \cap, \emptyset, M)$.

Ak $A = (A_{i,j})_{m \times n}$ a $B = (B_{i,j})_{n \times r}$ sú matice nad $(2^M, \cup, \cap, \emptyset, M)$, je ich súčin $A \cdot B = (C_{i,j})_{m \times r}$ definovaný pre $i = 1, \dots, m$ a $j = 1, \dots, r$ predpisom

$$C_{i,j} = \bigcup_{k=1}^n (A_{i,k} \cap B_{k,j}).$$

Pre štandardne definovanú t -tu mocninu $A^t = (A_{i,j}[t])_{n \times n}$ matice susednosti $A = (A_{i,j})_{n \times n}$ ohodnoteného grafu $\mathcal{G}$ nad $(2^M, \cup, \cap, \emptyset, M)$ potom platí, že množina $A_{i,j}[t]$ pozostáva z práve tých vozidiel z množiny M , ktoré možno nejakou trasou dĺžky t vypraviť z vrcholu i do vrcholu j .

Nad algebrou $(2^M, \cup, \cap, \emptyset, M)$ možno definovať aj iteráciu $A^* = (A_{i,j}[*])_{n \times n}$ matice A ako

$$A^* = \bigcup_{t=0}^{\infty} A^t;$$

množina $A_{i,j}[*]$ potom pozostáva z práve tých vozidiel z množiny M , ktoré možno nejakou trasou vypraviť z vrcholu i do vrcholu j .

Poznámka 1.2.8. Vo všetkých príkladoch vyššie bolo v uvažovanej algebre dobre definované maticové násobenie. Treba však upozorniť na skutočnosť, že táto definícia vždy skryte využívala niektoré vlastnosti operácií danej algebry: ťažko si napríklad predstaviť ozaj zmysluplnú definíciu násobenia matíc nad algebrou, ktorej aditívna operácia nie je komutatívna alebo asociatívna. Analýzou požiadaviek nutných na „rozumnú“ definíciu maticového násobenia dospejeme v oddiele 1.3 k pojmu *polokruhu*.

Poznámka 1.2.9. Iteráciu matice bolo možné definovať vo všetkých vyššie opísaných algebrách okrem tej z príkladu 1.2.3. Kľúčovým faktorom sa tu zo zrejmých príčin ukazuje byť existencia nekonečných súčtov. Analýzou tejto požiadavky dospejeme v oddiele 1.4 k pojmom *úplného* a *spočítateľne úplného polokruhu*; všeobecnejší prístup neskôr opíšeme v oddiele 3.13.

1.3 Polokruhy

Príklady z predchádzajúceho oddielu vykazujú relatívne veľkú mieru podobnosti: zakaždým sme definovali algebru $(S, \oplus, \otimes, 0, 1)$ s dvoma binárnymi operáciami a ich neutrálnymi prvkami. Nad týmito štruktúrami boli – vďaka určitým ich vlastnostiam – „rozumne“ definované operácie maticového násobenia a mocniny štvorcovej matice. Mocninu matice susednosti ohodnoteného grafu nad S sme potom interpretovali ako riešenie nejakej prirodzenej úlohy na ohodnotených grafoch. V podstate išlo vždy o tie isté pozorovania, akurát zakaždým v jemne odlišnom kabáte.

Posúva sa teda otázka, či možno príklady z predchádzajúceho oddielu študovať v spoločnom rámci. Ten by mala tvoriť trieda algebier $(S, \oplus, \otimes, 0, 1)$ s „rozumne“ definovaným maticovým násobením a mocnením, ktorá zahŕňa všetky konkrétne algebry z predchádzajúceho oddielu. V nasledujúcich úvahách dospejeme k pozorovaniu, že takýto rámec môže poskytnúť pojem *polokruhu* – jeho definíciu teraz „odvodíme“ analýzou vlastností algebry $(S, \oplus, \otimes, 0, 1)$ potrebných na definíciu „rozumného“ maticového násobenia a mocniny štvorcovej matice.

Súčin matíc $A = (a_{i,j})_{m \times n}$ a $B = (b_{i,j})_{n \times r}$ nad $(S, \oplus, \otimes, 0, 1)$ by sme chceli definovať ako maticu $A \cdot B = (c_{i,j})_{m \times r}$, kde pre $i = 1, \dots, m$ a $j = 1, \dots, r$ je

$$c_{i,j} = \bigoplus_{k=1}^n a_{i,k} \otimes b_{k,j}. \quad (1.3)$$

V tejto definícii sa implicitne predpokladá asociatívnosť operácie $\oplus$ – inak by použitie „veľkého“ $\bigoplus$ nebolo namieste. Každá algebra $(S, \oplus, \otimes, 0, 1)$ z nami hľadanej triedy teda musí mať asociatívnu operáciu $\oplus$. Je tiež žiaduce, aby bolo možné vzťah (1.3) prepísať do podoby

$$c_{i,j} = \bigoplus_{k \in [n]} a_{i,k} \otimes b_{k,j},$$

čiže aby výsledok nezávisel od poradia sčítancov – umožnia sa tým bežné „prečíslovania“ indexovej množiny matice (množiny vrcholov grafu). Tomu zodpovedá požiadavka komutatívnosti operácie $\oplus$. Operácia $\oplus$ teda musí byť asociatívna a komutatívna.

Akonáhle bez ďalšieho upresnenia hovoríme o umocňovaní matíc, obyčajne implicitne predpokladáme, že pre všetky štvorcové matice A je

$$A \cdot (A \cdot A) = (A \cdot A) \cdot A; \quad (1.4)$$

napríklad zápis A^3 by totiž v opačnom prípade nebol jednoznačný. Pre všetky štvorcové matice $A = (a_{i,j})_{n \times n}$ by teda mala platiť rovnosť (1.4) a z definície násobenia matíc tak pre $i, j = 1, \dots, n$ dostávame

$$\bigoplus_{k=1}^n a_{i,k} \otimes \left(\bigoplus_{\ell=1}^n a_{k,\ell} \otimes a_{\ell,j} \right) = \bigoplus_{\ell=1}^n \left(\bigoplus_{k=1}^n a_{i,k} \otimes a_{k,\ell} \right) \otimes a_{\ell,j}. \quad (1.5)$$

Od „rozumných“ algebier $(S, \oplus, \otimes, 0, 1)$ pritom očakávame, že sa táto spoločná hodnota bude dať zapísať ako

$$\bigoplus_{k=1}^n \bigoplus_{\ell=1}^n a_{i,k} \otimes a_{k,\ell} \otimes a_{\ell,j}. \quad (1.6)$$

V uvedenom zápise je implicitná asociatívnosť operácie $\otimes$; aby sa navyše obidve strany rovnosti (1.5) skutočne rovnali výrazu (1.6), mali by v algebre $(S, \oplus, \otimes, 0, 1)$ platiť distributívne zákony.

Od umocňovania matíc je tiež „rozmenné“ pre každú maticu A typu $n \times n$ očakávať platnosť rovnosti $A^0 = \mathbf{I}_n$, kde $\mathbf{I}_n$ je jednotková matica typu $n \times n$. To je jedným z možných odôvodnení nášho predpokladu existencie neutrálnych prvkov – jednotková matica by totiž mala mať na diagonále jednotky a mimo nej nuly. Od jednotkovej matice navyše očakávame, že bude spĺňať rovnosť

$$\mathbf{I}_n \cdot A = A \cdot \mathbf{I}_n = A$$

pre všetky matice A typu $n \times n$. Z definície maticového násobenia a neutrality prvku 1 vzhľadom na $\otimes$ vyplýva, že túto vlastnosť možno zaručiť požiadavkou $a \otimes 0 = 0 \otimes a = 0$ pre všetky $a \in S$.

Takýmto spôsobom napokon prichádzame k definícii triedy algebier, ktoré sú v literatúre známe ako *polokruhy* – v podstate totiž ide o „okruhy bez odčítania“. ¹¹ Polokruhy tak z istého pohľadu možno chápať ako algebry s „rozumným“ násobením a umocňovaním matíc.

¹¹To je pravda v prípade, že v definícii okruhu požadujeme existenciu jednotky. V tomto texte budeme pod okruhom vždy rozumieť okruh s jednotkou.

Definícia 1.3.1. *Polokruh* je usporiadaná päťica $(S, \oplus, \otimes, 0, 1)$, kde S je množina, $\oplus, \otimes: S \times S \rightarrow S$ sú binárne operácie na S a $0, 1$ sú prvky S , spĺňajúca nasledujúce podmienky:

(A1) Pre všetky $a, b, c \in S$ je $a \oplus (b \oplus c) = (a \oplus b) \oplus c$.

(A2) Pre všetky $a \in S$ je $a \oplus 0 = 0 \oplus a = a$.

(A3) Pre všetky $a, b \in S$ je $a \oplus b = b \oplus a$.

(M1) Pre všetky $a, b, c \in S$ je $a \otimes (b \otimes c) = (a \otimes b) \otimes c$.

(M2) Pre všetky $a \in S$ je $a \otimes 1 = 1 \otimes a = a$.

(D1) Pre všetky $a, b, c \in S$ je $a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$.

(D2) Pre všetky $a, b, c \in S$ je $(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$.

(Z) Pre všetky $a \in S$ je $0 \otimes a = a \otimes 0 = 0$.

Poznámka 1.3.2. Päťica $(S, \oplus, \otimes, 0, 1)$ je teda polokruh práve vtedy, keď $(S, \oplus, 0)$ je komutatívny monoid, $(S, \otimes, 1)$ je monoid a súčasne sú splnené podmienky (D1), (D2) a (Z).

Poznámka 1.3.3. Ak nebude hroziť nedorozumenie, budeme o polokruhu $(S, \oplus, \otimes, 0, 1)$ niekedy hovoriť len ako o polokruhu S .

Definícia 1.3.4. Polokruh $(S, \oplus, \otimes, 0, 1)$ nazveme *komutatívnym*, ak je komutatívna operácia $\otimes$, čiže ak $a \otimes b = b \otimes a$ pre všetky $a, b \in S$.

Je jednoduchým mechanickým cvičením presvedčiť sa, že všetky algebry $(S, \oplus, \otimes, 0, 1)$ z príkladov predchádzajúceho oddielu sú v skutočnosti polokruhmi. Pri niektorých z týchto polokruhov sa navyše zaužívali ich samostatné pomenovania: napríklad $(\mathbb{B}, \vee, \wedge, 0, 1)$ tvorí tzv. *booleovský polokruh*, pri päťici $(2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\})$ hovoríme o *polokruhu formálnych jazykov* nad abecedou Σ , pri algebre $([0, 1], \max, \cdot, 0, 1) = ([0, 1], \sup, \cdot, 0, 1)$ o *Viterbiho polokruhu* a algebry dané päťicami ako napríklad $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$, $(\mathbb{Z} \cup \{\infty\}, \min, +, \infty, 0)$, alebo $(\mathbb{R}_{\geq 0} \cup \{\infty\}, \min, +, \infty, 0)$ sa zvyknú nazývať *min-plus polokruhmi* alebo tiež *tropickými polokruhmi*.¹²

Špeciálnym prípadom polokruhu je *triviálny polokruh* $(\{0\}, \circ, \circ, 0, 0)$, kde $\circ: \{0\} \times \{0\} \rightarrow \{0\}$ je jediná binárna operácia na jednoprvkovej množine $\{0\}$. Očividne ide (až na izomorfizmus¹³) o jediný jednoprvkový polokruh. Všetky ostatné – čiže aspoň dvojprvkové – polokruhy nazývame *netriviálnymi*. Nasledujúce jednoduché tvrdenie ukazuje, že v netriviálnom polokruhu sú – v súlade s intuíciou – nula a jednotka vždy dva *rozdielne* prvky.

Tvrdenie 1.3.5. *Nech $(S, \oplus, \otimes, 0, 1)$ je netriviálny polokruh. Potom $0 \neq 1$.*

Dôkaz. V netriviálnom polokruhu S môžeme zvoliť aspoň jeden prvok $a \neq 0$. Preň vďaka podmienke (Z) platí $a \otimes 0 = 0$, kým s použitím (M2) dostávame $a \otimes 1 = a$. Z predpokladu $0 = 1$ by sme teda odvodili rovnosť $a = 0$: spor. $\square$

¹²Pomenovanie *tropický polokruh* vzniklo na počesť matematika I. Simona pôsobiaceho v Brazílii [79, 92]. Úzko súvisiacimi algebrami sú tzv. *max-plus polokruhy* ako napríklad $(\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$, $(\mathbb{Z} \cup \{-\infty\}, \max, +, -\infty, 0)$ a $(\mathbb{R}_{\geq 0} \cup \{-\infty\}, \max, +, -\infty, 0)$, prípadne ich varianty navyše obsahujúce aj $+\infty$. Max-plus polokruhy sa tiež niekedy zvyknú nazývať *arktíckymi polokruhmi* [29]. Mnemotechnickou pomôckou tu môže byť nasledovná interpretácia: v tropickom polokruhu je „horúco“ ($+\infty$), preto je žiaduce teplotu čo najviac znížiť (min); v arktickom polokruhu je naopak „zima“ ($-\infty$) a cieľom tak môže byť čo najlepšie zakúriť (max). Iní autori však nazývajú tropickými ako min-plus polokruhy, tak aj max-plus polokruhy [79].

¹³Predpokladáme, že čitateľ dokáže pojmy ako homomorfizmus alebo izomorfizmus polokruhov, ktoré sa prakticky ničím nelíšia od ich obdôb pre okruhy alebo polia, definovať aj samostatne ako ľubovoľné resp. bijektívne zobrazenia zachovávajúce obidve binárne operácie a *obidva neutrálne prvky* (prípadne možno tieto definície nájsť v zadaní cvičenia 6 na konci tejto kapitoly). Izomorfne algebry sa obvyčajne stotožňujú a budeme tak činiť aj v tomto texte.

Pri aplikáciách sa niekedy zvykne predpokladať netriviálnosť uvažovaných polokruhov – triviálny polokruh je prakticky bezvýznamný a z predchádzajúcich pozorovaní je zrejmé, že občas môže mať trochu osobité vlastnosti. Z pohľadu algebry je naopak existencia triviálneho polokruhu dôležitá a tento polokruh nemožno tak ľahko z úvah vylúčiť; iba vďaka nemu je napríklad trieda všetkých polokruhov uzavretá na homomorfné obrazy (pri homomorfizmoch *algebier* príslušného typu).

Poznámka 1.3.6. V nasledujúcom texte budeme *každý* polokruh $(S, \oplus, \otimes, 0, 1)$ súčasne považovať aj za množinu s nulou $(S, 0)$. To znamená, že pri grafoch ohodnotených prvkami polokruhu budeme dvojice vrcholov ohodnotené nulou v tomto polokruhu považovať za neexistujúce hrany.

Môžeme teraz definovať kľúčové operácie na maticiach nad polokruhmi: popri operáciách sčítania a násobenia matice skalárom, definovaných pre všetky matice $A = (a_{i,j})_{m \times n}$, $B = (b_{i,j})_{m \times n}$ nad polokruhom $(S, \oplus, \otimes, 0, 1)$ a všetky $q \in S$ ako $A + B = (a_{i,j} \oplus b_{i,j})_{m \times n}$ a $qA = (q \otimes a_{i,j})_{m \times n}$, pre nás budú dôležité najmä operácie násobenia a umocňovania matíc.

Definícia 1.3.7. Nech $(S, \oplus, \otimes, 0, 1)$ je polokruh a $A = (a_{i,j})_{m \times n}$, $B = (b_{i,j})_{n \times r}$ sú matice nad S . *Súčin* matíc A a B je matica $A \cdot B = (c_{i,j})_{m \times r}$ daná pre $i = 1, \dots, m$ a $j = 1, \dots, r$ predpisom

$$c_{i,j} = \bigoplus_{k=1}^n a_{i,k} \otimes b_{k,j}.$$

Definícia 1.3.8. Nech $(S, \oplus, \otimes, 0, 1)$ je polokruh a n je prirodzené číslo. *Jednotkovou maticou* typu $n \times n$ nad polokruhom S nazveme maticu

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Je elementárnym cvičením dokázať, že $\mathbf{I}_n \cdot A = A$ pre všetky matice $A = (a_{i,j})_{n \times r}$ nad S a $A \cdot \mathbf{I}_n = A$ pre všetky matice $A = (a_{i,j})_{m \times n}$ nad S .

Definícia 1.3.9. Nech $(S, \oplus, \otimes, 0, 1)$ je polokruh, $n \in \mathbb{N}$ a A je štvorcová matica typu $n \times n$ nad S . *Mocniny* matice A definujeme indukčným predpisom $A^0 = \mathbf{I}_n$ a $A^{t+1} = A^t \cdot A$ pre všetky $t \in \mathbb{N}$.

Naším najbližším cieľom teraz bude urobiť všeobecné pozorovanie o grafovej interpretácii mocniny štvorcovej matice nad polokruhom $(S, \oplus, \otimes, 0, 1)$, ktorého špeciálnymi prípadmi by mali byť jednotlivé interpretácie z príkladov predchádzajúceho oddielu. Pod *sledom* z vrcholu u do vrcholu v v ohodnotenom grafe $\mathcal{G} = (V, \varphi)$ nad $(S, \oplus, \otimes, 0, 1)$ budeme rozumieť konečnú postupnosť

$$\gamma = (u_0, u_1, u_2, \dots, u_t)$$

vrcholov $u_0, u_1, \dots, u_t \in V(\mathcal{G})$ takú, že $t \in \mathbb{N}$, $u_0 = u$, $u_t = v$ a pre $i = 1, \dots, t$ je $(u_{i-1}, u_i) \in E(\mathcal{G})$; pod *dĺžkou* sledu γ budeme rozumieť číslo $|\gamma| := t$. Ak navyše pre $i = 1, \dots, t$ položíme $e_i := (u_{i-1}, u_i)$, definujeme *ohodnotenie* sledu γ ako prvok $\varphi(\gamma)$ polokruhu S daný predpisom

$$\varphi(\gamma) := \varphi(e_1) \otimes \varphi(e_2) \otimes \dots \otimes \varphi(e_t).$$

Označenie $\mathcal{W}_{u,v}(\mathcal{G}, t)$ napokon budeme používať pre *množinu všetkých sledov dĺžky t vedúcich v grafe $\mathcal{G}$ z vrcholu u do vrcholu v* .

Nasledujúca veta poskytuje grafovú interpretáciu t -tej mocniny A^t matice susednosti A ohodnoteného grafu nad polokruhom $(S, \oplus, \otimes, 0, 1)$: matica A^t obsahuje na pozícii v i -tom riadku a j -tom stĺpci súčet ohodnotení všetkých sledov dĺžky t z vrcholu i do vrcholu j .

Veta 1.3.10. Nech $(S, \oplus, \otimes, 0, 1)$ je polokruh, $n \in \mathbb{N}$ a $\mathcal{G} = ([n], \varphi)$ je ohodnotený graf nad S . Pre $t \in \mathbb{N}$ označme $\mathbf{A}(\mathcal{G})^t = (a_{i,j}[t])_{n \times n}$. Potom pre všetky $t \in \mathbb{N}$ a $i, j = 1, \dots, n$ je

$$a_{i,j}[t] = \bigoplus_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G}, t)} \varphi(\gamma).$$

Dôkaz. Množina $\mathcal{W}_{i,j}(\mathcal{G}, t)$ je zjavne konečná pre všetky i, j, t ; znenie vety teda dáva zmysel.

Pre $t = 0$ je dokazované tvrdenie zrejmé: množina $\mathcal{W}_{i,j}(\mathcal{G}, 0)$ totiž pre $i = j$ obsahuje jediný sled $\gamma = (i)$ s $\varphi(\gamma) = 1$ a pre $i \neq j$ neobsahuje žiaden sled. To súhlasí s rovnosťou $\mathbf{A}(\mathcal{G})^0 = \mathbf{I}_n$.

Predpokladajme, že tvrdenie platí pre $t = s$ a dokážme ho pre $t = s + 1$. Nech $\mathbf{A}(\mathcal{G}) = (a_{i,j})_{n \times n}$. Keďže $\mathbf{A}(\mathcal{G})^{s+1} = \mathbf{A}(\mathcal{G})^s \cdot \mathbf{A}(\mathcal{G})$, z definície maticového násobenia pre $i, j = 1, \dots, n$ dostávame

$$a_{i,j}[s+1] = \bigoplus_{k=1}^n a_{i,k}[s] \otimes a_{k,j}.$$

To možno podľa indukčného predpokladu a definície matice susednosti prepísať ako

$$\bigoplus_{k=1}^n \left(\bigoplus_{\beta \in \mathcal{W}_{i,k}(\mathcal{G}, s)} \varphi(\beta) \right) \otimes \varphi(k, j) \stackrel{(D2)}{=} \bigoplus_{k=1}^n \bigoplus_{\beta \in \mathcal{W}_{i,k}(\mathcal{G}, s)} (\varphi(\beta) \otimes \varphi(k, j)).$$

Keďže má každý sled γ dĺžky $s+1$ vedúci z i do j práve jeden predposledný vrchol k a keďže sled β dĺžky s z i do k nemožno „predĺžiť“ na sled dĺžky $s+1$ z i do j iba ak $\varphi(k, j) = \varphi(\beta) \otimes \varphi(k, j) = 0$, dá sa uvedený výraz ďalej prepísať ako

$$\bigoplus_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G}, s+1)} \varphi(\gamma),$$

čo bolo treba dokázať. □

1.4 Úplné a spočítateľne úplné polokruhy

Vráťme sa ešte raz k príkladom z oddielu 1.2, v ktorých sme – ako sme následne ukázali v oddiele 1.3 – zakaždým definovali nejaký konkrétny *polokruh* $(S, +, \cdot, 0, 1)$: čiže algebru s „rozumne“ definovaným násobením a umocňovaním matíc, ktoré má často prirodzenú grafovú interpretáciu.

Poznámka 1.4.1. Počnúc týmto oddielom už väčšinou nebudeme operácie polokruhu označovať symbolmi $\oplus$ a $\otimes$, ktoré sme používali na zdôraznenie abstraktnosti uvažovanej algebry. Naopak budeme – podobne ako pred chvíľou – používať bežnejšie symboly $+$ pre sčítanie a $\cdot$ pre násobenie.

Nad *takmer* všetkými polokruhmi skonštruovanými v príkladoch oddielu 1.2 bolo navyše možné definovať *iteráciu matice*, opäť s prirodzenou interpretáciou v reči grafov. Výnimkou tu bol polokruh $(\mathbb{N}, +, \cdot, 0, 1)$ prirodzených čísel so štandardným sčítaním a násobením (príklad 1.2.3). Keby mali všetky štvorcové matice A nad týmto polokruhom dobre definovanú iteráciu $A^* = A^0 + A^1 + A^2 + \dots$, museli by byť definované aj nekonečné súčty ako napríklad $1 + 1 + 1 + \dots$ – to však zjavne nie je možné za predpokladu „čo i len trochu rozumných“ vlastností týchto nekonečných súčtov a už toľko nie za predpokladu, že by malo ísť o štandardné súčty nekonečných radov. Po pridaní nového prvku ∞ a dodefinovaní operácií $+$ a $\cdot$ pre tento novopridaný prvok (príklad 1.2.4) sme už ale získali polokruh, v ktorom boli definované ako nekonečné súčty, tak aj iterácia matice.

Lahko možno vidieť, že existencia nekonečných súčtov v polokruhu je *postačujúcou* podmienkou existencie iterácie matice – nekonečný súčet matíc je len matica nekonečných súčtov. V nasledujúcom túto požiadavku formalizujeme, čím definujeme hneď dva pojmy: *úplné polokruhy* budú mať definované súčty cez ľubovoľnú konečnú alebo nekonečnú indexovú množinu a v *spočítateľne úplných polokruhoch* budú definované súčty cez ľubovoľnú konečnú alebo *spočítateľne* nekonečnú indexovú množinu – každý úplný polokruh teda súčasne bude aj spočítateľne úplným polokruhom.¹⁴ V oddiele 3.13 neskôr naznačíme o niečo všeobecnejší prístup ku skúmaniu polokruhov s iteráciou.

¹⁴V podstate všetky vlastnosti úplných polokruhových relevantné z hľadiska ohodnotených grafov a teórie automatov sú zachované aj pri spočítateľne úplných polokruhoch. Pojem úplného polokruhu tak uvádzame iba preto, že sa v literatúre vyskytuje častejšie, než pojem spočítateľne úplného polokruhu [30, 94] (dôvodom môže byť skutočnosť, že spočítateľne úplné polokruhy významné z hľadiska aplikácií sú väčšinou aj úplnými polokruhmi).

Systémom prvkov množiny X nazveme ľubovoľné zobrazenie $F: I \rightarrow X$, kde I je množina indexov – v zásade teda ide o „postupnosť“ prvkov množiny X indexovaných množinou I , ktorá však nemusí byť ani usporiadaná, ani spočítateľná. Ak navyše $F(i) = a_i$ pre všetky $i \in I$, píšeme $F = (a_i \mid i \in I)$. Systém nazveme konečným, nekonečným, resp. spočítateľne nekonečným, ak má príslušnú vlastnosť množina I . Triedu¹⁵ všetkých systémov prvkov X budeme označovať $\mathcal{F}(X)$ a triedu všetkých konečných alebo spočítateľne nekonečných systémov prvkov X budeme označovať $\mathcal{F}_\omega(X)$. *Zovšeobecneným rozkladom* množiny Y nazveme systém $(Y_j \mid j \in J)$ podmnožín Y taký, že $\bigcup_{j \in J} Y_j = Y$ a $Y_j \cap Y_k = \emptyset$ pre ľubovoľné $j, k \in J$ také, že $j \neq k$. *Rozkladom* množiny Y nazveme zovšeobecnený rozklad $(Y_j \mid j \in J)$ množiny Y taký, že množiny Y_j sú neprázdne pre všetky $j \in J$.

Existenciu nekonečných (resp. spočítateľne nekonečných) súčtov v polokruhu $(S, +, \cdot, 0, 1)$ môžeme formalizovať pomocou zobrazenia $\Phi: \mathcal{F}(S) \rightarrow S$ (resp. $\Phi: \mathcal{F}_\omega(S) \rightarrow S$), pričom pre systém $(a_i \mid i \in I)$ budeme písať

$$\Phi(a_i \mid i \in I) =: \sum_{i \in I} a_i.$$

Je však zrejmé, že od takto chápaných nekonečných súčtov je na zmysluplnú prácu s nimi nutné požadovať určité vlastnosti. Nekonečné súčty by v prvom rade mali byť konzistentné s konečnými súčtami: pre všetky konečné množiny $I = \{i_1, i_2, \dots, i_k\}$ by teda malo byť

$$\Phi(a_i \mid i \in I) = \sum_{i \in I} a_i = a_{i_1} + a_{i_2} + \dots + a_{i_k}.$$

Zápis $\sum_{i \in I} a_i$ implicitne predpokladá „asociatívnosť“ a „komutatívnosť“ nekonečných súčtov – vo všeobecnosti totiž sumujeme cez neusporiadané indexové množiny. Tieto dve vlastnosti možno vyjadriť požiadavkou, aby pre ľubovoľný (resp. ľubovoľný spočítateľný) systém $(a_i \mid i \in I)$ prvkov S a ľubovoľný (resp. ľubovoľný spočítateľný) zovšeobecnený rozklad $(I_j \mid j \in J)$ množiny I bolo

$$\sum_{j \in J} \left(\sum_{i \in I_j} a_i \right) = \sum_{i \in I} a_i.$$

Nakoniec ešte budeme požadovať „distributívnosť“ nekonečných súčtov: pre všetky (resp. pre všetky spočítateľné) systémy $(a_i \mid i \in I)$ prvkov S a každé $a \in S$ by malo byť

$$a \cdot \left(\sum_{i \in I} a_i \right) = \sum_{i \in I} (a \cdot a_i) \quad \text{a} \quad \left(\sum_{i \in I} a_i \right) \cdot a = \sum_{i \in I} (a_i \cdot a).$$

Dostávame tak definíciu *úplných* resp. *spočítateľne úplných* polokruhov.

Definícia 1.4.2. *Úplný polokruh* (resp. *spočítateľne úplný polokruh*) je šesticca $(S, +, \cdot, 0, 1, \Phi)$, kde $(S, +, \cdot, 0, 1)$ je polokruh a $\Phi: \mathcal{F}(S) \rightarrow S$ (resp. $\Phi: \mathcal{F}_\omega(S) \rightarrow S$) je zobrazenie, pre ktoré píšeme

$$\Phi(a_i \mid i \in I) =: \sum_{i \in I} a_i$$

pre všetky $(a_i \mid i \in I) \in \mathcal{F}(S)$ (resp. všetky $(a_i \mid i \in I) \in \mathcal{F}_\omega(S)$), pričom:

- (i) Pre ľubovoľný konečný systém $(a_i \mid i \in I)$ s $I = \{i_1, i_2, \dots, i_k\}$ je $\sum_{i \in I} a_i = a_{i_1} + a_{i_2} + \dots + a_{i_k}$.
- (ii) Pre ľubovoľný (resp. ľubovoľný konečný alebo spočítateľne nekonečný) systém $(a_i \mid i \in I)$ prvkov polokruhu S , ľubovoľnú (resp. ľubovoľnú konečnú alebo spočítateľne nekonečnú) indexovú množinu J a ľubovoľný zovšeobecnený rozklad $(I_j \mid j \in J)$ množiny I je $\sum_{j \in J} \left(\sum_{i \in I_j} a_i \right) = \sum_{i \in I} a_i$.
- (iii) Pre ľubovoľný (resp. ľubovoľný konečný alebo spočítateľne nekonečný) systém $(a_i \mid i \in I)$ prvkov polokruhu S a ľubovoľné $a \in S$ je $a \cdot \left(\sum_{i \in I} a_i \right) = \sum_{i \in I} (a \cdot a_i)$ a $\left(\sum_{i \in I} a_i \right) \cdot a = \sum_{i \in I} (a_i \cdot a)$.

¹⁵Množina všetkých systémov prvkov X , ba dokonca ani len množina všetkých konečných systémov, neexistuje. Po vhodnom obmedzení univerza indexových množín by bolo možné skonštruovať množiny všetkých konečných alebo spočítateľne nekonečných systémov prvkov X ; takýto prístup by však nebol veľmi účelný.

Poznámka 1.4.3. V prípade, že nebude hroziť nedorozumenie, budeme pre (spočítateľne) úplné polokruhy namiesto $(S, +, \cdot, 0, 1, \Phi)$ písať len $(S, +, \cdot, 0, 1)$, prípadne S .

Nie je ťažké presvedčiť sa o tom, že – až na ten z príkladu 1.2.3 – sú všetky polokruhy z príkladov oddielu 1.2 úplné. Z podmienok (i) a (ii) predchádzajúcej definície ďalej vyplýva nasledujúca vlastnosť (spočítateľne) úplných polokruhov, ktorú budeme často – aj implicitne – využívať pri dôkazoch.

Tvrdenie 1.4.4. *Nech $(a_i \mid i \in I)$ je ľubovoľný (konečný alebo spočítateľne nekonečný) systém prvkov (spočítateľne) úplného polokruhu S a $f: J \rightarrow I$ je bijekcia. Potom $\sum_{i \in I} a_i = \sum_{j \in J} a_{f(j)}$.*

Dôkaz. Uvažujme rozklad $(\{f(j)\} \mid j \in J)$ indexovej množiny I . Potom

$$\sum_{i \in I} a_i \stackrel{(ii)}{=} \sum_{j \in J} \left(\sum_{i \in \{f(j)\}} a_i \right) \stackrel{(i)}{=} \sum_{j \in J} a_{f(j)}. \quad \square$$

Každý úplný polokruh je súčasne aj spočítateľne úplný polokruh. V nasledujúcom teda budeme naše tvrdenia formulovať pre spočítateľne úplné polokruhy, pričom analogické tvrdenia budú vždy „automaticky“ platné aj pre úplné polokruhy.

Definícia 1.4.5. Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $a \in S$ a $A \in S^{n \times n}$ pre nejaké prirodzené číslo n . *Iteráciu prvku a a potom definujeme ako*

$$a^* := \sum_{t \in \mathbb{N}} a^t$$

a *iteráciu matice A definujeme ako*

$$A^* := \sum_{t \in \mathbb{N}} A^t,$$

kde nekonečný súčet matíc je definovaný po zložkách.

Nasledujúca veta poskytuje grafovú interpretáciu iterácie štvorcovej matice – ide o zovšeobecnenie interpretácií nad konkrétnymi úplnými polokruhmi z oddielu 1.2. Ako $\mathcal{W}_{u,v}(\mathcal{G})$ budeme označovať množinu všetkých sledov vedúcich v ohodnotenom grafe $\mathcal{G}$ z vrcholu u do vrcholu v .

Veta 1.4.6. *Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $n \in \mathbb{N}$ a $\mathcal{G} = ([n], \varphi)$ je ohodnotený graf nad S . Nech $\mathbf{A}(\mathcal{G})^* = (a_{i,j}[*])_{n \times n}$. Potom pre $i, j = 1, \dots, n$ je*

$$a_{i,j}[*] = \sum_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G})} \varphi(\gamma).$$

Dôkaz. Predpokladajme, že pre každé $t \in \mathbb{N}$ je $\mathbf{A}(\mathcal{G})^t = (a_{i,j}[t])_{n \times n}$. Z definície iterácie matice

$$a_{i,j}[*] = \sum_{t \in \mathbb{N}} a_{i,j}[t]. \quad (1.7)$$

Z vety 1.3.10 ďalej pre všetky $t \in \mathbb{N}$ vyplýva

$$a_{i,j}[t] = \sum_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G}, t)} \varphi(\gamma). \quad (1.8)$$

Keďže ale množinu $\mathcal{W}_{i,j}(\mathcal{G})$ možno vyjadriť ako *disjunktné* zjednotenie

$$\mathcal{W}_{i,j}(\mathcal{G}) = \bigcup_{t \in \mathbb{N}} \mathcal{W}_{i,j}(\mathcal{G}, t),$$

určujúce zovšeobecnený rozklad množiny $\mathcal{W}_{i,j}(\mathcal{G})$, postupným použitím (1.7), (1.8) a podmienky (ii) definície 1.4.2 dostávame

$$a_{i,j}[*] = \sum_{t \in \mathbb{N}} a_{i,j}[t] = \sum_{t \in \mathbb{N}} \sum_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G}, t)} \varphi(\gamma) = \sum_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G})} \varphi(\gamma),$$

čo bolo treba dokázať. □

1.5 Výpočet iterácie matice nad spočítateľne úplným polokruhom

Videli sme niekoľko príkladov úloh na ohodnotených grafoch, ktoré možno interpretovať ako hľadanie iterácie štvorcovej matice nad vhodným spočítateľne úplným polokruhom; to sa javí ako dobrý základ metódy riešenia týchto grafových úloh. Doposiaľ sme ale s iteráciou matice pracovali len ako s maticou nekonečných súčtov, čo je pohľad, ktorý neposkytuje žiaden návod ako iteráciu danej matice aj skutočne vypočítať. Nájdením efektívneho algoritmu na výpočet iterácie matice je pritom očividne podmienená praktická použiteľnosť celého prístupu.

Opíšeme teraz algoritmus založený na dynamickom programovaní, ktorý realizuje výpočet iterácie matice nad spočítateľne úplným polokruhom S s využitím elementárnych operácií (skalárneho) sčítania, násobenia a iterácie prvkov polokruhu S . Reálne teda možno algoritmus efektívne implementovať len za predpokladu, že sú efektívne implementovateľné tieto elementárne operácie;¹⁶ pre vstupnú maticu typu $n \times n$ sa pritom vykoná $\Theta(n^3)$ takýchto elementárnych operácií.

Algoritmus 1.5.1 (Výpočet iterácie štvorcovej matice nad spočítateľne úplným polokruhom).

Vstup: Štvorcová matica $A = (a_{i,j})_{n \times n}$ nad spočítateľne úplným polokruhom $(S, +, \cdot, 0, 1)$.

Výstup: Štvorcová matica $R = (r_{i,j})_{n \times n}$ nad $(S, +, \cdot, 0, 1)$.

1. Pre $i, j = 1, \dots, n$ polož $r_{i,j} := a_{i,j}$.
2. Opakuj pre $k = 1, \dots, n$:
 - 2.1. Pre $i, j = 1, \dots, n$ také, že $i \neq k$ a $j \neq k$ polož $r_{i,j} := r_{i,j} + r_{i,k} \cdot r_{k,j}^*$.
 - 2.2. Pre $i = 1, \dots, n$ také, že $i \neq k$ polož $r_{i,k} := r_{i,k} + r_{i,k} \cdot r_{k,k}^*$.
 - 2.3. Pre $j = 1, \dots, n$ také, že $j \neq k$ polož $r_{k,j} := r_{k,j} + r_{k,k} \cdot r_{k,j}^*$.
 - 2.4. Polož $r_{k,k} := r_{k,k} + r_{k,k} \cdot r_{k,k}^*$.
3. Pre $i = 1, \dots, n$ polož $r_{i,i} := r_{i,i} + 1$.

Poznámka 1.5.2. Algoritmus 1.5.1 je zovšeobecnením dvoch známych algoritmov nad konkrétnymi polokruhmi: nad tropickým polokruhom $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ resp. $(\mathbb{R}_{\geq 0} \cup \{\infty\}, \min, +, \infty, 0)$ ide o *Floydov-Warshallov algoritmus* na hľadanie cien najlacnejších ciest v grafe¹⁷ a nad polokruhmi jazykov $(2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\})$ pre vstupné matice nad $2^{\Sigma \cup \{\varepsilon\}}$ v podstate o *McNaughtonov-Jamadov algoritmus* konštrukcie racionálneho (regulárneho) výrazu k danému konečnému automatu.¹⁸

Poznámka 1.5.3. Hoci tropický polokruh $(\mathbb{Z} \cup \{\infty\}, \min, +, \infty, 0)$ nie je spočítateľne úplný, možno z neho úplný polokruh získať pridaním najmenšieho prvku $-\infty$ s $-\infty + \infty = \infty$ a $-\infty + a = -\infty$ pre všetky $a \in \mathbb{Z} \cup \{-\infty\}$. Aj pre ohodnotený grafy nad týmto polokruhom $(\mathbb{Z} \cup \{\pm\infty\}, \min, +, \infty, 0)$ potom možno aplikovať algoritmus 1.5.1, ktorý realizuje hľadanie infím cien *sledov* medzi jednotlivými dvojicami vrcholov. Nejde však v pravom slova zmysle o zovšeobecnenie Floydovho-Warshallovho algoritmu; nájdené infimá cien *sledov* sa už navyše nemusia rovnať cenám najlacnejších *ciest* medzi príslušnými dvojicami vrcholov. Pri Floydovom-Warshallovom algoritme sa koniec koncov obvykle predpokladá neexistencia orientovaných kružníc so zápornou cenou – hľadanie cien najlacnejších *ciest* vo všeobecnom grafe s celočíselnými ohodnoteniami hrán je **NP**-ťažký problém.

¹⁶Pre mnoho spočítateľne úplných polokruhov, ako napríklad $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ alebo $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$, ide o pomerne prirodzený predpoklad. Napríklad pri polokruhoch formálnych jazykov však vzniká problém s reprezentáciou nekonečných jazykov, ktoré môžu vzniknúť aj ako iterácie konečných jazykov. Nad takýmito polokruhmi ale možno algoritmus interpretovať ako konštrukciu racionálnych (regulárnych) výrazov pre jednotlivé prvky matice A^* . K tomuto pohľadu na vec sa ešte vrátíme v oddiele 1.8.

¹⁷Nad týmito polokruhmi je totiž vždy $r_{k,k}^* = 0$ – to je neutrálny prvok vzhľadom na násobenie v tropickom polokruhu – a teda možno tieto činitele z jednotlivých priradení vypustiť. Algoritmus hľadá ceny najlacnejších *sledov*, ktoré sa ale nad tropickými polokruhmi nezáporných čísel vždy rovnajú cenám najlacnejších *ciest*.

¹⁸Stačí namiesto operácií na 2^{Σ^*} vykonávať operácie na racionálnych výrazoch a na záver vziať súčet výrazov na tých pozíciách výslednej matice, ktorých riadky zodpovedajú počiatočným a stĺpce koncovým stavom automatu.

Intuitívne je správnosť algoritmu 1.5.1 založená na vete 1.4.6 a na platnosti nasledujúceho invariantu: po k opakovaníach cyklu v kroku 2 je pre $k = 0, \dots, n$ premenná $r_{i,j}$ rovná súčtu ohodnotení všetkých sledov v grafe $\text{Gr}(A)$ vedúcich z vrcholu i do vrcholu j , ktoré sú *nenulovej dĺžky* a súčasne všetky vrcholy sledu okrem prvého a posledného – budeme im hovoriť *vnútorné vrcholy* sledu – patria do množiny $[k]$. Pre $k = 1, \dots, n$ je pritom každý neprázdny sled z i do j s vnútornými vrcholmi v $[k]$ buď sledom s vnútornými vrcholmi v $[k-1]$, alebo ho možno jednoznačne „rozložiť“ na neprázdny sled z i do k s vnútornými vrcholmi v $[k-1]$, niekoľko neprázdnych sledov z k do k s vnútornými vrcholmi v $[k-1]$ a neprázdny sled z k do j s vnútornými vrcholmi v $[k-1]$. Ak teda označíme hodnotu premennej $r_{i,j}$ po k opakovaníach cyklu z kroku 2 ako $r_{i,j}^{(k)}$, je

$$r_{i,j}^{(k)} = r_{i,j}^{(k-1)} + r_{i,k}^{(k-1)} \cdot \left(r_{k,k}^{(k-1)} \right)^* \cdot r_{k,j}^{(k-1)}.$$

V algoritme 1.5.1 sú priradenia hodnôt do premenných $r_{i,j}$ rozdelené pre každé k do štyroch po sebe idúcich fáz 2.1 až 2.4. Tým sa zaručí, že hodnoty na pravých stranách priradení sú ešte skutočne hodnotami po $k-1$ opakovaníach cyklu a nie už hodnotami prepísanými v jeho k -tom opakovaní – priradenia tak možno realizovať *in situ* v jedinej matici R . Výrazy na pravých stranách priradení v krokoch 2.2 až 2.4 možno zjednodušiť; neurobili sme tak len pre väčšiu názornosť.

Po poslednom opakovaní cyklu v kroku 2 je teda každá premenná $r_{i,j}$ rovná súčtu ohodnotení neprázdnych sledov z i do j . V kroku 3 sa už len pripočíta ohodnotenie prípadného prázdneho sledu z i do j , ktorý existuje iba ak $i = j$. Premenná $r_{i,j}$ je už potom rovná súčtu ohodnotení *všetkých* sledov z i do j , ktorý je podľa vety 1.4.6 rovný prvku na pozícii $[i, j]$ matice A^* .

Pomerne komplikovane vyzerajúci dôkaz nasledujúceho tvrdenia je v skutočnosti len priamočiarou formalizáciou tejto jednoduchej argumentácie.

Tvrdenie 1.5.4. *Algoritmus 1.5.1 sa pre každú vstupnú maticu $A = (a_{i,j})_{n \times n}$ zastaví a na výstupe vráti maticu $R = (r_{i,j})_{n \times n}$, pre ktorú platí $R = A^*$. Počas behu algoritmu sa vykoná $\Theta(n^3)$ elementárnych operácií sčítania, násobenia a iterácie prvkov spočítateľne úplného polokruhu S .*

Dôkaz. Zastavenie algoritmu pri vykonaní $\Theta(n^3)$ elementárnych operácií v každom jeho behu možno považovať za zrejmé. Zostáva teda dokázať správnosť algoritmu 1.5.1.

Nech $\mathcal{G} = \text{Gr}(A)$ je graf matice $A = (a_{i,j})_{n \times n}$ a nech $\mathcal{G} = ([n], \varphi)$. Vnútorným vrcholom sledu $\gamma = (u_0, u_1, \dots, u_{t-1}, u_t)$ v grafe $\mathcal{G}$ nazveme ľubovoľný vrchol z množiny

$$I(\gamma) := \{u_1, \dots, u_{t-1}\}.$$

Pre $i, j = 1, \dots, n$ a $k = 0, \dots, n$ označme ako $\mathcal{U}_{i,j}^{(k)}(\mathcal{G})$ množinu všetkých sledov $\gamma \in \mathcal{W}_{i,j}(\mathcal{G})$ takých, že $|\gamma| > 0$ a $I(\gamma) \subseteq [k]$ a ako $r_{i,j}^{(k)}$ hodnotu premennej $r_{i,j}$ po k vykonaniach cyklu z kroku 2.

Zrejme

$$\mathcal{W}_{i,j}(\mathcal{G}) = \begin{cases} \mathcal{U}_{i,j}^{(n)}(\mathcal{G}) \cup \{(i)\} & \text{ak } i = j, \\ \mathcal{U}_{i,j}^{(n)}(\mathcal{G}) & \text{ak } i \neq j. \end{cases}$$

Keďže navyše pre všetky sledy γ nulovej dĺžky – čiže sledy $\gamma = (i)$ s $i \in V(\mathcal{G})$ – je $\varphi(\gamma) = 1$, zjavne stačí dokázať, že pre $i, j = 1, \dots, n$ a $k = 0, \dots, n$ je

$$r_{i,j}^{(k)} = \sum_{\gamma \in \mathcal{U}_{i,j}^{(k)}(\mathcal{G})} \varphi(\gamma). \quad (1.9)$$

Po vykonaní kroku 3 sa už totiž stačí odvolať na vetu 1.4.6.

Matematickou indukciou vzhľadom na k teda dokážeme vzťah (1.9). Pre $k = 0$ je tvrdenie zrejmé, pretože sledom nenulovej dĺžky z i do j s prázdnu množinou vnútorných vrcholov je pre $\varphi(i, j) \neq 0$ iba sled (i, j) a pre $\varphi(i, j) = 0$ takýto sled neexistuje. V oboch prípadoch je súčet ohodnotení týchto sledov rovný $a_{i,j} = r_{i,j}^{(0)}$. Nech teda tvrdenie platí pre $k = s < n$. Ukážeme, že platí aj pre $k = s + 1$.

Množinu $\mathcal{U}_{i,j}^{(s+1)}(\mathcal{G})$ možno vyjadriť ako *disjunktné* zjednotenie

$$\mathcal{U}_{i,j}^{(s+1)}(\mathcal{G}) = \mathcal{U}_{i,j}^{(s)}(\mathcal{G}) \cup \mathcal{V}_{i,j}^{(s+1)}(\mathcal{G}), \quad (1.10)$$

kde $\mathcal{V}_{i,j}^{(s+1)}(\mathcal{G})$ je množina tých sledov $\gamma \in \mathcal{U}_{i,j}^{(s+1)}(\mathcal{G})$, pre ktoré $s+1 \in I(\gamma)$. Množinu $\mathcal{V}_{i,j}^{(s+1)}(\mathcal{G})$ možno ďalej vyjadriť ako spočítateľné *disjunktné* zjednotenie

$$\mathcal{V}_{i,j}^{(s+1)}(\mathcal{G}) = \bigcup_{r \in \mathbb{N} \setminus \{0\}} \mathcal{V}_{i,j}^{(s+1,r)}(\mathcal{G}), \quad (1.11)$$

kde množina $\mathcal{V}_{i,j}^{(s+1,r)}(\mathcal{G})$ obsahuje tie sledy $\gamma = (u_0, u_1, \dots, u_{t-1}, u_t) \in \mathcal{V}_{i,j}^{(s+1)}(\mathcal{G})$, ktoré obsahujú práve r výskytov $s+1$ ako vnútorného vrcholu – čiže $|\{\ell \in \{1, \dots, t-1\} \mid u_\ell = s+1\}| = r$.

Rozdeľme každý sled z $\mathcal{V}_{i,j}^{(s+1,r)}(\mathcal{G})$ na sled z $\mathcal{U}_{i,s+1}^{(s)}(\mathcal{G})$, práve $r-1$ sledov z $\mathcal{U}_{s+1,s+1}^{(s)}(\mathcal{G})$ a sled z $\mathcal{U}_{s+1,j}^{(s)}(\mathcal{G})$; evidentne tak definujeme bijekciu. S použitím definície ohodnotenia sledu, tvrdenia 1.4.4 a „nekonečných verzií“ asociatívnosti, komutatívnosti a distributívnosti v spočítateľne úplnom polokruhu dostávame

$$\sum_{\gamma \in \mathcal{V}_{i,j}^{(s+1,r)}(\mathcal{G})} \varphi(\gamma) = \left(\sum_{\gamma \in \mathcal{U}_{i,s+1}^{(s)}(\mathcal{G})} \varphi(\gamma) \right) \cdot \left(\sum_{\gamma \in \mathcal{U}_{s+1,s+1}^{(s)}(\mathcal{G})} \varphi(\gamma) \right)^{r-1} \cdot \left(\sum_{\gamma \in \mathcal{U}_{s+1,j}^{(s)}(\mathcal{G})} \varphi(\gamma) \right). \quad (1.12)$$

Vďaka poradiu aplikácií jednotlivých priradení v krokoch 2.1 až 2.4 ďalej platí

$$r_{i,j}^{(s+1)} = r_{i,j}^{(s)} + r_{i,s+1}^{(s)} \cdot \left(r_{s+1,s+1}^{(s)} \right)^* \cdot r_{s+1,j}^{(s)},$$

z čoho vďaka indukčnému predpokladu vyplýva

$$r_{i,j}^{(s+1)} = \sum_{\gamma \in \mathcal{U}_{i,j}^{(s)}(\mathcal{G})} \varphi(\gamma) + \left(\sum_{\gamma \in \mathcal{U}_{i,s+1}^{(s)}(\mathcal{G})} \varphi(\gamma) \right) \cdot \left(\sum_{\gamma \in \mathcal{U}_{s+1,s+1}^{(s)}(\mathcal{G})} \varphi(\gamma) \right)^* \cdot \left(\sum_{\gamma \in \mathcal{U}_{s+1,j}^{(s)}(\mathcal{G})} \varphi(\gamma) \right).$$

To možno s využitím podmienky (iii) definície 1.4.2 upraviť na

$$r_{i,j}^{(s+1)} = \sum_{\gamma \in \mathcal{U}_{i,j}^{(s)}(\mathcal{G})} \varphi(\gamma) + \sum_{r \in \mathbb{N} \setminus \{0\}} \left[\left(\sum_{\gamma \in \mathcal{U}_{i,s+1}^{(s)}(\mathcal{G})} \varphi(\gamma) \right) \cdot \left(\sum_{\gamma \in \mathcal{U}_{s+1,s+1}^{(s)}(\mathcal{G})} \varphi(\gamma) \right)^{r-1} \cdot \left(\sum_{\gamma \in \mathcal{U}_{s+1,j}^{(s)}(\mathcal{G})} \varphi(\gamma) \right) \right]$$

a vďaka rovnostiam (1.10), (1.11) a (1.12) nakoniec z podmienky (ii) definície 1.4.2 dostávame

$$r_{i,j}^{(s+1)} = \sum_{\gamma \in \mathcal{U}_{i,j}^{(s+1)}(\mathcal{G})} \varphi(\gamma),$$

čo bolo treba dokázať. □

Neskôr sa nám pri dôkazoch zídu vzorce vyjadrujúce iteráciu blokovej matice nad spočítateľne úplným polokruhom pomocou iterácií jej blokov. Výsledky tohto typu navyše môžu slúžiť aj ako základ pre alternatívny algoritmus výpočtu iterácie matice, založený namiesto na dynamickom programovaní na metóde *divide et impera*. Zamerajme sa najprv na blokové matice typu 2×2 .

Lema 1.5.5. *Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad S . Nech*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

je blokový rozklad matice A , kde $A_{1,1} = (a_{i,j})_{n_1 \times n_1}$, $A_{1,2} = (a_{i,n_1+j})_{n_1 \times n_2}$, $A_{2,1} = (a_{n_1+i,j})_{n_2 \times n_1}$ a $A_{2,2} = (a_{n_1+i,n_1+j})_{n_2 \times n_2}$ pre nejaké $n_1, n_2 \in \mathbb{N}$. Potom

$$A^* = \begin{pmatrix} A_{1,1}[*] & A_{1,2}[*] \\ A_{2,1}[*] & A_{2,2}[*] \end{pmatrix},$$

pričom jednotlivé bloky $A_{1,1}[] = (a_{i,j}[*])_{n_1 \times n_1}$, $A_{1,2}[*] = (a_{i,n_1+j}[*])_{n_1 \times n_2}$, $A_{2,1}[*] = (a_{n_1+i,j}[*])_{n_2 \times n_1}$ a $A_{2,2}[*] = (a_{n_1+i,n_1+j}[*])_{n_2 \times n_2}$ sú dané nasledujúcimi vzťahmi:*

$$\begin{aligned} A_{1,1}[*] &= (A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^*, \\ A_{1,2}[*] &= (A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^*A_{1,2}A_{2,2}^*, \\ A_{2,1}[*] &= (A_{2,2} + A_{2,1}A_{1,1}^*A_{1,2})^*A_{2,1}A_{1,1}^*, \\ A_{2,2}[*] &= (A_{2,2} + A_{2,1}A_{1,1}^*A_{1,2})^*. \end{aligned}$$

Dôkaz. Dokážeme, že $A_{1,1}[*] = (A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^*$; pre zvyšné bloky matice A^* možno postupovať analogicky. Nech $\mathcal{G} := \text{Gr}(A) = ([n], \varphi)$, $V_1 = \{1, \dots, n_1\}$ a $V_2 = \{n_1 + 1, \dots, n\}$.

Intuitívne dokazovaná rovnosť vyplýva z možnosti rozložiť každý sled v grafe $\mathcal{G}$ z vrcholu množiny V_1 do vrcholu množiny V_1 na nejaký počet sledov vedúcich z V_1 do V_1 , ktoré sa do vrcholu z V_1 dostanú iba na svojom začiatku a konci. Takýmito sledmi môžu byť sledy (i, j) , kde i aj j patria do V_1 (tým prislúcha blok $A_{1,1}$), alebo sledy typu $(i, k_0, k_1, \dots, k_t, j)$, kde $i, j \in V_1$, $t \in \mathbb{N}$ a $k_0, \dots, k_t \in V_2$ (týmto prislúcha blok $A_{1,2}A_{2,2}^*A_{2,1}$).

Formálne: zaveďme označenia

$$\begin{aligned} A_{2,2}^* &= (b_{i,j})_{n_2 \times n_2}, \\ A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1} &= (c_{i,j})_{n_1 \times n_1}, \\ (A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^s &= (c_{i,j}[s])_{n_1 \times n_1} \quad \text{pre všetky } s \in \mathbb{N}, \\ (A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^* &= (c_{i,j}[*])_{n_1 \times n_1}. \end{aligned}$$

Dokazovanú rovnosť teda možno vyjadriť aj ako $a_{i,j}[*] = c_{i,j}[*]$ pre všetky $i, j \in V_1$.

Z vety 1.4.6 pre všetky $i, j \in V_1$ vyplýva

$$a_{i,j}[*] = \sum_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G})} \varphi(\gamma).$$

Nech $\mathcal{X}_{i,j}(\mathcal{G})$ pozostáva zo sledov *nenulovej dĺžky* v $\mathcal{W}_{i,j}(\mathcal{G})$. Množinu $\mathcal{X}_{i,j}(\mathcal{G})$ potom možno vyjadriť ako *disjunktné* zjednotenie

$$\mathcal{X}_{i,j}(\mathcal{G}) = \bigcup_{r \in \mathbb{N}} \mathcal{X}_{i,j}^{(r)}(\mathcal{G}), \quad (1.13)$$

kde $\mathcal{X}_{i,j}^{(r)}(\mathcal{G})$ pozostáva z tých sledov v $\mathcal{X}_{i,j}(\mathcal{G})$, ktoré obsahujú práve $r + 2$ výskytov vrcholov z V_1 (čiže práve r výskytov vrcholov z V_1 vo vnútri sledu).

Každý sled $\gamma \in \mathcal{X}_{i,j}^{(0)}(\mathcal{G})$ je daný buď ako $\gamma = (i, j)$, alebo ako $\gamma = (i, k_0, \dots, k_t, j)$ pre nejaké $t \in \mathbb{N}$ a $k_0, \dots, k_t \in V_2$. V prvom prípade je pritom $\varphi(\gamma) = \varphi(i, j) = a_{i,j}$, kým v druhom prípade dostávame $\varphi(\gamma) = \varphi(i, k_0) \cdot \varphi(\beta) \cdot \varphi(k_t, j) = a_{i,k_0} \cdot \varphi(\beta) \cdot a_{k_t,j}$ pre nejaký sled β z vrcholu k_0 do vrcholu k_t v podgrafe

$\mathcal{G}_2 := (V_2, \varphi|_{V_2})$ grafu $\mathcal{G}$.¹⁹ Z vlastností spočítateľne úplných polokruhov a vety 1.4.6 preto vyplýva

$$\begin{aligned} \sum_{\gamma \in \mathcal{X}_{i,j}^{(0)}(\mathcal{G})} \varphi(\gamma) &= a_{i,j} + \sum_{k_0, k_t \in V_2} a_{i,k_0} \cdot \left(\sum_{\beta \in \mathcal{W}_{k_0, k_t}(\mathcal{G}_2)} \varphi(\beta) \right) \cdot a_{k_t, j} = \\ &= a_{i,j} + \sum_{k_0, k_t \in V_2} a_{i,k_0} \cdot b_{k_0, k_t} \cdot a_{k_t, j} = c_{i,j} = c_{i,j}[1]. \end{aligned} \quad (1.14)$$

Pre každé $r \in \mathbb{N} \setminus \{0\}$ navyše existuje bijekcia medzi množinami $\mathcal{X}_{i,j}^{(r)}(\mathcal{G})$ a

$$\bigcup_{\ell_1, \dots, \ell_r \in V_1} \mathcal{X}_{i, \ell_1}^{(0)}(\mathcal{G}) \times \mathcal{X}_{\ell_1, \ell_2}^{(0)}(\mathcal{G}) \times \dots \times \mathcal{X}_{\ell_{r-1}, \ell_r}^{(0)}(\mathcal{G}) \times \mathcal{X}_{\ell_r, j}^{(0)}(\mathcal{G}),$$

ktorá každý sled $\gamma \in \mathcal{X}_{i,j}^{(r)}(\mathcal{G})$ rozdelí na $r+1$ nadväzujúcich sledov postupne náležiacich do množín $\mathcal{X}_{i, \ell_1}^{(0)}(\mathcal{G}), \mathcal{X}_{\ell_1, \ell_2}^{(0)}(\mathcal{G}), \dots, \mathcal{X}_{\ell_{r-1}, \ell_r}^{(0)}(\mathcal{G}), \mathcal{X}_{\ell_r, j}^{(0)}(\mathcal{G})$ pre nejaké $\ell_1, \dots, \ell_r \in V_1$. Keďže ide o bijekciu, z vlastností spočítateľne úplných polokruhov a (1.14) vyplýva

$$\sum_{\gamma \in \mathcal{X}_{i,j}^{(r)}(\mathcal{G})} \varphi(\gamma) = \sum_{\ell_1, \dots, \ell_r \in V_1} \left[\left(\sum_{\gamma \in \mathcal{X}_{i, \ell_1}^{(0)}(\mathcal{G})} \varphi(\gamma) \right) \dots \left(\sum_{\gamma \in \mathcal{X}_{\ell_r, j}^{(0)}(\mathcal{G})} \varphi(\gamma) \right) \right] = c_{i,j}[r+1].$$

Vďaka (1.13) potom dostávame

$$\sum_{\gamma \in \mathcal{X}_{i,j}(\mathcal{G})} \varphi(\gamma) = \sum_{r \in \mathbb{N}} \sum_{\gamma \in \mathcal{X}_{i,j}^{(r)}(\mathcal{G})} \varphi(\gamma) = \sum_{r \in \mathbb{N}} c_{i,j}[r+1]$$

a s použitím Kroneckerovej delty²⁰ tak môžeme písať

$$a_{i,j}[*] = \sum_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G})} \varphi(\gamma) = \delta_{i,j} + \sum_{\gamma \in \mathcal{X}_{i,j}(\mathcal{G})} \varphi(\gamma) = \delta_{i,j} + \sum_{r \in \mathbb{N}} c_{i,j}[r+1] = c_{i,j}[*],$$

čo bolo treba dokázať. □

S použitím práve dokázanej lemy možno pomerne ľahko odvodiť vzorce pre iterácie trojuholníkových blokových matíc: v nasledujúcom tak urobíme pre matice o 2×2 a 3×3 blokoch. Blokové trojuholníkové matice často vznikajú pri konštrukciách automatov, pričom vzorce pre ich iterácie sú o niečo jednoduchšie, než pre všeobecné blokové matice. Ako $\mathbf{0}_{m \times n}$ označíme pre všetky $m, n \in \mathbb{N}$ nulovú maticu $(0)_{m \times n}$; ak sú rozmery nulovej matice zrejmé z kontextu, píšeme namiesto $\mathbf{0}_{m \times n}$ len $\mathbf{0}$.

Dôsledok 1.5.6. *Nech S je spočítateľne úplný polokruh, $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad S s blokovým rozkladom*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ \mathbf{0} & A_{2,2} \end{pmatrix},$$

kde $A_{1,1} = (a_{i,j})_{n_1 \times n_1}$, $A_{1,2} = (a_{i, n_1+j})_{n_1 \times n_2}$ a $A_{2,2} = (a_{n_1+i, n_1+j})_{n_2 \times n_2}$ pre nejaké $n_1, n_2 \in \mathbb{N}$. Potom

$$A^* = \begin{pmatrix} A_{1,1}^* & A_{1,1}^* A_{1,2} A_{2,2}^* \\ \mathbf{0} & A_{2,2}^* \end{pmatrix}.$$

Dôkaz. Stačí použiť lemu 1.5.5 pre $A_{2,1} = \mathbf{0}_{n_2 \times n_1}$. □

¹⁹Symbol $\varphi|_{V_2}$ tu označuje zúženie zobrazenia φ na obor V_2 .

²⁰Pre $i = j$ je $\delta_{i,j} = 1$; inak $\delta_{i,j} = 0$.

Dôsledok 1.5.7. *Nech S je spočítateľne úplný polokruh, $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad S s blokovým rozkladom*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ \mathbf{0} & A_{2,2} & A_{2,3} \\ \mathbf{0} & \mathbf{0} & A_{3,3} \end{pmatrix},$$

kde $A_{1,1} = (a_{i,j})_{n_1 \times n_1}$, $A_{1,2} = (a_{i,n_1+j})_{n_1 \times n_2}$, $A_{1,3} = (a_{i,n_1+n_2+j})_{n_1 \times n_3}$, $A_{2,2} = (a_{n_1+i,n_1+j})_{n_2 \times n_2}$, $A_{2,3} = (a_{n_1+i,n_1+n_2+j})_{n_2 \times n_3}$ a $A_{3,3} = (a_{n_1+n_2+i,n_1+n_2+j})_{n_3 \times n_3}$ pre nejaké $n_1, n_2, n_3 \in \mathbb{N}$. Potom

$$A^* = \begin{pmatrix} A_{1,1}^* & A_{1,1}^* A_{1,2} A_{2,2}^* & A_{1,1}^* A_{1,3} A_{3,3}^* + A_{1,1}^* A_{1,2} A_{2,2}^* A_{2,3} A_{3,3}^* \\ \mathbf{0} & A_{2,2}^* & A_{2,2}^* A_{2,3} A_{3,3}^* \\ \mathbf{0} & \mathbf{0} & A_{3,3}^* \end{pmatrix}.$$

Dôkaz. Nech

$$B_{1,1} = \begin{pmatrix} A_{1,1} & A_{1,2} \\ \mathbf{0} & A_{2,2} \end{pmatrix} \quad \text{a} \quad B_{1,3} = \begin{pmatrix} A_{1,3} \\ A_{2,3} \end{pmatrix}.$$

Potom

$$A = \begin{pmatrix} B_{1,1} & B_{1,3} \\ \mathbf{0} & A_{3,3} \end{pmatrix}$$

a z dôsledku 1.5.6 dostávame

$$A^* = \begin{pmatrix} B_{1,1}^* & B_{1,1}^* B_{1,3} A_{3,3}^* \\ \mathbf{0} & A_{3,3}^* \end{pmatrix}. \quad (1.15)$$

Z dôsledku 1.5.6 ale tiež vyplýva

$$B_{1,1}^* = \begin{pmatrix} A_{1,1}^* & A_{1,1}^* A_{1,2} A_{2,2}^* \\ \mathbf{0} & A_{2,2}^* \end{pmatrix},$$

z čoho

$$B_{1,1}^* B_{1,3} = \begin{pmatrix} A_{1,1}^* A_{1,3} + A_{1,1}^* A_{1,2} A_{2,2}^* A_{2,3} \\ A_{2,2}^* A_{2,3} \end{pmatrix}.$$

Dosadením do (1.15) potom dostávame

$$A^* = \begin{pmatrix} A_{1,1}^* & A_{1,1}^* A_{1,2} A_{2,2}^* & A_{1,1}^* A_{1,3} A_{3,3}^* + A_{1,1}^* A_{1,2} A_{2,2}^* A_{2,3} A_{3,3}^* \\ \mathbf{0} & A_{2,2}^* & A_{2,2}^* A_{2,3} A_{3,3}^* \\ \mathbf{0} & \mathbf{0} & A_{3,3}^* \end{pmatrix},$$

čo bolo treba dokázať. □

1.6 Konečné automaty nad polokruhmi

V príklade 1.2.2 sme videli, že nedeterministické konečné automaty nad abecedou Σ možno chápať ako špeciálne ohodnotené grafy nad polokruhom formálnych jazykov $(2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\})$ obohatené o trochu dodatočnej informácie. Ohodnotením prípadnej hrany (prechodu) z vrcholu (stavu) p do vrcholu (stavu) q je tu jazyk pozostávajúci z práve všetkých písmen – a prípadne aj prázdneho slova – pomocou ktorých možno medzi týmito dvoma stavmi prejsť „na jeden krok“. Konečnému automatu teda prislúcha ohodnotený graf nad polokruhom 2^{Σ^*} taký, že všetky ohodnotenia hrán sú prvkami $2^{\Sigma \cup \{\varepsilon\}}$. Onou dodatočnou informáciou sú dané množiny počiatočných a koncových stavov automatu. Pod behom automatu môžeme rozumieť sled v práve opísanom ohodnotenom grafe – ak, pravda, pamätáme na to, že pri našej interpretácii nepovoľujeme násobné prechody medzi stavmi a namiesto toho pracujeme s prechodmi ohodnotenými konečnými jazykmi z $2^{\Sigma \cup \{\varepsilon\}}$. Každému takto chápanému behu zodpovedá jazyk,

ktorý získame postupným zreťazení jazykov jednotlivých jeho prechodov. Jazyk rozoznávaný automatom $\mathcal{A}$ potom môžeme definovať ako (konečné alebo spočítateľne nekonečné) zjednotenie jazykov prislúchajúcich behom automatu $\mathcal{A}$ vedúcim z počiatočného do koncového stavu.

To isté možno urobiť nad ľubovoľným spočítateľne úplným polokruhom S : namiesto automatov s prechodmi ohodnotenými jazykmi z $2^{\Sigma \cup \{\varepsilon\}}$ možno uvažovať automaty s prechodmi ohodnotenými prvkami nejakej podmnožiny polokruhu S . Každý beh z počiatočného do koncového stavu potom tiež realizuje prvok polokruhu S získaný postupným prenášobením ohodnotení jeho prechodov. Konečným alebo spočítateľne nekonečným súčtom cez všetky takéto behy – v spočítateľne úplnom polokruhu je takýto súčet dobre definovaný – napokon získame prvok polokruhu S realizovaný daným konečným automatom.

Týmito úvahami sa postupne dostávame k pojmu *konečného S' -automatu nad spočítateľne úplným polokruhom S* , kde $S' \subseteq S$ je množina možných ohodnotení prechodov (hrán) obsahujúca prvky 0 a 1.²¹ V skutočnosti ale naša definícia bude o niečo všeobecnejšia: ku každému stavu q priradíme prvok $\iota(q)$ množiny S' udávajúci jeho počiatočné ohodnotenie a prvok $\tau(q)$ udávajúci jeho koncové ohodnotenie. Prvok polokruhu S realizovaný behom zo stavu p do stavu q je potom ešte prenášobený zľava prvkom $\iota(p)$ a sprava prvkom $\tau(q)$, ako keby boli prvkami polokruhu ohodnotené aj „polovičné šípky“ vedúce do počiatočných a z koncových stavov. Vyššie opísaný model s množinami počiatočných a koncových stavov tak dostávame v prípade, že sa všetky počiatočné aj koncové ohodnotenia stavov automatu rovnajú niektorému z prvkov 0 alebo 1 – neskôr dokážeme, že ide o normálny tvar.

Z dôvodu, ktorý bude zrejmý až neskôr, definujeme práve opísaný model po formálnej stránke nielen pre spočítateľne úplné polokruhy, ale pre *ľubovoľný* polokruh S . Prvok polokruhu realizovaný automatom však teraz definujeme iba v prípade, že je polokruh S spočítateľne úplný.

Definícia 1.6.1. Nech S je polokruh a $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 polokruhu S . *Konečný S' -automat nad S* je štvorica $\mathcal{A} = (Q, \varphi, \iota, \tau)$, kde Q je konečná množina stavov, $\varphi: Q^2 \rightarrow S'$ je zobrazenie určujúce ohodnotenia prechodov, $\iota: Q \rightarrow S'$ je zobrazenie určujúce počiatočné ohodnotenia stavov a $\tau: Q \rightarrow S'$ je zobrazenie určujúce koncové ohodnotenia stavov.

Grafom automatu $\mathcal{A} = (Q, \varphi, \iota, \tau)$ nazveme ohodnotený graf $\mathcal{G}[\mathcal{A}] = (Q, \varphi)$ nad S . *Behom* automatu $\mathcal{A}$ zo stavu p do stavu q nazveme ľubovoľný sled γ z vrcholu p do vrcholu q v grafe $\mathcal{G}[\mathcal{A}]$, čo nám umožňuje hovoriť aj o dĺžke $|\gamma|$ a ohodnotení $\varphi(\gamma)$ behu γ . *Prvok polokruhu S realizovaný behom γ zo stavu p do stavu q je potom daný ako*

$$\|\gamma\| := \iota(p)\varphi(\gamma)\tau(q). \quad (1.16)$$

Množinu všetkých behov automatu $\mathcal{A}$ označíme $\mathcal{R}(\mathcal{A})$. Ak je navyše polokruh S spočítateľne úplný, definujeme *prvok polokruhu S realizovaný automatom $\mathcal{A}$* ako

$$\|\mathcal{A}\| := \sum_{\gamma \in \mathcal{R}(\mathcal{A})} \|\gamma\|. \quad (1.17)$$

Príklad 1.6.2. Nedeterministické konečné automaty nad abecedou Σ , tak ako boli definované v rámci príkladu 1.2.2, možno v zmysle vyššie uvedeného chápať aj ako konečné $2^{\Sigma \cup \{\varepsilon\}}$ -automaty nad polokruhom formálnych jazykov $(2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\})$ s počiatočnými a koncovými ohodnoteniami z $\{\emptyset, \{\varepsilon\}\}$. Podobne *konečnestavové prekladače* možno chápať aj ako konečné $2_{\text{fin}}^{\Sigma^* \times \Sigma^*}$ -automaty nad polokruhom $(2^{\Sigma^* \times \Sigma^*}, \cup, \cdot, \emptyset, \{(\varepsilon, \varepsilon)\})$, kde $2_{\text{fin}}^{\Sigma^* \times \Sigma^*}$ označuje množinu všetkých *konečných* podmnožín monoidu $\Sigma^* \times \Sigma^*$ a počiatočné a koncové ohodnotenia všetkých stavov sú alebo $\emptyset$, alebo $\{(\varepsilon, \varepsilon)\}$.

²¹Príslušnosť prvku 0 do S' je vlastne vyjadrením požiadavky, že medzi niektorými dvojicami stavov (vrcholov) nemusí viesť žiaden prechod (hrana). Prechody ohodnotené prvkom 1 až tak kľúčové nie sú. Požiadavka príslušnosti obidvoch prvkov 0, 1 do S' sa však zakrátko ukáže byť prirodzenou v súvislosti s počiatočnými a koncovými ohodnoteniami stavov.

Ak je množina stavov automatu $\mathcal{A}$ daná ako $Q = [n]$ pre nejaké $n \in \mathbb{N}$ – to možno predpokladať bez ujmy na všeobecnosti – nazveme maticu $A[\mathcal{A}] := \mathbf{A}(\mathcal{G}[\mathcal{A}])$ *prechodovou maticou* automatu $\mathcal{A}$, riadkový vektor $\mathbf{i}[\mathcal{A}] := (\iota(1), \dots, \iota(n))$ *vektorom počiatočných ohodnotení* automatu $\mathcal{A}$ a stĺpcový vektor $\mathbf{f}[\mathcal{A}] := (\tau(1), \dots, \tau(n))^T$ *vektorom koncových ohodnotení* automatu $\mathcal{A}$. Nasledujúca veta poskytuje maticovú charakterizáciu prvku polokruhu realizovaného automatom $\mathcal{A}$.

Veta 1.6.3. *Nech S je spočítateľne úplný polokruh a $S' \subseteq S$ je množina obsahujúca prvky 0 a 1. Nech $\mathcal{A} = (Q, \varphi, \iota, \tau)$ je konečný S' -automat nad S , kde $Q = [n]$ pre nejaké $n \in \mathbb{N}$, $\mathbf{i}[\mathcal{A}] = \mathbf{i}$, $A[\mathcal{A}] = A$ a $\mathbf{f}[\mathcal{A}] = \mathbf{f}$. Potom $\|\mathcal{A}\| = \mathbf{i}A\mathbf{f}$.*

Dôkaz. Vyplýva bezprostredne z (1.16), (1.17), vety 1.4.6 a podmienky (iii) definície 1.4.2. $\square$

Poznámka 1.6.4. Prechodová matica $A[\mathcal{A}] = A$ a vektory $\mathbf{i}[\mathcal{A}] = \mathbf{i}$ a $\mathbf{f}[\mathcal{A}] = \mathbf{f}$ zrejme obsahujú kompletnú informáciu o automate $\mathcal{A}$. Takýto automat s množinou stavov $[n]$ pre $n \in \mathbb{N}$ teda budeme písať aj ako štvoricu $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$.

1.7 Normálny tvar konečných automatov nad polokruhmi

Pre automaty nad spočítateľne úplnými polokruhmi teraz dokážeme vetu o normálnom tvare, ktorá je zovšeobecnením dobre známeho tvrdenia z klasickej teórie automatov a jazykov: ku každému automatu možno skonštruovať ekvivalentný, v ktorom existuje *práve jeden* stav s nenulovým počiatočným ohodnotením, ktoré je pre tento stav rovné jednotke v polokruhu a *práve jeden* stav s nenulovým koncovým ohodnotením, ktoré je pre tento stav tiež rovné jednotke v polokruhu; tieto dva stavy sú pritom rôzne. V automate navyše nevedú žiadne prechody *do* stavu s nenulovým počiatočným ohodnotením, ani *zo* stavu s nenulovým koncovým ohodnotením. Takéto automaty sa v literatúre často nazývajú *normalizované*, prípadne sa hovorí o *automatoch v normálnom tvare* [36, 94]; rovnakú terminológiu budeme používať aj my (nie je tu veľké riziko nedorozumenia, pretože žiaden ďalší normálny tvar automatov nad polokruhmi v tomto texte uvažovať nebudeme).

Definícia 1.7.1. Nech S je spočítateľne úplný polokruh a $S' \subseteq S$ je množina obsahujúca prvky 0 a 1. Konečný S' -automat $\mathcal{A} = (Q, \varphi, \iota, \tau)$ nad polokruhom S je v *normálnom tvare*, ak existujú stavy $q_i, q_f \in Q$ také, že $q_i \neq q_f$, $\iota(q_i) = \tau(q_f) = 1$, $\iota(q) = 0$ pre všetky $q \in Q \setminus \{q_i\}$, $\tau(q) = 0$ pre všetky $q \in Q \setminus \{q_f\}$ a pre všetky $p, q \in Q$ s $\varphi(p, q) \neq 0$ je $p \neq q_f$ a $q \neq q_i$.

Pomenovanie *normálny tvar* je tu odôvodnené nasledujúcou vetou, podľa ktorej pre ľubovoľný konečný S' -automat nad spočítateľne úplným polokruhom S existuje k nemu ekvivalentný konečný S' -automat nad S v normálnom tvare. Z dôkazu vety bude zrejímavý aj postup konštrukcie takéhoto ekvivalentného automatu.

Veta 1.7.2. *Nech S je spočítateľne úplný polokruh, $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $\mathcal{A}$ je konečný S' -automat nad S . Potom existuje konečný S' -automat $\mathcal{A}'$ nad S v normálnom tvare taký, že $\|\mathcal{A}'\| = \|\mathcal{A}\|$.*

Dôkaz. Konštrukcia bude spočívať v premenovaní pôvodných stavov $1, \dots, n$ na $2, \dots, n+1$ a v pridaní nového „počiatočného“ stavu 1 a nového „koncového“ stavu $n+2$. Ak mal stav i v pôvodnom automate nenulové počiatočné ohodnotenie $\iota(i)$, zo stavu 1 bude v novom automate viesť prechod s ohodnotením $\iota(i)$ do stavu $i+1$, ktorý zodpovedá pôvodnému stavu i . Podobne pre každý stav j pôvodného automatu s nenulovým $\tau(j)$ bude v novom automate viesť prechod s ohodnotením $\tau(j)$ zo stavu $j+1$ do stavu $n+2$. Následne už len stačí zabezpečiť, aby stav 1 mal počiatočné ohodnotenie 1, stav $n+2$ mal koncové ohodnotenie 1 a aby zvyšné počiatočné a koncové ohodnotenia stavov boli nulové.

Formálne: bez ujmy na všeobecnosti predpokladajme, že pre nejaké $n \in \mathbb{N}$ je $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$. Položme $\mathcal{A}' = (n+2, \mathbf{i}', A', \mathbf{f}')$, kde $\mathbf{i}' = (1, 0, \dots, 0)$, $\mathbf{f}' = (0, \dots, 0, 1)^T$ a

$$A' = \begin{pmatrix} 0 & \mathbf{i} & 0 \\ \mathbf{0} & A & \mathbf{f} \\ 0 & \mathbf{0} & 0 \end{pmatrix}.$$

Z dôsledku 1.5.7 potom vyplýva

$$(A')^* = \begin{pmatrix} 1 & \mathbf{i}A^* & \mathbf{i}A^*\mathbf{f} \\ \mathbf{0} & A^* & A^*\mathbf{f} \\ 0 & \mathbf{0} & 1 \end{pmatrix}$$

a z vety 1.6.3 následne dostávame

$$\|\mathcal{A}'\| = \mathbf{i}'(A')^*\mathbf{f}' = \mathbf{i}A^*\mathbf{f} = \|\mathcal{A}\|.$$

Je navyše očividné, že automat $\mathcal{A}'$ je v normálnom tvare – stačí vziať $q_i = 1$ a $q_f = n+2$. Tvrdenie je tým pádom dokázané. $\square$

1.8 Racionálne prvky spočítateľne úplného polokruhu

Definujeme teraz *racionálne výrazy* (matematicky menej výstižne nazývané aj *regulárnymi výrazmi*) nad spočítateľne úplným polokruhom S a dokážeme ich ekvivalenciu s konečnými automatmi nad S . Prvky polokruhu realizované racionálnymi výrazmi – alebo ekvivalentne konečnými automatmi – nazveme jeho *racionálnymi prvkami*. Pomenovanie *racionálny výraz* resp. *racionálny prvok* tu vychádza z možnej interpretácie racionálnych funkcií ako racionálnych prvkov vhodne zvoleného polokruhu; podobné úvahy však prídu na pretras až v oddiele 3.5 v súvislosti s mocninovými radmi.

Podobne ako pri konečných automatoch definujeme racionálne výrazy po syntaktickej stránke pre všetky polokruhy a na spočítateľne úplné polokruhy sa obmedzíme až pri definícii ich sémantiky.

Definícia 1.8.1 (Syntax racionálnych výrazov). Nech S je polokruh a $S' \subseteq S$ je množina obsahujúca prvky 0 a 1. *Množina S' -racionálnych výrazov* $\text{RatE}(S', S)$ nad polokruhom S je daná jazykom nad (vo všeobecnosti aj nekonečnou) abecedou $\{\mathbf{a} \mid \mathbf{a} \in S'\} \cup \{+, \cdot, *, (,)\}$ takým, že:

- (i) Jazyk $\text{RatE}(S', S)$ obsahuje pre každé $\mathbf{a} \in S'$ slovo $\mathbf{a}$.
- (ii) Pre všetky $\mathbf{E}, \mathbf{F} \in \text{RatE}(S', S)$ obsahuje jazyk $\text{RatE}(S', S)$ slová $(\mathbf{E} + \mathbf{F})$, $(\mathbf{E} \cdot \mathbf{F})$ a $(\mathbf{E}^*)$.
- (iii) Jazyk $\text{RatE}(S', S)$ neobsahuje žiadne ďalšie slová.

Slová z jazyka $\text{RatE}(S', S)$ nazývame *S' -racionálnymi výrazmi* nad S .

Definícia 1.8.2 (Sémantika racionálnych výrazov). Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh a $S' \subseteq S$ je množina obsahujúca prvky 0 a 1. *Prvok $\|\mathbf{G}\|$ realizovaný S' -racionálnym výrazom $\mathbf{G}$* nad S je daný nasledovne:

- (i) Ak $\mathbf{G} = \mathbf{a}$ pre nejaké $\mathbf{a} \in S'$, tak $\|\mathbf{G}\| = \mathbf{a}$.
- (ii) Nech $\mathbf{E}, \mathbf{F} \in \text{RatE}(S', S)$. Ak $\mathbf{G} = (\mathbf{E} + \mathbf{F})$, tak $\|\mathbf{G}\| = \|\mathbf{E}\| + \|\mathbf{F}\|$; ak $\mathbf{G} = (\mathbf{E} \cdot \mathbf{F})$, tak $\|\mathbf{G}\| = \|\mathbf{E}\| \cdot \|\mathbf{F}\|$; ak $\mathbf{G} = (\mathbf{E}^*)$, tak $\|\mathbf{G}\| = \|\mathbf{E}\|^*$.

Poznámka 1.8.3. Definícia sémantiky nám okrem iného umožňuje vynechávať v zápise racionálnych výrazov niektoré zátvorky. Keďže je napríklad symbol $+$ interpretovaný ako asociatívna operácia sčítania, možno pre ktorýkoľvek z výrazov $(E + (F + G))$ a $((E + F) + G)$ písať len $(E + F + G)$, pričom pod takýmto zápisom treba chápať jeho ľubovoľné uzátvorkovanie; podobne aj pre symbol $\cdot$ interpretovaný ako násobenie. Iterácia je navyše operácia s vyššou precedenciou ako násobenie a násobenie má vyššiu precedenciu ako sčítanie. Napríklad namiesto $((E^*) \cdot F)$ teda píšeme len $(E^* \cdot F)$ a namiesto $(E + (F \cdot G))$ píšeme len $(E + F \cdot G)$. Podobne možno vynechať aj vonkajšie zátvorky okolo celého výrazu.

Dokážeme teraz postupne ekvivalenciu S' -racionálnych výrazov nad S a konečných S' -automatov nad S . Začneme jednoduchým pozorovaním, podľa ktorého možno na konštrukciu racionálneho výrazu k danému konečnému automatu použiť algoritmus 1.5.1.

Lema 1.8.4. *Nech S je spočítateľne úplný polokruh, $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $\mathcal{A}$ je konečný S' -automat nad S . Potom existuje S' -racionálny výraz E nad S taký, že $\|E\| = \|\mathcal{A}\|$.*

Dôkaz. Bez ujmy na všeobecnosti predpokladajme, že $\mathcal{A} = (Q, \varphi, \iota, \tau)$ je v normálnom tvare, pričom $Q = [n]$ pre nejaké prirodzené číslo $n \geq 2$ a $\iota(1) = \tau(n) = 1$ (ostatné počiatočné a koncové ohodnotenia stavov sú z definície normálneho tvaru nulové). Ak potom $A = (a_{i,j})_{n \times n}$ je prechodová matica automatu $\mathcal{A}$ a $A^* = (a_{i,j}[*])$ je jej iterácia, z vety 1.6.3 vyplýva $\|\mathcal{A}\| = a_{1,n}[*]$.

Stačí už teda upraviť algoritmus 1.5.1 tak, aby namiesto vyhodnocovania súčtov, súčinov a iterácií konštruoval príslušný racionálny výraz – výstupom algoritmu je tak matica S' -racionálnych výrazov $(E_{i,j})_{n \times n}$. Matematickou indukciou je potom možné dokázať, že pre $i, j = 1, \dots, n$ je $\|E_{i,j}\| = a_{i,j}[*]$. Stačí teda položiť $E = E_{1,n}$. $\square$

Zostáva ukázať, že ku každému racionálnemu výrazu možno zostrojiť ekvivalentný konečný automat. Dôkaz tohto tvrdenia rozdelíme do nasledujúcich štyroch lem. Najprv v leme 1.8.5 ukážeme, že existujú konečné automaty ekvivalentné atomickým výrazom a pre $a \in S'$. V ďalších troch lemach potom postupne zostrojíme konečné automaty realizujúce súčet, súčin a iteráciu.

Lema 1.8.5. *Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $a \in S'$. Potom existuje konečný S' -automat $\mathcal{A}$ nad S taký, že $\|\mathcal{A}\| = a$.*

Dôkaz. Stačí vziať $\mathcal{A} = (1, \mathbf{i}, A, \mathbf{f})$, kde $\mathbf{i} = (a)$, $A = (0)$ a $\mathbf{f} = (1)^T$. Z vety 1.6.3 potom dostávame $\|\mathcal{A}\| = (a)(0)^*(1)^T = (a)(1)(1)^T = a$. $\square$

Lema 1.8.6. *Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $\mathcal{A}_1, \mathcal{A}_2$ sú konečné S' -automaty nad polokruhom S . Potom existuje konečný S' -automat $\mathcal{A}$ nad S taký, že $\|\mathcal{A}\| = \|\mathcal{A}_1\| + \|\mathcal{A}_2\|$.*

Dôkaz. Bez ujmy na všeobecnosti predpokladajme, že $\mathcal{A}_1 = (n_1, \mathbf{i}_1, A_1, \mathbf{f}_1)$ a $\mathcal{A}_2 = (n_2, \mathbf{i}_2, A_2, \mathbf{f}_2)$ pre nejaké $n_1, n_2 \in \mathbb{N}$. Položme $\mathcal{A} = (n_1 + n_2, \mathbf{i}, A, \mathbf{f})$, kde $\mathbf{i} = (\mathbf{i}_1, \mathbf{i}_2)$,

$$A = \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & A_2 \end{pmatrix}$$

a $\mathbf{f} = (\mathbf{f}_1^T, \mathbf{f}_2^T)^T$. Z vety 1.6.3 a dôsledku 1.5.6 potom vyplýva

$$\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f} = \begin{pmatrix} \mathbf{i}_1 & \mathbf{i}_2 \end{pmatrix} \begin{pmatrix} A_1^* & \mathbf{0} \\ \mathbf{0} & A_2^* \end{pmatrix} \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix} = \mathbf{i}_1 A_1^* \mathbf{f}_1 + \mathbf{i}_2 A_2^* \mathbf{f}_2 = \|\mathcal{A}_1\| + \|\mathcal{A}_2\|,$$

čo dokazuje správnosť konštrukcie automatu $\mathcal{A}$. $\square$

Lema 1.8.7. *Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $\mathcal{A}_1, \mathcal{A}_2$ sú konečné S' -automaty nad polokruhom S . Potom existuje konečný S' -automat $\mathcal{A}$ nad S taký, že $\|\mathcal{A}\| = \|\mathcal{A}_1\| \cdot \|\mathcal{A}_2\|$.*

Dôkaz. Bez ujmy na všeobecnosti opäť predpokladajme, že $\mathcal{A}_1 = (n_1, \mathbf{i}_1, A_1, \mathbf{f}_1)$ a $\mathcal{A}_2 = (n_2, \mathbf{i}_2, A_2, \mathbf{f}_2)$ pre nejaké $n_1, n_2 \in \mathbb{N}$. Predpokladajme tiež, že pre všetky stavy $i \in [n_1]$ automatu $\mathcal{A}_1$ s koncovým ohodnotením $\tau_1(i)$ a pre všetky stavy $j \in [n_2]$ automatu $\mathcal{A}_2$ s počiatočným ohodnotením $\iota_2(j)$ je $\tau_1(i) \cdot \iota_2(j) \in S'$; môžeme si to dovoliť vďaka vete 1.7.2. Automat $\mathcal{A}$ skonštruujeme „zapojením oboch týchto automatov za seba“ tak, aby z každého stavu i automatu $\mathcal{A}_1$ s nenulovým koncovým ohodnotením $\tau_1(i)$ viedol do každého stavu j automatu $\mathcal{A}_2$ s nenulovým počiatočným ohodnotením $\iota_2(j)$ prechod ohodnotený prvkom $\tau_1(i) \cdot \iota_2(j)$; samozrejme ale musíme tieto stavy vhodne prečíslovať. To znamená, že $\mathcal{A} = (n_1 + n_2, \mathbf{i}, A, \mathbf{f})$, kde $\mathbf{i} = (\mathbf{i}_1, \mathbf{0})$,

$$A = \begin{pmatrix} A_1 & \mathbf{f}_1 \mathbf{i}_2 \\ \mathbf{0} & A_2 \end{pmatrix}$$

a $\mathbf{f} = (\mathbf{0}, \mathbf{f}_2^T)^T$. Z vety 1.6.3 a dôsledku 1.5.6 potom dostávame

$$\|\mathcal{A}\| = \mathbf{i} A^* \mathbf{f} = \begin{pmatrix} \mathbf{i}_1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} A_1^* & A_1^* \mathbf{f}_1 \mathbf{i}_2 A_2^* \\ \mathbf{0} & A_2^* \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{f}_2 \end{pmatrix} = \mathbf{i}_1 A_1^* \mathbf{f}_1 \mathbf{i}_2 A_2^* \mathbf{f}_2 = \|\mathcal{A}_1\| \cdot \|\mathcal{A}_2\|,$$

čo dokazuje správnosť konštrukcie. $\square$

Lema 1.8.8. *Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $\mathcal{B}$ je konečný S' -automat nad S . Potom existuje konečný S' -automat $\mathcal{A}$ nad S taký, že $\|\mathcal{A}\| = \|\mathcal{B}\|^*$.*

Dôkaz. Bez ujmy na všeobecnosti predpokladajme, že $\mathcal{B} = (n, \mathbf{i}', A', \mathbf{f}')$ pre nejaké $n \in \mathbb{N}$. Položme $\mathcal{A} = (n + 1, \mathbf{i}, A, \mathbf{f})$, kde $\mathbf{i} = (1, \mathbf{0})$,

$$A = \begin{pmatrix} 0 & \mathbf{i}' \\ \mathbf{f}' & A' \end{pmatrix}$$

a $\mathbf{f} = (1, \mathbf{0})^T$. Z vety 1.6.3 a lemy 1.5.5 potom dostávame

$$\|\mathcal{A}\| = \mathbf{i} A^* \mathbf{f} = \begin{pmatrix} 1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} (\mathbf{i}'(A')^* \mathbf{f}')^* & (\mathbf{i}'(A')^* \mathbf{f}')^* \mathbf{i}'(A')^* \\ (A' + \mathbf{f}' \mathbf{i}')^* \mathbf{f}' & (A' + \mathbf{f}' \mathbf{i}')^* \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{0} \end{pmatrix} = (\mathbf{i}'(A')^* \mathbf{f}')^* = \|\mathcal{B}\|^*,$$

čo dokazuje správnosť konštrukcie. $\square$

Tvrdenia dokázané v predchádzajúcich piatich lemach možno zhrnúť v podobe nasledujúcej vety o ekvivalencii konečných automatov a racionálnych výrazov nad spočítateľne úplnými polokruhmi, ktorá je zovšeobecnením dobre známeho tvrdenia z teórie formálnych jazykov.

Veta 1.8.9. *Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh, $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $a \in S$. Potom sú nasledujúce tvrdenia ekvivalentné:*

- (i) *Existuje konečný S' -automat $\mathcal{A}$ nad S taký, že $\|\mathcal{A}\| = a$.*
- (ii) *Existuje S' -racionálny výraz G nad S taký, že $\|G\| = a$.*

Dôkaz. Ak existuje konečný S' -automat $\mathcal{A}$ nad S taký, že $\|\mathcal{A}\| = a$, podľa lemy 1.8.4 existuje aj S' -racionálny výraz G taký, že $\|G\| = \|\mathcal{A}\| = a$.

Opačnú implikáciu možno dokázať štruktúrnou indukciou: ak $G = a$ pre nejaké $a \in S'$, existuje konečný automat $\mathcal{A}$ spĺňajúci $\|\mathcal{A}\| = \|G\|$ podľa lemy 1.8.5; ak $G = E + F$ pre nejaké $E, F \in \text{RatE}(S', S)$, existuje takýto automat $\mathcal{A}$ podľa lemy 1.8.6; ak $G = E \cdot F$ pre nejaké $E, F \in \text{RatE}(S', S)$, možno na existenciu automatu $\mathcal{A}$ usúdiť podľa lemy 1.8.7; ak napokon $G = E^*$ pre nejaké $E \in \text{RatE}(S', S)$, je existencia automatu $\mathcal{A}$ ekvivalentného výrazu G dôsledkom lemy 1.8.8. $\square$

Definícia 1.8.10. Nech $(S, +, \cdot, 0, 1)$ je spočítateľne úplný polokruh a $S' \subseteq S$ je množina obsahujúca prvky 0 a 1. Prvok $a \in S$ je *S' -racionálnym prvkom polokruhu S* , ak existuje S' -racionálny výraz E nad S taký, že $\|E\| = a$. Množinu všetkých S' -racionálnych prvkov polokruhu S označíme $\text{Rat}(S', S)$.

Poznámka 1.8.11. Množina $\text{Rat}(S', S)$ všetkých S' -racionálnych prvkov polokruhu S je očividne uzavretá na súčet, na súčin a na iteráciu.

Teóriu racionálnych prvkov spočítateľne úplného polokruhu nateraz ďalej rozvíjať nebudeme. K podobným myšlienkam sa však ešte vrátíme v kapitole 3 v súvislosti s racionálnymi mocninovými radmi a konečnými automatmi s váhami, ktoré tvoria prakticky zďaleka najvýznamnejšiu triedu konečných automatov nad polokruhmi.

Cvičenia

1. Pridajte k príkladom z oddielu 1.2 ďalšie polokruhy s prirodzenou grafovou interpretáciou maticevého násobenia a prípadne aj iterácie.
2. Interpretujte operáciu transpozície štvorcovej matice nad polokruhom $(2^{\Sigma^*}, \cup, \cdot, \emptyset, \{\varepsilon\})$ z pohľadu jazykov rozoznávaných nedeterministickými konečnými automatmi nad abecedou Σ .
3. Dokážte, že axióma (Z) je v definícii polokruhu skutočne podstatná – nie je teda dôsledkom platnosti zvyšných axiém. Inými slovami: nájdite príklad algebry $(S, \oplus, \otimes, 0, 1)$ spĺňajúcej (A1), (A2), (A3), (M1), (M2), (D1) a (D2) a zároveň nespĺňajúcej (Z).
4. *Hemiokruh* (angl. *hemiring*) sa od polokruhu (angl. *semiring*) líši tým, že v ňom nemusí existovať multiplikatívny neutrálny prvok 1. Ide teda o štvoricu $(S, \oplus, \otimes, 0)$ spĺňajúcu podmienky (A1), (A2), (A3), (M1), (D1), (D2) a (Z) z definície polokruhu. Definujte *nenulové* mocniny matice nad hemiokruhom a dokážte pre ne tvrdenie analogické k vete 1.3.10.
5. *Silný bimonoid* [32] sa od polokruhu líši tým, že v ňom nemusia platiť distributívne zákony – ide teda o algebru $(S, +, \cdot, 0, 1)$ spĺňajúcu podmienky (A1), (A2), (A3), (M1), (M2) a (Z) z definície polokruhu. Nájdite príklady silných bimonoidov, ktoré nie sú polokruhmi. Možno na prípad silných bimonoidov zovšeobecniť tvrdenie 1.3.5? Ako je to s vetou 1.3.10?
6. *Homomorfizmom polokruhov* $(S, +, \cdot, 0_S, 1_S)$ a $(T, \oplus, \otimes, 0_T, 1_T)$ nazveme zobrazenie $h: S \rightarrow T$ také, že $h(0_S) = 0_T$, $h(1_S) = 1_T$ a súčasne pre všetky $a, b \in S$ je $h(a + b) = h(a) \oplus h(b)$ a $h(a \cdot b) = h(a) \otimes h(b)$. Bijektívny homomorfizmus polokruhov nazveme *izomorfizmom polokruhov* a dva polokruhy nazveme *izomorfnými*, ak medzi nimi existuje izomorfizmus.
Dokážte nasledujúce vlastnosti homomorfizmov a izomorfizmov polokruhov: identické zobrazenie na ľubovoľnom polokruhu S je izomorfizmus; zloženie dvoch homomorfizmov (resp. izomorfizmov) polokruhov je opäť homomorfizmus (resp. izomorfizmus) polokruhov; inverzné zobrazenie k izomorfizmu polokruhov je opäť izomorfizmus polokruhov.
7. Dokážte, že Viterbiho polokruh $([0, 1], \max, \cdot, 0, 1)$ je izomorfný s tropickým polokruhom nezáporných reálnych čísel $(\mathbb{R}_{\geq 0} \cup \{\infty\}, \min, +, \infty, 0)$.
8. Nájdite spočítateľne úplný polokruh $(S, +, \cdot, 0, 1, \Phi)$, ktorý nie je úplný.²²
9. Polokruh $(S, +, \cdot, 0, 1)$ sa nazýva *idempotentný*, ak je idempotentná jeho aditívna operácia – čiže ak pre všetky $a \in S$ platí $a + a = a$. Príkladmi idempotentných polokruhov sú napríklad tropické polokruhy alebo polokruhy formálnych jazykov. Musí byť každý idempotentný polokruh úplný?²³

²²Teda *neexistuje* žiadne zobrazenie $\Psi: \mathcal{F}(S) \rightarrow S$ také, že pre všetky $F \in \mathcal{F}_\omega(S)$ platí $\Psi(F) = \Phi(F)$ a $(S, +, \cdot, 0, 1, \Psi)$ je úplný polokruh.

²³V zmysle, že *existuje* zobrazenie $\Phi: \mathcal{F}(S) \rightarrow S$ také, že $(S, +, \cdot, 0, 1, \Phi)$ je úplný polokruh.

10. Musí byť každý *konečný* polokruh úplný alebo aspoň spočítateľne úplný?
11. Nech $(S, +, \cdot, 0, 1)$ je polokruh a $T \subseteq S$ je množina uzavretá na obidve operácie polokruhu S taká, že $0, 1 \in T$; algebra $(T, +, \cdot, 0, 1)$ je v takom prípade opäť polokruh, ktorý nazveme *podpolokruhom* polokruhu S . Musí byť podpolokruh (spočítateľne) úplného polokruhu (spočítateľne) úplný?
12. Existuje netriviálny okruh (s jednotkou), ktorý je súčasne aj spočítateľne úplným polokruhom?
13. Polokruh $(S, +, \cdot, 0, 1)$ sa nazýva *lokálne uzavretý* [35], ak pre každé $a \in S$ existuje $k \in \mathbb{N}$ také, že $a^0 + a^1 + \dots + a^k = a^0 + a^1 + \dots + a^{k+1}$. Dokážte, že potom $a^0 + a^1 + \dots + a^k = a^0 + a^1 + \dots + a^m$ pre všetky $m \geq k$. Iteráciu a^* prvku a tak možno prirodzeným spôsobom definovať pomocou predpisu $a^* := a^0 + a^1 + \dots + a^k$. Dokážte, že existuje lokálne uzavretý polokruh, ktorý nie je spočítateľne úplný (z čoho vyplýva, že trieda spočítateľne úplných polokruhových nezáhrňa všetky polokruhy so zmysluplne definovanou skalárnou iteráciou).
14. Dokážte, že pre každú štvorcovú maticu $A = (a_{i,j})_{n \times n}$ nad tropickým polokruhom prirodzených čísel – čiže nad $(S, \oplus, \otimes, 0, 1) = (\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ – platí $A^* = A^0 \oplus A^1 \oplus \dots \oplus A^{n-1}$.
15. Nech A^T je transpozícia matice $A = (a_{i,j})_{n \times n}$ nad komutatívnym spočítateľne úplným polokruhom $(S, +, \cdot, 0, 1)$. S využitím ohodnotených grafov dokážte alebo vyvráťte: $(A^T)^* = (A^*)^T$.
16. Nech S je spočítateľne úplný polokruh, nech $S' \subseteq S$ je množina obsahujúca prvky 0 a 1 a $\mathcal{A} = (Q, \varphi, \iota, \tau)$ je konečný S' -automat nad S . Beh γ automatu $\mathcal{A}$ zo stavu p do stavu q nazveme *úspešným*, ak $\iota(p) \neq 0$ a $\tau(q) \neq 0$. Označme ako $\mathcal{S}(\mathcal{A})$ množinu všetkých úspešných behov automatu $\mathcal{A}$. Dokážte, že

$$\|\mathcal{A}\| = \sum_{\gamma \in \mathcal{S}(\mathcal{A})} \|\gamma\|.$$

17. Dokážte vetu 1.7.2 len s použitím elementárnych metód, bez odvolávania sa na vzorce pre iteráciu blokovej matice dokázané v oddiele 1.5.
18. V dôkaze lemy 1.8.4 sa využíva modifikácia algoritmu 1.5.1, pri ktorej tento algoritmus namiesto vyhodnocovania súčtov, súčinov a iterácií konštruuje zodpovedajúci racionálny výraz. Zapište takto upravený algoritmus vo forme pseudokódu a dokážte, že pre vstupnú maticu $A = (a_{i,j})_{n \times n}$ s iteráciou $A^* = (a_{i,j}[*])_{n \times n}$ je jeho výstupom matica racionálnych výrazov $A = (E_{i,j})_{n \times n}$ taká, že pre $i, j = 1, \dots, n$ platí $\|E_{i,j}\| = a_{i,j}[*]$.
19. Nech S je spočítateľne úplný polokruh a $S' \subseteq S$ je množina obsahujúca prvky 0 a 1. *Kladnou iteráciou* prvku $a \in S$ nazveme prvok $a^+ = \sum_{t \in \mathbb{N} \setminus \{0\}} a^t$. Zistite, či je množina $\text{Rat}(S', S)$ uzavretá na kladnú iteráciu a svoje tvrdenie dokážte.

Poznámky a odkazy na samoštúdium

Súvis medzi ohodnotenými grafmi a maticami nad polokruhmi sa spomína ako v literatúre o grafových algoritmoch, tak aj v literatúre o teórii automatov. Algoritmicke orientovaný pohľad na problematiku je podaný napríklad v učebniciach [2, 27, 71]. Pohľad z perspektívy teórie automatov možno nájsť najmä v literatúre o automatoch s váhami, ktoré sú prakticky najvýznamnejšou triedou automatov nad polokruhmi, a ktorými sa budeme zaoberať v kapitole 3 – spomeňme predovšetkým monografiu [30] a jej kapitoly [29, 36, 87].

Algebraicky orientované monografie o polokruhoch [50, 52] sú z veľkej časti obsiahlymi zbierkami definícií – skôr než na systematické štúdium ich možno odporúčať ako užitočné referenčné zdroje.

Formalizácie polokruhov s dobre definovanou iteráciou matice sa líšia od zdroja k zdroju. Algoritmicky orientovaná literatúra sa často obmedzuje na rôzne triedy idempotentných polokruhov, ako sú napríklad (trochu neporiadne definované) *uzavreté polokruhy* [2] alebo *Kleeneho algebry* [71]. Pri definícii uzavretých polokruhov používanej v [27] ide v podstate o spočítateľne úplné polokruhy používané v tomto texte (naša definícia spočítateľne úplných polokruhov je zas viac-menej prebratá z [52]). V teórii automatov sa často používajú úplné polokruhy alebo ich podtriedy (najčastejšie tzv. *spojité polokruhy*) [29, 36, 87]. Je ale možný aj o niečo všeobecnejší prístup založený na čisto axiomatickom prístupe k definícii iterácie a vychádzajúci z práce J. H. Conwaya [22]. Ten v stručnosti naznačíme v oddiele 3.13, kde definujeme tzv. *conwayovské polokruhy* [29, 36] a *častočné conwayovské polokruhy nad ideálom* [13, 37]. S menom J. Sakarovitcha [94] je spojený alternatívny prístup založený na *topologických polokruhoch*.

Materiál spracovaný v rámci tejto kapitoly má svoje historické korene v práci S. C. Kleeneho [64] a predovšetkým v jeho konštrukcii racionálnych (regulárnych) výrazov ku konečným automatom. Variant tohto algoritmu opísaný R. McNaughtonom a H. Jamadom (*angl.* prepis Yamada) [83] dnes možno chápať ako špeciálny prípad algoritmu 1.5.1 na výpočet iterácie matice nad spočítateľne úplným (resp. úplným) polokruhom. V rovnakej dobe boli objavené aj ďalšie algoritmy dnes chápané ako špeciálne prípady algoritmu 1.5.1: B. Roy [93] a S. Warshall [102] nezávisle na sebe opísali algoritmus na konštrukciu reflexívno-tranzitívneho uzáveru grafu a R. W. Floyd [45] prišiel s algoritmom hľadania najlacnejších ciest v grafe dnes známym ako Floydov-Warshallov algoritmus [23] (už samotné toto pomenovanie je pritom dokladom toho, že výpočet reflexívno-tranzitívneho uzáveru grafu a hľadanie najlacnejších sledov v grafe sú v skutočnosti „rovnaké“ problémy). Prechod od podobných výsledkov nad konkrétnymi polokruhmi k ucelenej teórii ohodnotených grafov, matíc a konečných automatov nad polokruhmi je do veľkej miery folklórnym dielom; vychádza však z myšlienkového odkazu M.-P. Schützenbergera [98] a J. H. Conwaya [22].

Prezentácia oddielu 1.2 je sčasti inšpirovaná knihou [27], kde možno nájsť aj ďalšie príklady podobného charakteru.

Kapitola 2

Formálne mocninové rady

Myšlienky načaté v predchádzajúcej kapitole budeme rozvíjať ďalej, ale zameriame sa teraz na o niečo špecifickejšiu triedu polokruhov významnú z hľadiska teórie automatov. Naším cieľom, ktorý budeme naplňať v nasledujúcich dvoch kapitolách, je postupne dospieť ku skúmaniu pravdepodobne najdôležitejšej triedy konečných automatov nad polokruhmi, tvorenej *konečnými automatmi s váhami*. Pôjde o zovšeobecnenie klasických nedeterministických konečných automatov rozoznávajúcich formálne jazyky, pri ktorom sú prechodom automatu navyše priradené váhy z nejakého polokruhu reprezentujúce napríklad ich cenu, spoľahlivosť, užitočnosť, či multiplicitu. Ku každému behu v automate s váhami možno priradiť prvok polokruhu, ktorý sa dá vypočítať postupným prenásobením váh jeho prechodov.¹ K slovu w potom prislúcha prvok získaný ako súčet hodnôt všetkých behov automatu na slove w .

Automaty s váhami už teda nerozoznávajú jazyky, ale realizujú funkcie priraďujúce jednotlivým slovám prvky polokruhu.² *Kvalitatívna* vlastnosť príslušnosti do jazyka je tak nahradená vlastnosťou *kvantitatívnou*. Takto „vylepšené“ automaty nachádzajú uplatnenie napríklad pri analýze prirodzeného jazyka alebo vo formálnej verifikácii a možno ich využiť aj ako nástroj na kódovanie obrazu.

Ako aditívnu a multiplikatívnu operáciu na funkciách realizovaných automatmi s váhami budeme uvažovať prirodzené zovšeobecnenia operácií zjednotenia a zretazovania jazykov. Ukazuje sa však, že takto získaná multiplikatívna operácia nepripomína ani tak typické násobenie funkcií, ako skôr Cauchyho súčin radov. Prijmeme teda notáciu a terminológiu radov a funkcie realizované automatmi s váhami budeme chápať ako *formálne mocninové rady o niekoľkých nekomutujúcich premenných*.

Pochopenie tohto prechodu vyžaduje určitý kontext – hovoriť o *mocninových* radoch v súvislosti s objektmi, v ktorých sa mocniny nevyskytujú, sa môže na prvý pohľad zdať zvláštne, ba až paradoxné. Pôvod tejto terminológie treba hľadať vo vzťahu k formálnym mocninovým radom s *komutujúcimi* premennými, ktoré sú v podobe vytvárajúcich – alebo generujúcich – funkcií hojne využívané v kombinatorike. Keďže je teória formálnych mocninových radov s komutujúcimi premennými pre teóriu ich nekomutatívnych náprotivkov myšlienkovým východiskom a hlavným zdrojom inšpirácie, venujeme jej prvé tri oddiely tejto kapitoly, v ktorých naznačíme aj význam tejto teórie pre kombinatoriku. Obmedzíme sa ale na základné fakty nutné na získanie predstavy o kontexte, v ktorom je teória automatov s váhami a radov s *nekomutujúcimi* premennými zasadená. O formálnych mocninových radoch s komutujúcimi premennými, vytvárajúcich funkciách a ich súvisi s analytickými funkciami sa možno dozvedieť viac napríklad na prednáškach o *analytickej a enumeratívnej kombinatorike* [70].

Hlavnou náplňou tejto kapitoly sú teda základy teórie formálnych mocninových radov o niekoľkých nekomutujúcich premenných – tú možno chápať ako kvantitatívnu analógiu teórie formálnych jazykov. Charakter kapitoly je skôr prípravný: dáme tu dohromady formalizmus potrebný na prácu s automatmi s váhami, ktorými sú formálne mocninové rady s nekomutujúcimi premennými realizované a ktoré sú zas kvantitatívnou obdobou klasických konečných automatov.

¹Do tohto súčinu ešte treba zakomponovať počiatočnú váhu prvého a koncovú váhu posledného stavu behu.

²Jazyky ale stále možno chápať ako špeciálny prípad takýchto funkcií, v rámci ktorého sú slovám priradené prvky booleovského polokruhu $\mathbb{B}$. Ide tak o skutočné *zovšeobecnenie* klasických konečných automatov.

2.1 Formálne mocninové rady o jednej premennej

Veľká časť súčasnej kombinatoriky sa nejakým spôsobom zaoberá postupnosťami prirodzených čísel, čo je dané charakterom otázky, ktorej rôzne varianty sa snaží zodpovedať *enumeratívna kombinatorika*: koľko existuje kombinatorických objektov daného typu a veľkosti $t \in \mathbb{N}$? Snaha tento vážny problém náležite sformalizovať nás vedie k základnému objektu skúmania enumeratívnej kombinatoriky, ktorým sú takzvané *kombinatorické triedy*. Ide o dvojice $\mathcal{C} = (\mathcal{C}, |\cdot|)$, kde $\mathcal{C}$ je množina nejakých kombinatorických objektov a $|\cdot|: \mathcal{C} \rightarrow \mathbb{N}$ je zobrazenie udávajúce veľkosť objektov z množiny $\mathcal{C}$. Požaduje sa pritom, aby vzor ľubovoľného prirodzeného čísla t v tomto zobrazení – čiže množina $\mathcal{C}_t = \{x \in \mathcal{C} \mid |x| = t\}$ všetkých objektov veľkosti t v triede $\mathcal{C}$ – bola konečná.

Príkladmi kombinatorických tried sú napríklad triedy všetkých grafov určitého typu s veľkosťou danou počtom vrcholov, triedy všetkých slov určitého typu s veľkosťou danou dĺžkou slova, triedy všetkých množín alebo multimnožín určitého typu s veľkosťou danou počtom prvkov, atď.

Ak pre uvažovanú kombinatorickú triedu $\mathcal{C}$ a všetky $t \in \mathbb{N}$ označíme c_t počet prvkov množiny $\mathcal{C}_t$, možno za základný problém enumeratívnej kombinatoriky považovať nájdenie čo možno najpresnejšieho opisu *enumeráčnej postupnosti* $(c_0, c_1, c_2, \dots)$ triedy $\mathcal{C}$. V jednoduchších prípadoch pritom môže ísť o vyjadrenie hodnôt c_t pre $t \in \mathbb{N}$ v uzavretom tvare, častejšie však ide napríklad o skúmanie asymptotických vlastností členov enumeráčnej postupnosti pre $t \rightarrow \infty$.

Za dve najdôležitejšie konštrukcie v elementárnej kombinatorike možno považovať *disjunktné zjednotenie* (pravidlo súčtu) a *karteziánsky súčin* (pravidlo súčinu). Obidve tieto operácie možno uvažovať aj na kombinatorických triedach.

Disjunktným zjednotením kombinatorických tried $(\mathcal{C}, |\cdot|_{\mathcal{C}})$ a $(\mathcal{D}, |\cdot|_{\mathcal{D}})$ nazveme triedu $(\mathcal{C} + \mathcal{D}, |\cdot|)$, kde $\mathcal{C} + \mathcal{D} = (\mathcal{C} \times \{1\}) \cup (\mathcal{D} \times \{2\})$, pre všetky $x \in \mathcal{C}$ je $|(x, 1)| = |x|_{\mathcal{C}}$ a pre všetky $y \in \mathcal{D}$ je $|(y, 2)| = |y|_{\mathcal{D}}$.

Pri *karteziánskom súčine* kombinatorických tried $(\mathcal{C}, |\cdot|_{\mathcal{C}})$ a $(\mathcal{D}, |\cdot|_{\mathcal{D}})$ je veľkosť dvojice daná *súčtom* veľkostí jednotlivých jej zložiek. To znamená, že ide o triedu $(\mathcal{C} \times \mathcal{D}, |\cdot|)$, kde pre všetky $x \in \mathcal{C}$ a $y \in \mathcal{D}$ je $|(x, y)| = |x|_{\mathcal{C}} + |y|_{\mathcal{D}}$. Prvky veľkosti t súčinu $\mathcal{C} \times \mathcal{D}$ sú teda dvojicami objektov, kde jeden z nich patrí do $\mathcal{C}$ a je veľkosti $k \in \{0, \dots, t\}$, kým druhý je z $\mathcal{D}$ a veľkosti $t - k$. Množinu $(\mathcal{C} \times \mathcal{D})_t$ preto dostaneme disjunktným zjednotením množín $\mathcal{C}_k \times \mathcal{D}_{t-k}$ takýchto dvojíc pre $k = 0, \dots, t$.

Na zmysluplnú prácu s enumeráčnymi postupnosťami je na nich potrebné uvažovať operácie zodpovedajúce najdôležitejším operáciám na príslušných kombinatorických triedach. Ak teda $(c_0, c_1, c_2, \dots)$ je enumeráčna postupnosť kombinatorickej triedy $\mathcal{C}$ a $(d_0, d_1, d_2, \dots)$ je enumeráčna postupnosť triedy $\mathcal{D}$, mala by použitím *aditívnej operácie* na tieto dve postupnosti vzniknúť postupnosť $(s_0, s_1, s_2, \dots)$ taká, že pre všetky $t \in \mathbb{N}$ je

$$s_t = |(\mathcal{C} + \mathcal{D})_t| = |\mathcal{C}_t| + |\mathcal{D}_t| = c_t + d_t.$$

Podobne by sme mali použitím *multiplikatívnej operácie* na uvedené dve postupnosti dostať postupnosť $(p_0, p_1, p_2, \dots)$ takú, že pre všetky $t \in \mathbb{N}$ je

$$p_t = |(\mathcal{C} \times \mathcal{D})_t| = \left| \bigcup_{k=0}^t (\mathcal{C}_k \times \mathcal{D}_{t-k}) \right| = \sum_{k=0}^t c_k d_{t-k}. \quad (2.1)$$

Vidíme tak, že kým za aditívnu operáciu operáciu môžeme vziať bežný súčet postupností, požadovaná multiplikatívna operácia nezodpovedá bežnému súčinu postupností po zložkách, ale skôr operácii, ktorá sa zvykne nazývať *konvolúciou* postupností.

Rovnakým spôsobom ako v (2.1) naopak počítame koeficienty *Cauchyho súčinu* mocninových radov, ktorý je prirodzenou multiplikatívnou operáciou na týchto radoch. V nasledujúcom preto zmeníme uhol pohľadu a enumeráčne postupnosti budeme chápať ako mocninové rady, pri ktorých nás ale nebudú zaujímať žiadne otázky okolo ich konvergencie. Na takýchto objektoch, ktoré nazveme *formálnymi mocninovými radmi*, následne budeme môcť očakávaným spôsobom zaviesť aditívnu a multiplikatívnu operáciu, ktoré nad prirodzenými číslami priamo zodpovedajú disjunktnému zjednoteniu a karteziánskemu súčinu kombinatorických tried. Koeficientmi formálnych mocninových radov však nebudú vždy iba prirodzené čísla, ale budú nimi môcť byť napríklad aj prvky ľubovoľného polokruhu.

Formálny mocninový rad r o jednej premennej – či neurčitej – z nad množinou koeficientov S možno intuitívne chápať ako *formálny súčet*

$$r(z) = \sum_{t \in \mathbb{N}} a_t z^t, \quad (2.2)$$

kde koeficienty $a_0, a_1, a_2, \dots$ patria do S . Prívlastok „formálny“ tu znamená, že v skutočnosti nemáme žiadnu snahu daný súčet vypočítať. To koniec koncov ani nemusí byť možné – z matematickej analýzy je napríklad dobre známe, že existujú mocninové rady nad $\mathbb{R}$ s nulovým polomerom konverencie; tu navyše (zatiaľ) pracujeme s množinou koeficientov S , na ktorej ani len nemusí byť definovaný súčet dvoch prvkov. Pri formálnom mocninovom rade otázky sčítateľnosti a konverencie nehrajú žiadnu rolu – a ani nás z logiky vecí nezaujímajú. Formálny mocninový rad ako matematický objekt nezodpovedá „hodnote nekonečného súčtu“, ale jeho „formálnemu zápisu“ v duchu (2.2).

Herbert Wilf [104] túto myšlienku vyjadril slovami, že formálny mocninový rad je iba „vešiakom na koeficienty“. To sa odráža aj v nasledujúcej definícii, v ktorej formálny mocninový rad stotožňujeme s postupnosťou jeho koeficientov – nimi je totiž jednoznačne určená všetka podstatná informácia vo formálnom zápise typu (2.2). Formálny mocninový rad je teda len nekonečná postupnosť, ktorú zapisujeme a interpretujeme trochu neobvyklým spôsobom.

Definícia 2.1.1. Nech S je množina. *Formálny mocninový rad o jednej premennej z s koeficientmi v S je nekonečná postupnosť*

$$r = (a_0, a_1, a_2, \dots),$$

kde pre všetky $t \in \mathbb{N}$ je $a_t \in S$. Namiesto $r = (a_0, a_1, a_2, \dots)$ potom píšeme

$$r = r(z) = a_0 z^0 + a_1 z^1 + a_2 z^2 + \dots = \sum_{t \in \mathbb{N}} a_t z^t.$$

Koeficient a_t pri z^t označujeme aj

$$a_t =: [z^t]r(z) = [z^t]r$$

a koeficient a_0 nazývame *konštantným*. Množinu všetkých formálnych mocninových radov o premennej z nad S označujeme $S[[z]]$.

Zapísať nekonečnú postupnosť v podobe mocninového radu by samozrejme nemalo význam, keby sme na formálnych mocninových radoch nezavedli žiadne zmysluplné operácie. Definovať najprirodzenejšie z týchto operácií – ako napríklad súčet alebo Cauchyho súčin radov – ale znamená do veľkej miery využiť nejaké už „existujúce“ operácie na množine koeficientov S . Ďalej si už teda nevystačíme s pojmom formálnych mocninových radov *nad množinou*; budeme musieť predpokladať, že na množine koeficientov S je vopred daná algebraická štruktúra.

Na definovanie niektorých základných operácií na formálnych mocninových radoch stačí pomerne slabá algebraická štruktúra na množine koeficientov; pri jej postupnom zosilňovaní následne možno pridávať ďalšie operácie. Pre nás budú najzaujímavejšie dva prípady: všeobecnejší, keď množina koeficientov tvorí polokruh a špeciálnejší, keď sú koeficientmi prvky poľa komplexných čísel $\mathbb{C}$. Formálne mocninové rady s *koeficientmi v polokruhu* pritom budú kľúčové pre náš neskorší prechod k automatom s váhami a radom s niekoľkými nekomutujúcimi premennými; rady s *komplexnými koeficientmi* – špeciálnym prípadom takýchto radov sú rady s reálnymi a prirodzenými koeficientmi – sú zas základom teórie vytvárajúcich funkcií, na ktorej je do veľkej miery založená moderná kombinatorika.

Zamerajme sa teraz na prípad formálnych mocninových radov *nad polokruhom* – operácie súčtu a súčinu na polokruhu koeficientov tu možno využiť na definovanie operácií *súčtu* a *Cauchyho súčinu*³ formálnych mocninových radov.

³Pri pomenovaniach tohto typu ide samozrejme len o zneužitie mien slávnych matematikov na zdôraznenie analógie s dobre známymi myšlienkami matematickej analýzy.

Definícia 2.1.2. Nech S je polokruh a nech $r = \sum_{t \in \mathbb{N}} a_t z^t$ a $s = \sum_{t \in \mathbb{N}} b_t z^t$ sú formálne mocninové rady v $S[[z]]$. Súčet radov r a s je rad $r + s \in S[[z]]$ daný ako

$$r + s = \sum_{t \in \mathbb{N}} (a_t + b_t) z^t$$

a Cauchyho súčin radov r a s je rad $r \cdot s \in S[[z]]$ daný ako

$$r \cdot s = \sum_{t \in \mathbb{N}} c_t z^t,$$

kde koeficient c_t je pre všetky $t \in \mathbb{N}$ definovaný predpisom

$$c_t = \sum_{k=0}^t a_k b_{t-k}.$$

Prvok a polokruhu S budeme obvykle stotožňovať s formálnym mocninovým radom $r(z)$ takým, že $[z^0]r(z) = a$ a $[z^t]r(z) = 0$ pre všetky prirodzené $t \geq 1$; budeme teda písať $r(z) = a$. Podobne budeme písať iba z pre formálny mocninový rad $r(z)$ taký, že $[z^1]r(z) = 1$ a $[z^t]r(z) = 0$ pre všetky prirodzené $t \neq 1$. Umocňovanie radov definujeme štandardným spôsobom: pre $r \in S[[z]]$ kladieme $r^0 = 1$ a $r^{t+1} = r^t \cdot r$ pre všetky $t \in \mathbb{N}$. Ľahko potom vidieť, že pre všetky $n \in \mathbb{N}$ a $a_0, a_1, \dots, a_n \in S$ je výraz $a_0 z^0 + a_1 z^1 + \dots + a_n z^n$ rovný radu $r(z)$ takému, že $[z^t]r(z) = a_t$ pre $t = 0, \dots, n$ a $[z^t]r(z) = 0$ pre všetky prirodzené $t > n$. Týmto spôsobom možno získať práve všetky formálne mocninové rady $r(z)$ také, že $[z^t]r(z) \neq 0$ pre konečne veľa rôznych $t \in \mathbb{N}$; takéto formálne mocninové rady nazývame *polynómy* a množinu všetkých polynómov o jednej premennej z nad polokruhom S označujeme $S[z]$.

Je triviálnym mechanickým cvičením dokázať nasledujúce tvrdenie, podľa ktorého množiny $S[[z]]$ a $S[z]$ s operáciami súčtu a Cauchyho súčinu radov tvoria opäť polokruhy.

Tvrdenie 2.1.3. Nech $(S, +, \cdot, 0, 1)$ je polokruh. Algebry $(S[[z]], +, \cdot, 0, 1)$ a $(S[z], +, \cdot, 0, 1)$ sú potom tiež polokruhy.

Je zrejmé, že pokiaľ polokruh S nie je (spočítateľne) úplný, nemožno vo všeobecnosti definovať ani (spočítateľne) nekonečné súčty formálnych mocninových radov z množiny $S[[z]]$. Súčty niektorých špeciálnych nekonečných systémov radov ale možno definovať vždy. Ak napríklad v danom nekonečnom systéme existuje pre každé $t \in \mathbb{N}$ iba konečne veľa radov s nenulovým koeficientom pri z^t , možno nekonečný súčet definovať po zložkách: každý koeficient nekonečného súčtu radov je tak konečným súčtom prvkov polokruhu S . Takéto systémy formálnych mocninových radov nazveme *lokálne konečnými*.

Definícia 2.1.4. Nech S je polokruh. Systém $(r_i \mid i \in I)$ formálnych mocninových radov z $S[[z]]$ je *lokálne konečný*, ak je pre každé $t \in \mathbb{N}$ konečná množina $I(t) := \{i \in I \mid [z^t]r_i(z) \neq 0\}$.

Definícia 2.1.5. Nech S je polokruh a $(r_i \mid i \in I)$ je lokálne konečný systém formálnych mocninových radov z $S[[z]]$. Súčet systému radov $(r_i \mid i \in I)$ potom definujeme ako

$$\sum_{i \in I} r_i = r,$$

kde r je formálny mocninový rad v $S[[z]]$ taký, že pre všetky $t \in \mathbb{N}$ je

$$[z^t]r(z) = \sum_{i \in I(t)} [z^t]r_i(z).$$

Poznámka 2.1.6. Je jednoduchým cvičením overiť, že každý *konečný* systém $(r_i \mid i \in \{i_1, \dots, i_n\})$ je lokálne konečný a jeho súčet je rovný $r_{i_1} + \dots + r_{i_n}$. Definícia 2.1.5 je teda konzistentná s definíciou sčítania v polokruhu $S[[z]]$.

Mocninové rady sú v matematickej analýze predovšetkým reprezentáciami funkcií. To samozrejme neplatí o *formálnych* mocninových radoch, pri ktorých otázky sčítateľnosti nehrajú rolu. V analytickom pohľade na mocninové rady ako na funkcie ale majú korene definície niektorých *operácií* na mocninových radoch, ktoré sa ukazujú byť užitočnými aj v teórii formálnych mocninových radov.⁴ Jednou z takýchto operácií je *skladanie mocninových radov* vychádzajúce z bežného skladania funkcií. Pre dvojicu formálnych mocninových radov $r(z)$ a $s(z)$ chceme definovať ich zloženie $(r \circ s)(z) = r(s(z))$ tak, že za každý výskyt premennej z v rade $r(z)$ „dosadíme“ rad $s(z)$. Nad všeobecným polokruhom to zrejme nepôjde pre všetky dvojice radov: ak napríklad $r(z) = \sum_{t \in \mathbb{N}} z^t$ a $s(z) = 1$, dostávame zložením týchto dvoch radov nekonečný súčet jednotiek, ktorý nemusí byť definovaný. Akonáhle ale $[z^0]s(z) = 0$, je rad $(r \circ s)(z)$ daný súčtom cez lokálne konečný systém radov, a teda je dobre definovaný – to dokážeme v tvrdení 2.1.7. Zloženie $(r \circ s)(z)$ formálnych mocninových radov $r(z)$ a $s(z)$ teda nedefinujeme pre všetky dvojice formálnych mocninových radov, ale iba pre tie dvojice radov, kde $s(z)$ má nulový konštantný koeficient.

Tvrdenie 2.1.7. *Nech S je polokruh, $(a_0, a_1, a_2, \dots)$ je postupnosť prvkov S a $s(z)$ je formálny mocninový rad z $S[[z]]$ taký, že $[z^0]s(z) = 0$. Potom je systém $(a_t s(z)^t \mid t \in \mathbb{N})$ lokálne konečný.*

Dôkaz. Stačí indukciou vzhľadom na t dokázať, že pre všetky $t \in \mathbb{N}$ a $k = 0, \dots, t-1$ je $[z^k]s(z)^t = 0$.⁵ Tvrdenie platí triviálne pre $t = 0$ a vďaka predpokladu na $s(z)$ aj pre $t = 1$. Predpokladajme teraz platnosť tvrdenia pre $t = q$ a uvažujme $t = q + 1$. Pre $k = 0, \dots, q$ z definície Cauchyho súčtu dostávame

$$[z^k]s(z)^{q+1} = [z^k](s(z)^q s(z)) = \sum_{j=0}^k ([z^j]s(z)^q)([z^{k-j}]s(z)). \quad (2.3)$$

Pre $j = 0, \dots, k-1$ je vďaka indukčnému predpokladu $[z^j]s(z)^q = 0$; pre $j = k$ zas z predpokladu o rade $s(z)$ vyplýva $[z^{k-j}]s(z) = [z^0]s(z) = 0$. Dosadením týchto nulových hodnôt do (2.3) získavame kýženu rovnosť $[z^k]s(z)^{q+1} = 0$. $\square$

Z dokázaného tvrdenia vyplýva korektnosť nasledujúcej definície.

Definícia 2.1.8. Nech S je polokruh, $r(z) = \sum_{t \in \mathbb{N}} a_t z^t$ a $s(z) = \sum_{t \in \mathbb{N}} b_t z^t$ sú formálne mocninové rady z $S[[z]]$ a $b_0 = 0$. *Zloženie radov* $r(z)$ a $s(z)$ je formálny mocninový rad $(r \circ s)(z) = r(s(z)) \in S[[z]]$ definovaný nekonečným súčtom

$$(r \circ s)(z) = r(s(z)) = \sum_{t \in \mathbb{N}} a_t s(z)^t.$$

Prejdime teraz k formálnym mocninovým radom nad špeciálnejšími algebrami, než sú všeobecné polokruhy. Uvažujme najprv prípad, keď sú koeficienty radu prvkami nejakého *okruhu*⁶ $(R, +, \cdot, 0, 1)$ – čiže polokruhu R takého, že ku každému $a \in R$ navyše v R existuje aditívny inverzný prvok $-a$ taký, že $a + (-a) = (-a) + a = 0$. Všimnime si, že existencia aditívnych inverzných prvkov v okruhu koeficientov R nám umožňuje definovať po zložkách *aditívne inverzné prvky* aj na $R[[z]]$ a $R[z]$.

Definícia 2.1.9. Nech R je okruh a $r = \sum_{t \in \mathbb{N}} a_t z^t$ je formálny mocninový rad z $R[[z]]$. Potom $-r$ označuje formálny mocninový rad

$$-r = -r(z) = \sum_{t \in \mathbb{N}} (-a_t) z^t.$$

Ak $s \in R[[z]]$ je nejaký ďalší formálny mocninový rad, píšeme namiesto $s + (-r)$ zvyčajne len $s - r$.

⁴Myšlienkové východisko v teórii analytických funkcií sa koniec koncov odráža už v označení $r(z)$, ktoré miestami používame pre formálny mocninový rad r o jednej premennej z .

⁵Potom nutne aj $[z^k]a_t s(z)^t = 0$ a v systéme radov $(a_t s(z)^t \mid t \in \mathbb{N})$ tak môžeme pre každé $k \in \mathbb{N}$ nájsť najviac konečne veľa radov $a_t s(z)^t$ takých, že $[z^k]a_t s(z)^t \neq 0$ – v takom prípade totiž musí byť $t \in \{0, \dots, k\}$.

⁶V tomto texte pod *okruhom* vždy rozumieme okruh *s jednotkou*.

Tvrdenie 2.1.10. *Nech R je okruh. Pre každý formálny mocninový rad $r \in R[[z]]$ potom:*

- (i) $r - r = -r + r = 0$;
- (ii) $-r = (-1) \cdot r$.

Dôkaz. Jednoduché cvičenie. □

Polokruhy $R[[z]]$ a $R[z]$ sa tak nad okruhom koeficientov R takisto stávajú okruhmi.

Dôsledok 2.1.11. *Nech R je okruh. Algebry $(R[[z]], +, \cdot, 0, 1)$ a $(R[z], +, \cdot, 0, 1)$ sú potom tiež okruhy.*

Pripomeňme si, že pod *oborom integrity*⁷ rozumieme netriviálny komutatívny okruh (s jednotkou) bez netriviálnych deliteľov nuly – ide teda o okruh R obsahujúci aspoň dva rôzne prvky taký, že pre ľubovoľnú dvojicu prvkov $a, b \in R$ je $a \cdot b = b \cdot a$, pričom táto spoločná hodnota je nenulová kedykoľvek sú nenulové obidva prvky a, b .

Obory integrity budú pre nás zaujímavé hlavne vďaka tomu, že v nich možno *krátiť nenulovými prvkami*: pre ľubovoľnú trojicu prvkov a, b, c oboru integrity, kde $c \neq 0$, je $a = b$ kedykoľvek $ac = bc$ (alebo $ca = cb$). V dôsledku uvedenej rovnosti je totiž $(a - b)c = 0$, čo vďaka predpokladu o nenulovosti prvku c môže byť v obore integrity pravda iba v prípade, že $a = b$.

Dokážeme teraz, že okruhy $(R[[z]], +, \cdot, 0, 1)$ a $(R[z], +, \cdot, 0, 1)$ sú pre obor integrity R tiež vždy obormi integrity. Toto tvrdenie uvádzame najmä preto, že v jeho dôsledku sú obormi integrity algebry $(\mathbb{F}[[z]], +, \cdot, 0, 1)$ a $(\mathbb{F}[z], +, \cdot, 0, 1)$ pre ľubovoľné pole $\mathbb{F}$ a špeciálne aj pre $\mathbb{F} = \mathbb{C}$ – v týchto algebrách tak možno krátiť nenulovými radmi.

Tvrdenie 2.1.12. *Nech R je obor integrity. Okruhy $(R[[z]], +, \cdot, 0, 1)$ a $(R[z], +, \cdot, 0, 1)$ sú potom tiež obory integrity.*

Dôkaz. Netriviálnosť a komutatívnosť oboch okruhov je zrejmá a okruh $(R[z], +, \cdot, 0, 1)$ je tiež evidentne podokruhom okruhu $(R[[z]], +, \cdot, 0, 1)$. Stačí teda ukázať, že pre ľubovoľné $r, s \in R[[z]]$ môže byť $r \cdot s = 0$ iba ak $r = 0$ alebo $s = 0$.

Sporom: nech $r = \sum_{t \in \mathbb{N}} a_t z^t$ a $s = \sum_{t \in \mathbb{N}} b_t z^t$ sú formálne mocninové rady nad R také, že $r \neq 0$, $s \neq 0$ a $r \cdot s = 0$. Nech m je *najmenšie* prirodzené číslo také, že $a_m \neq 0$ a n je *najmenšie* prirodzené číslo také, že $b_n \neq 0$ – keďže je R obor integrity, musí byť aj $a_m b_n \neq 0$. Potom

$$[z^{m+n}](r \cdot s)(z) = \sum_{k=0}^{m+n} a_k b_{m+n-k} = a_m b_n \neq 0,$$

čo je v spore s predpokladom $r \cdot s = 0$. □

Budeme sa teraz nejaký čas venovať formálnym mocninovým radom *nad* *polom* koeficientov $\mathbb{F}$, pričom pre aplikácie v kombinatorike je zďaleka najvýznamnejší prípad, keď ide o pole komplexných čísel. Hoci sme práve dokázali, že $(\mathbb{F}[[z]], +, \cdot, 0, 1)$ aj $(\mathbb{F}[z], +, \cdot, 0, 1)$ sú vždy obory integrity, nejde o polia – ľahko napríklad ukázať, že v $(\mathbb{F}[[z]], +, \cdot, 0, 1)$ neexistuje multiplikatívny inverzný prvok k radu z . Napriek tomu je často užitočné študovať rady, pre ktoré v algebre $(\mathbb{F}[[z]], +, \cdot, 0, 1)$ multiplikatívny inverzný prvok existuje.

Definícia 2.1.13. *Nech $\mathbb{F}$ je pole a $r \in \mathbb{F}[[z]] \setminus \{0\}$ je formálny mocninový rad. Multiplikatívnym inverzným prvkom radu r nazveme, ak existuje, rad $r^{-1} = r(z)^{-1}$ taký, že $r \cdot r^{-1} = r^{-1} \cdot r = 1$. Namiesto r^{-1} resp. $r(z)^{-1}$ budeme písať aj $1/r$ resp. $1/r(z)$.*

Bezprostredným dôsledkom asociatívnosti Cauchyho súčiny formálnych mocninových radov je jednoznačnosť radu r^{-1} , kedykoľvek tento rad pre $r \in \mathbb{F}[[z]]$ existuje; v obore integrity $\mathbb{F}[[z]]$ táto skutočnosť vyplýva aj z možnosti krátenia nenulovými radmi.

⁷Definíciu oborov integrity možno chápať aj ako výsledok snahy o abstrakciu niektorých najdôležitejších vlastností celých čísel [47].

Dokážeme teraz, že multiplikatívny inverzný prvok k radu $r \in \mathbb{F}[[z]]$ existuje práve vtedy, keď má tento rad nenulový konštantný koeficient.

Tvrdenie 2.1.14. *Nech $\mathbb{F}$ je pole a $r = \sum_{t \in \mathbb{N}} a_t z^t$ je formálny mocninový rad z $\mathbb{F}[[z]] \setminus \{0\}$. Rad r^{-1} potom existuje práve vtedy, keď $a_0 \neq 0$.*

Dôkaz. Nech $a_0 = 0$. Za účelom sporu predpokladajme, že rad r^{-1} existuje, pričom $r^{-1} = \sum_{t \in \mathbb{N}} b_t z^t$. Keďže $r \cdot r^{-1} = 1$, z definície Cauchyho súčinu dostávame

$$1 = [z^0](r \cdot r^{-1}) = \sum_{k=0}^0 a_k b_{0-k} = a_0 b_0 = 0 b_0;$$

zjavne ale neexistuje žiadne b_0 vyhovujúce rovnici $1 = 0 b_0$ – spor.

Predpokladajme teraz, že $a_0 \neq 0$. Skonstruujeme postupnosť koeficientov $b_0, b_1, b_2, \dots$ formálneho mocninového radu $s = \sum_{t \in \mathbb{N}} b_t z^t$ spĺňajúceho rovnosť $r \cdot s = 1$. Z komutatívnosti okruhu $\mathbb{F}[[z]]$ potom vyplynie aj rovnosť $s \cdot r = 1$; rad s teda bude hľadaným radom r^{-1} .

Aby bolo $r \cdot s = 1$, musí byť z definície Cauchyho súčinu splnená rovnosť

$$1 = [z^0](r \cdot s) = \sum_{k=0}^0 a_k b_{0-k} = a_0 b_0.$$

Keďže $a_0 \neq 0$, môžeme položiť $b_0 := 1/a_0$. Predpokladajme teraz, že máme skonštruované všetky koeficienty $b_0, b_1, \dots, b_t$ pre nejaké $t \in \mathbb{N}$. Aby potom bolo $r \cdot s = 1$, musí byť z definície Cauchyho súčinu splnená rovnosť

$$0 = [z^{t+1}](r \cdot s) = \sum_{k=0}^{t+1} a_k b_{t+1-k} = a_0 b_{t+1} + a_1 b_t + a_2 b_{t-1} + \dots + a_{t+1} b_0,$$

z čoho vyplýva, že možno položiť $b_{t+1} := -(a_1 b_t + a_2 b_{t-1} + \dots + a_{t+1} b_0)/a_0$; opäť pritom využívame nenulovosť koeficientu a_0 . Rad s je týmito vzťahmi jednoznačne určený a z vyššie uvedeného tiež vyplýva, že nutne $r \cdot s = 1$. $\square$

Príklad 2.1.15. Dôkaz predchádzajúceho tvrdenia poskytuje pre každý formálny mocninový rad r s nenulovým konštantným koeficientom návod na výpočet koeficientov radu r^{-1} . Ten možno využiť napríklad na odvodenie nasledujúcich mimoriadne užitočných vzorcov, ktorých *overenie* je ešte jednoduchšou úlohou (v nasledujúcom $a \in \mathbb{F}$ a $b \in \mathbb{F} \setminus \{0\}$ pre nejaké pole $\mathbb{F}$):

$$\begin{aligned} \frac{1}{1-z} &= 1 + z + z^2 + z^3 + \dots = \sum_{t \in \mathbb{N}} z^t, \\ \frac{1}{1-az} &= 1 + az + a^2 z^2 + a^3 z^3 + \dots = \sum_{t \in \mathbb{N}} a^t z^t, \\ \frac{1}{b-az} &= \frac{1}{b} + \frac{az}{b^2} + \frac{a^2 z^2}{b^3} + \frac{a^3 z^3}{b^4} + \dots = \sum_{t \in \mathbb{N}} \frac{a^t z^t}{b^{t+1}}. \end{aligned}$$

Podobnosť so vzorcami pre Maclaurinove rady, známymi z matematickej analýzy, nie je náhodná. Ku koncu oddielu ju vysvetlíme ako špeciálny prípad oveľa všeobecnejšieho fenoménu.

Multiplikatívne inverzné prvky teda existujú len pre rady $r(z)$ s nenulovým koeficientom $[z^0]r(z)$. To okrem iného znamená, že *podiel* dvoch radov je zatiaľ definovaný len v prípade, že má menovateľ nenulový konštantný koeficient. Občas je ale užitočné uvažovať aj niektoré iné podiely, ako napríklad

$$\frac{z + z^3 + z^5 + \dots}{z} = 1 + z^2 + z^4 + \dots,$$

alebo

$$\frac{z^2 - z^4}{z - z^2} = z + z^2.$$

Takéto podiely je možné definovať priamočiarym rozšírením definície multiplikatívnych inverzných prvkov – podiel $r(z)/s(z)$ teda bude definovaný práve vtedy, keď existuje formálny mocninový rad $q(z)$ taký, že $r(z) = s(z)q(z)$; v takom prípade ním bude práve tento rad $q(z)$.

Definícia 2.1.16. Nech $\mathbb{F}$ je pole a $r, s \in \mathbb{F}[[z]]$ sú formálne mocninové rady také, že $s \neq 0$. *Podielom* radov r a s nazveme, ak existuje, rad $r/s = r(z)/s(z)$ taký, že $r = (r/s) \cdot s$.

Z možnosti krátiť nenulovými prvkami oboru integrity opäť vyplýva jednoznačnosť takto definovaných podielov.

Kritérium existencie podielu dvojice formálnych mocninových radov nad poľom $\mathbb{F}$ využíva pojem *kostupňa* radu. Ten je pre ľubovoľný *nenulový* rad $r \in \mathbb{F}[[z]]$ definovaný ako prirodzené číslo

$$\text{codeg } r(z) = \min\{t \in \mathbb{N} \mid [z^t]r(z) \neq 0\}.$$

Tvrdenie 2.1.17. Nech $\mathbb{F}$ je pole a $r, s \in \mathbb{F}[[z]]$ sú formálne mocninové rady také, že $s \neq 0$. Rad r/s potom existuje práve vtedy, keď $r = 0$ alebo $r \neq 0$ a $\text{codeg } s(z) \leq \text{codeg } r(z)$.

Dôkaz. Pre $r = 0$ je evidentne $r/s = 0$. Po zvyšok dôkazu teda predpokladajme, že $r \neq 0$. Nech najprv $b = \text{codeg } s(z) \leq \text{codeg } r(z) = a$. Potom $r(z) = z^a \bar{r}(z)$ a $s(z) = z^b \bar{s}(z)$ pre nejakú dvojicu formálnych mocninových radov $\bar{r}, \bar{s} \in \mathbb{F}[[z]]$ takých, že $\text{codeg } \bar{r}(z) = \text{codeg } \bar{s}(z) = 0$. Podiel radov r a s je teda daný vzťahom

$$\frac{r(z)}{s(z)} = z^{a-b} \frac{\bar{r}(z)}{\bar{s}(z)},$$

ako možno overiť nasledujúcim výpočtom:

$$\left(z^{a-b} \frac{\bar{r}(z)}{\bar{s}(z)} \right) s(z) = z^{a-b} \frac{\bar{r}(z)}{\bar{s}(z)} z^b \bar{s}(z) = z^{a-b} z^b \bar{r}(z) = z^a \bar{r}(z) = r(z).$$

Opačnú implikáciu dokážeme sporom – predpokladajme, že existuje dvojica formálnych mocninových radov $r, s \in \mathbb{F}[[z]] \setminus \{0\}$ takých, že $b = \text{codeg } s(z) > \text{codeg } r(z) = a$ a súčasne existuje podiel r/s . Nech

$$s(z) = \sum_{t \in \mathbb{N}} b_t z^t$$

a

$$\frac{r(z)}{s(z)} = \sum_{t \in \mathbb{N}} q_t z^t.$$

Z definície podielu potom

$$\frac{r(z)}{s(z)} s(z) = r(z),$$

z čoho pre koeficient $[z^a]r(z)$, nenulový vďaka rovnosti $\text{codeg } r(z) = a$, dostávame

$$[z^a]r(z) = [z^a] \left(\frac{r(z)}{s(z)} s(z) \right) = \sum_{k=0}^a q_k b_{a-k} = 0,$$

keďže z $a < b = \text{codeg } s(z)$ vyplýva $b_0 = \dots = b_a = 0$. To samozrejme odporuje nenulovosti tohto koeficientu. $\square$

Dalšou operáciou, ktorú možno na formálnych mocninových radoch nad poľom uvažovať, je n -tá odmocnina pre nenulové $n \in \mathbb{N}$. Ľahko vidieť, že pre formálne mocninové rady $r(z)$ s $[z^0]r(z) = 0$ vo všeobecnosti *nemusí* existovať žiaden formálny mocninový rad $s(z)$, pre ktorý je $s(z)^n = r(z)$ – stačí vziať napríklad $n = 3$ a $r(z) = z^2$. Hoci pre iné dvojice n a $r(z)$ takýto rad $s(z)$ existovať môže, ide v takom prípade o relatívne náhodnú skutočnosť s neveľkým kombinatorickým významom. Odmocniny teda budeme definovať len pre formálne mocninové rady $r(z)$ s $[z^0]r(z) \neq 0$.

Nad *všeobecným poľom* sa pritom obmedzíme na rady s konštantným koeficientom rovným jednej, pre ktoré ukážeme existenciu práve jednej n -tej odmocniny, ktorej konštantný koeficient sa tiež rovná jednej – tú nazveme *kanonickou*. Nad poľom *komplexných čísel* už budeme uvažovať naozaj všetky rady s nenulovým konštantným koeficientom a ukážeme, že každý takýto rad r má presne n odmocnín rádu n . Ak je navyše $[z^0]r(z)$ kladné reálne číslo, je jednoznačne daná *kanonická* odmocnina radu $r(z)$, ktorá má tiež kladný reálny konštantný koeficient.

Začneme teda nad všeobecným poľom $\mathbb{F}$, a to definíciou kanonickej n -tej odmocniny formálneho mocninového radu $r(z)$ takého, že $[z^0]r(z) = 1$. Tou bude – ako dokážeme – jednoznačne určený formálny mocninový rad $s(z)$, pre ktorý platí $[z^0]s(z) = 1$ a zároveň $s(z)^n = r(z)$. Prezentácia dôkazu nasledujúceho tvrdenia je v zásade prebratá z [89].

Tvrdenie 2.1.18. *Nech $\mathbb{F}$ je pole a $r = \sum_{t \in \mathbb{N}} a_t z^t$ je formálny mocninový rad v $\mathbb{F}[[z]]$ taký, že $a_0 = 1$. Nech $n \in \mathbb{N} \setminus \{0\}$. Potom existuje jednoznačne určený formálny mocninový rad $s = \sum_{t \in \mathbb{N}} b_t z^t \in \mathbb{F}[[z]]$ taký, že $b_0 = 1$ a $s(z)^n = r(z)$.*

Dôkaz. Nech $q = \sum_{t \in \mathbb{N}} d_t z^t$ je ľubovoľný formálny mocninový rad v $\mathbb{F}[[z]]$ taký, že $d_0 = 1$. Matematickou indukciou vzhľadom na n potom možno dokázať, že ak $q^n = \sum_{t \in \mathbb{N}} c_t z^t$, tak $c_0 = 1$, $c_1 = nd_1$ a pre všetky prirodzené $t \geq 2$ platí $c_t = nd_t + p_{t,n}(d_1, \dots, d_{t-1})$, kde $p_{t,n}$ je nejaká polynomickeá funkcia o $t - 1$ premenných.

Ak teraz zo vzťahov $a_0 = 1$, $a_1 = nb_1$ a $a_t = nb_t + p_{t,n}(b_1, \dots, b_{t-1})$ pre $t \geq 2$ vyjadríme koeficienty $b_1, b_2, b_3, \dots$ a ak položíme $b_0 = 1$, z uvedeného pre $s = \sum_{t \in \mathbb{N}} b_t z^t$ vyplýva $s^n = r$. Keďže navyše existuje práve jedna postupnosť $(b_1, b_2, b_3, \dots)$ spĺňajúca vyššie uvedené vzťahy, je rad s určený jednoznačne. $\square$

Definícia 2.1.19. Nech $\mathbb{F}$ je pole a $r = \sum_{t \in \mathbb{N}} a_t z^t$ je formálny mocninový rad v $\mathbb{F}[[z]]$ taký, že $a_0 = 1$. Pre jednoznačne určený rad $s = \sum_{t \in \mathbb{N}} b_t z^t \in \mathbb{F}[[z]]$ taký, že $b_0 = 1$ a $s(z)^n = r(z)$ píšeme $s =: r^{1/n} = \sqrt[n]{r}$ resp. $s(z) =: r(z)^{1/n} = \sqrt[n]{r(z)}$, pričom tento rad nazývame *kanonickou n -tou odmocninou* radu $r(z)$. Pre $n = 2$ navyše píšeme aj $\sqrt{r} := \sqrt[n]{r}$ resp. $\sqrt{r(z)} := \sqrt[n]{r(z)}$.

Rozšírenie pojmu odmocniny na formálne mocninové rady $r(z) \in \mathbb{C}[[z]]$ s nenulovým konštantným koeficientom je už teraz priamočiarou záležitosťou.

Tvrdenie 2.1.20. *Nech $r = \sum_{t \in \mathbb{N}} a_t z^t$ je formálny mocninový rad v $\mathbb{C}[[z]]$ taký, že $a_0 \neq 0$. Nech $n \geq 1$ je prirodzené číslo a nech $\alpha_0, \dots, \alpha_{n-1}$ sú práve všetky n -té komplexné odmocniny čísla a_0 . Potom existuje práve n radov $s(z) \in \mathbb{C}[[z]]$ takých, že $s(z)^n = r(z)$, pričom tieto rady sú dané ako $s_0(z) = \alpha_0((1/a_0)r(z))^{1/n}$, $s_1(z) = \alpha_1((1/a_0)r(z))^{1/n}$, $\dots$, $s_{n-1}(z) = \alpha_{n-1}((1/a_0)r(z))^{1/n}$.*

Dôkaz. Ak má pre rad $s = \sum_{t \in \mathbb{N}} b_t z^t$ platiť $s^n = r$, zrejme musí byť aj $b_0^n = a_0$. Koeficient b_0 teda musí byť rovný jednému z čísel $\alpha_0, \dots, \alpha_{n-1}$. Nech $b_0 = \alpha_\ell$ pre nejaké $\ell \in \{0, \dots, n-1\}$. Potom $((1/\alpha_\ell)s(z))^n = (1/a_0)r(z)$, a teda $(1/\alpha_\ell)s(z) = ((1/a_0)r(z))^{1/n}$, z čoho $s(z) = \alpha_\ell((1/a_0)r(z))^{1/n}$. Pre každé $\ell \in \{0, \dots, n-1\}$ sme teda našli práve jeden rad $s(z)$ taký, že $s(z)^n = r(z)$, pričom žiadne ďalšie takéto rady neexistujú. Zostáva označiť rad $s(z)$ získaný pre to-ktoré ℓ ako $s_\ell(z)$. $\square$

Definícia 2.1.21. Nech $r = \sum_{t \in \mathbb{N}} a_t z^t$ je rad v $\mathbb{C}[[z]]$ taký, že $a_0 \neq 0$. Nech $n \geq 1$ je prirodzené číslo. Potom ľubovoľný z n existujúcich radov $s \in \mathbb{C}[[z]]$ takých, že $s(z)^n = r(z)$ nazveme *n -tou odmocninou* radu $r(z)$. Ak je navyše a_0 kladné reálne číslo, rad $s_0(z)$ daný ako $s_0(z) = \alpha_0((1/a_0)r(z))^{1/n}$, kde α_0 je kladnou reálnou odmocninou čísla a_0 , nazveme *kanonickou n -tou odmocninou* radu $r(z)$, pričom budeme písať $s_0 =: r^{1/n} = \sqrt[n]{r}$ resp. $s_0(z) =: r(z)^{1/n} = \sqrt[n]{r(z)}$. Pre $n = 2$ navyše píšeme aj $\sqrt{r} := \sqrt[n]{r}$ resp. $\sqrt{r(z)} := \sqrt[n]{r(z)}$.

Pre ľubovoľný rad $r(z)$ v $\mathbb{C}[[z]]$ s kladným reálnym $[z^0]r(z)$ sme definovali mocniny aj odmocniny prirodzeného rádu, ako aj multiplikatívny inverzný prvok. Všetky tieto operácie navyše navzájom „komutujú“ – ich vykonanie v inom poradí sa vo výsledku nijak neprejaví; priamočiary dôkaz prenechávame čitateľovi ako cvičenie 2 na konci kapitoly. Týmto pozorovaním je zaručená korektnosť nasledujúcej definície, ktorá spomenuté tri operácie využíva na zavedenie kanonickej *racionalnej mocniny* formálneho mocninového radu s kladným reálnym konštantným koeficientom.

Definícia 2.1.22. Nech $r = \sum_{t \in \mathbb{N}} a_t z^t$ je formálny mocninový rad v $\mathbb{C}[[z]]$ taký, že a_0 je kladné reálne číslo. Nech $m \geq 0$ a $n \geq 1$ sú prirodzené čísla. *Kanonicou m/n -tou mocninou* radu r nazveme rad $r^{m/n}$ resp. $r(z)^{m/n}$ definovaný ako $r^{m/n} := (r^m)^{1/n} = (r^{1/n})^m$. *Kanonicou $-m/n$ -tou mocninou* radu r nazveme rad $r^{-m/n}$ resp. $r(z)^{-m/n}$ definovaný ako $r^{-m/n} := (r^{m/n})^{-1} = (r^{-1})^{m/n}$.

Poznámka 2.1.23. Hoci vo väčšine kombinatorických aplikácií sú racionálne mocniny formálnych mocninových radov postačujúce, je možné definovať aj ich komplexné mocniny. Jeden z možných prístupov je na základe analógie s Maclaurinovými radmi – presnejšie s Newtonovou zovšeobecnenou binomickou vetou – *definovať* pre každé $\alpha \in \mathbb{C}$ formálny mocninový rad

$$(1+z)^\alpha := \sum_{t \in \mathbb{N}} \binom{\alpha}{t} z^t$$

a pomocou zloženia formálnych mocninových radov následne definovať komplexnú mocninu ľubovoľného radu $r \in \mathbb{C}[[z]]$ takého, že $[z^0]r(z) = 1$ – takýto rad r totiž možno vyjadriť ako $r = 1 + s$, kde $[z^0]s(z) = 0$; rad r^α potom možno definovať ako $(1+s)^\alpha$. Túto definíciu ďalej možno rozšíriť na všetky rady $r \in \mathbb{C}[[z]]$ také, že $[z^0]r(z)$ je kladné reálne – metóda je rovnaká ako pri racionálnych mocninách.

Takto definované komplexné mocniny radov sú konzistentné s našou definíciou racionálnych mocnín. Dôkaz tu robiť nebudeme – nie je ťažký, ale vyžadoval by si rozvinúť teóriu formálnych mocninových radov do väčšej miery, než to dovoľuje rámec tohto textu; napríklad by bolo potrebné zaviesť a preskúmať pojem *formálnej derivácie* mocninového radu, ako aj formálne obdoby Maclaurinových radov pre exponenciálne a logaritmické funkcie. Čitateľa teda na tomto mieste iba odkážeme na [89] alebo na kurz *analytickej a enumeratívnej kombinatoriky* [70].

Na formálnych mocninových radoch možno definovať aj viaceré ďalšie užitočné operácie (napríklad už spomenutú formálnu deriváciu) a na základe analógie s teóriou Maclaurinových radov možno definovať aj viaceré špeciálne rady – ako napríklad e^z , $\sin z$ a podobne – z ktorých možno vytvárať nové rady pomocou operácie zloženia mocninových radov.

Poznámka 2.1.24. Bolo spomenuté, že v teórii formálnych mocninových radov otázky konverencie nehrajú rolu. Dobre definovanými formálnymi mocninovými radmi sú napríklad aj $\sum_{t \in \mathbb{N}} t! z^t$ alebo $\sum_{t \in \mathbb{N}} 2^{2^t} z^t$, hoci Maclaurinove rady $\sum_{t=0}^{\infty} t! z^t$ a $\sum_{t=0}^{\infty} 2^{2^t} z^t$ majú nulový polomer konverencie, a teda nereprezentujú funkcie analytické v bode $z = 0$. Možno však dokázať, že obor integrity $\mathbf{H}_0$ všetkých funkcií komplexnej premennej analytických v bode $z = 0$ je izomorfný s určitým podokruhom oboru integrity $\mathbb{C}[[z]]$, pričom analytická funkcia s Maclaurinovým rozvojom $\sum_{t=0}^{\infty} a_t z^t$ sa v tomto izomorfizme zobrazí na $\sum_{t \in \mathbb{N}} a_t z^t$. Ľahko potom vidieť, že aj operácie ako racionálna mocnina alebo zloženie radov sa na $\mathbf{H}_0$ správajú rovnako ako na obraze $\mathbf{H}_0$ v $\mathbb{C}[[z]]$.

Dôsledkom tohto pozorovania napríklad je, že na výpočet koeficientov formálneho mocninového radu $r(z)$ zadaného pomocou „základných“ radov z a a pre $a \in \mathbb{C}$ a vyššie zavedených operácií, napríklad

$$r(z) = \frac{1 - \sqrt{1 - 4z}}{2z},$$

možno použiť štandardný vzorec na výpočet koeficientov Maclaurinovho radu známy z matematickej analýzy.

2.2 Vytvárajúce funkcie a ich použitie

Najmä z dôvodov opísaných na začiatku predchádzajúceho oddielu je pri analýze postupností objavujúcich sa v kombinatorike mimoriadne užitočnou technikou práca s *vytvárajúcimi* – alebo *generujúcimi* – *funkciami*. Pre danú postupnosť komplexných – a špeciálne aj prirodzených – čísel $(a_0, a_1, a_2, \dots)$ nazveme jej *vytvárajúcou funkciou*⁸ formálny mocninový rad $a(z) \in \mathbb{C}[[z]]$ daný ako

$$a(z) = a_0 z^0 + a_1 z^1 + a_2 z^2 + \dots = \sum_{t \in \mathbb{N}} a_t z^t.$$

Tento rad bol v predchádzajúcom oddiele definovaný práve ako postupnosť koeficientov $(a_0, a_1, a_2, \dots)$ – na formálnej úrovni teda *neexistuje žiaden rozdiel medzi postupnosťou a jej vytvárajúcou funkciou*. Rozlišovanie medzi týmito dvoma konceptmi má však veľký prínos z hľadiska metodiky – interpretovať postupnosť ako formálny mocninový rad znamená pracovať s ňou odlišným spôsobom a považovať za prirodzené odlišné operácie.

Ľubovoľnú postupnosť čísel – či už ide o enumeračnú postupnosť kombinatorickej triedy, alebo napríklad o postupnosť danú rekurentným vzťahom – tak môžeme chápať aj ako formálny mocninový rad, ktorý nazývame *vytvárajúcou funkciou* danej postupnosti; pôvodná postupnosť sa zas takto stáva *postupnosťou koeficientov* svojej vytvárajúcej funkcie. Operácie na formálnych mocninových radoch pritom vždy zodpovedajú určitým operáciám na ich postupnostiach koeficientov a často býva možné identifikovať aj ich kombinatorický význam (tzn. príslušnú operáciu na kombinatorických triedach).

Uvedme pár najdôležitejších vzťahov medzi operáciami na postupnostiach a ich náprotivkami na vytvárajúcich funkciách. Tabuľka 2.1 začína dvoma vytvárajúcimi funkciami $a(z)$ a $b(z)$ pre postupnosti $(a_0, a_1, a_2, \dots)$ resp. $(b_0, b_1, b_2, \dots)$. Ďalšie riadky hovoria o postupnostiach koeficientov vytvárajúcich funkcií získaných z funkcií $a(z)$ a $b(z)$ niektorými elementárnymi operáciami. Všetky tieto vzťahy sú zrejším dôsledkom definícií operácií na formálnych mocninových radoch z minulého oddielu.

Vytvárajúca funkcia	Postupnosť koeficientov
$a(z)$	$(a_0, a_1, a_2, \dots)$
$b(z)$	$(b_0, b_1, b_2, \dots)$
$c \cdot a(z)$ pre $c \in \mathbb{C}$	$(ca_0, ca_1, ca_2, \dots)$
$a(z) \pm b(z)$	$(a_0 \pm b_0, a_1 \pm b_1, a_2 \pm b_2, \dots)$
$a(z) \cdot b(z)$	$(c_0, c_1, c_2, \dots)$, kde $c_t = \sum_{k=0}^t a_k b_{t-k}$ pre všetky $t \in \mathbb{N}$
$\frac{a(z)-a_0}{z}$	$(a_1, a_2, a_3, \dots)$
$z \cdot a(z)$	$(0, a_0, a_1, a_2, \dots)$
$\frac{a(z)-a_0 z^0 - a_1 z^1 - \dots - a_{j-1} z^{j-1}}{z^j}$ pre $j \in \mathbb{N}$	$(a_j, a_{j+1}, a_{j+2}, \dots)$
$z^j \cdot a(z)$ pre $j \in \mathbb{N}$	$(\underbrace{0, 0, \dots, 0}_j, a_0, a_1, a_2, \dots)$
$a(z^j)$ pre $j \in \mathbb{N}$ také, že $j \geq 1$	$(a_0, \underbrace{0, 0, \dots, 0}_{j-1}, a_1, \underbrace{0, 0, \dots, 0}_{j-1}, a_2, \dots)$
$\vdots$	$\vdots$

Tabuľka 2.1: Vybrané operácie na formálnych mocninových radoch spolu so zodpovedajúcimi operáciami na postupnostiach koeficientov.

⁸Presnejšie ide o tzv. *obyčajnú* vytvárajúcu funkciu. S inými typmi vytvárajúcich funkcií v tomto texte pracovať nebudeme, a preto si dovoľíme toto drobné zjednodušenie terminológie. Samotný termín „vytvárajúca funkcia“ sa používa najmä z historických dôvodov – v minulosti sa pod vytvárajúcou funkciou často nerozumel *formálny* mocninový rad, ale rovno Maclaurinov rad reprezentujúci analytickú funkciu.

V tabuľke 2.2 sú ďalej zhrnuté postupnosti koeficientov niektorých význačných konkrétnych vytvárajúcich funkcií.⁹ Tento súhrn je značne minimalistický – existuje množstvo ďalších vytvárajúcich funkcií, ktoré je z hľadiska aplikácií v kombinatorike užitočné poznať.

Vytvárajúca funkcia	Postupnosť koeficientov
a pre $a \in \mathbb{C}$	$(a, 0, 0, \dots)$
z	$(0, 1, 0, 0, \dots)$
z^j pre $j \in \mathbb{N}$	$(\underbrace{0, 0, \dots, 0}_{j-1}, 1, 0, 0, \dots)$
$\frac{1}{1-z}$	$(1, 1, 1, \dots)$
$\frac{1}{1-az}$ pre $a \in \mathbb{C}$	$(a^0, a^1, a^2, \dots)$
$\frac{1}{b-az}$ pre $a \in \mathbb{C}$ a $b \in \mathbb{C} \setminus \{0\}$	$(\frac{a^0}{b^1}, \frac{a^1}{b^2}, \frac{a^2}{b^3}, \dots)$
$(1+z)^\alpha$ pre $\alpha \in \mathbb{C}$	$((\alpha)_0, (\alpha)_1, (\alpha)_2, \dots)$
$\vdots$	$\vdots$

Tabuľka 2.2: Niektoré konkrétne vytvárajúce funkcie.

Proces aplikácie vytvárajúcich funkcií v enumeratívnej kombinatorike väčšinou pozostáva z nasledujúcich dvoch fáz:

1. *Nájdenie vytvárajúcej funkcie* pre skúmanú postupnosť. Tým sa rozumie jej vyjadrenie pomocou „základných“ formálnych mocninových radov, akými sú napríklad rady a pre $a \in \mathbb{C}$, rad z a ďalšie, a pomocou známych operácií na formálnych mocninových radoch, ako sú napríklad tie zavedené v predchádzajúcom oddiele. Výstupom tejto fázy môže byť napríklad pozorovanie, že vytvárajúcou funkciou danej postupnosti $(a_0, a_1, a_2, \dots)$ je $a(z) = (1 - \sqrt{1 - 4z})/(2z)$.

Je známych viacero techník umožňujúcich tento proces hľadania vytvárajúcich funkcií do veľkej miery zmechanizovať. Napríklad pre postupnosti dané rekurentnými vzťahmi určitých typov býva nájdenie príslušnej vytvárajúcej funkcie často veľmi jednoduché. Existujú tiež rôzne formalizmy na špecifikáciu kombinatorických tried, pozostávajúcich z objektov ako napríklad stromy, grafy a podobne – tie sú často navrhnuté tak, aby bolo možné špecifikáciu danej kombinatorickej triedy mechanicky pretaviť do podoby vytvárajúcej funkcie jej enumeračnej postupnosti.

Medzi takéto špecifikačné mechanizmy patrí napríklad *symbolická metóda* opísaná v klasickej monografii Flajoleta a Sedgewicka [44], do veľkej miery inšpirovaná formálnymi gramatikami. Spomedzi niekoľkých alternatívnych, avšak v podstate veľmi podobných prístupov spomeňme takzvanú *Delestovu-Viennotovu-Schützenbergerovu metódu* spočívajúcu v zakódovaní objektov danej kombinatorickej triedy do slov nad určitou abecedou tak, aby dĺžka slova vždy odrážala veľkosť zakódovaného objektu. Následne sa pokračuje, ak je to možné, konštrukciou (obyčajne bezkontextovej) gramatiky generujúcej jazyk všetkých takýchto slov. Na základe danej bezkontextovej gramatiky $\mathcal{G}$ je už možné mechanicky skonštruovať vytvárajúcu funkciu pre postupnosť, ktorej t -ty člen udáva počet slov dĺžky t v jazyku generovanom gramatikou $\mathcal{G}$.

2. *Použitie nájdenej vytvárajúcej funkcie* na získanie informácií o skúmanej postupnosti. Niekedy sa napríklad môže podariť vyjadriť vytvárajúcu funkciu spôsobom umožňujúcim nájdenie vzorca pre t -ty člen zodpovedajúcej postupnosti v uzavretom tvare. To často nie je možné; pre omnoho väčšiu triedu vytvárajúcich funkcií je ale možné aplikovať metódy *analytickej kombinatoriky* na nájdenie asymptotického odhadu pre t -ty člen postupnosti a $t \rightarrow \infty$, prípadne aspoň nájsť ľubovoľný člen postupnosti štandardnými metódami na výpočet koeficientov Maclaurinovho radu. O ďalších aplikáciách vytvárajúcich funkcií sa možno dočítať napríklad v [44, 104].

⁹Treba prízvukovať, že v prípade vytvárajúcich funkcií $(1+z)^\alpha$ pre $\alpha \in \mathbb{C}$ ide o ich *definíciu*, ktorej konzistentnosť s racionálnymi mocninami sme postulovali, ale nedokázali.

Pokročilými metódami špecifikácie kombinatorických objektov a analytickými metódami umožňujúcimi z veľkej časti zmechanizovať proces nájdenia asymptotického odhadu pre počet kombinatorických objektov daného typu a veľkosti sa na tomto mieste zaoberať nebudeme; ide o hlavnú náplň kurzu *analytickej a enumeratívnej kombinatoriky* [70]. V nasledujúcom sa obmedzíme na jeden ukázkový príklad, v ktorom nájdeme vytvárajúcu funkciu postupnosti Fibonacciho čísel danej rekurentným vzťahom. Nájdenu vytvárajúcu funkciu následne využijeme – vďaka tomu, že bude „veľmi jednoduchého typu“ – na vyjadrenie t -teho Fibonacciho čísla v uzavretom tvare.

Príklad 2.2.1. Postupnosť Fibonacciho čísel $(F_t)_{t \in \mathbb{N}} = (F_0, F_1, F_2, \dots)$ je definovaná rekurentne:

$$\begin{aligned} F_0 &= 0, \\ F_1 &= 1, \\ F_{t+2} &= F_{t+1} + F_t \quad \text{pre všetky } t \in \mathbb{N}. \end{aligned} \tag{2.4}$$

Nech $r(z)$ je vytvárajúca funkcia postupnosti $(F_t)_{t \in \mathbb{N}}$. Potom

$$\frac{r(z) - F_0}{z} = \frac{r(z)}{z}$$

je vytvárajúca funkcia postupnosti $(F_{t+1})_{t \in \mathbb{N}}$ a

$$\frac{r(z) - F_0 - F_1 z}{z^2} = \frac{r(z) - z}{z^2}$$

je vytvárajúca funkcia postupnosti $(F_{t+2})_{t \in \mathbb{N}}$. Zo vzťahu (2.4) preto vyplýva

$$\frac{r(z) - z}{z^2} = \frac{r(z)}{z} + r(z),$$

z čoho úpravou pomocou operácií zavedených v predchádzajúcom oddiele dostávame

$$r(z) \cdot (1 - z - z^2) = z.$$

Teda

$$r(z) = \frac{z}{1 - z - z^2} = \frac{z}{(1 - \varphi z)(1 - \psi z)} = \frac{1}{\sqrt{5}} \left(\frac{1}{1 - \varphi z} - \frac{1}{1 - \psi z} \right),$$

kde

$$\varphi = \frac{1 + \sqrt{5}}{2} \quad \text{a} \quad \psi = \frac{1 - \sqrt{5}}{2}.$$

Vytvárajúcu funkciu $r(z)$ sme teda vyjadrili ako lineárnu kombináciu dvojice známych formálnych mocninových radov $1/(1 - \varphi z)$ a $1/(1 - \psi z)$. Z tabuliek 2.1 a 2.2 už teda možno vyčítať, že pre všetky $t \in \mathbb{N}$ je

$$F_t = [z^t]r(z) = [z^t] \frac{1}{\sqrt{5}} \left(\frac{1}{1 - \varphi z} - \frac{1}{1 - \psi z} \right) = \frac{1}{\sqrt{5}} (\varphi^t - \psi^t);$$

tým sme vyjadrili t -te Fibonacciho číslo v uzavretom tvare.

Metóda z predchádzajúceho príkladu je použiteľná na mnoho širšiu škálu rekurentne zadaných postupností. Obdobným spôsobom je možné, za predpokladu schopnosti faktorizovať polynómy ľubovoľného stupňa, vyjadriť v uzavretom tvare t -ty člen ľubovoľnej postupnosti $(a_t)_{t \in \mathbb{N}}$ danej lineárnou homogénnou rekurenciou s konštantnými koeficientmi – to jest rekurenciou typu

$$c_d a_{t+d} + c_{d-1} a_{t+d-1} + \dots + c_0 a_t = 0$$

pre všetky $t \in \mathbb{N}$, kde $d \geq 1$ je prirodzené číslo, $c_0, \dots, c_d$ sú komplexné konštanty, $c_d \neq 0$ a $a_0, \dots, a_{d-1}$ sú dané ako počiatočné podmienky. Zovšeobecnenie prístupu z príkladu 2.2.1 na takúto širšiu triedu rekurencií pre túto chvíľu prenechávame čitateľovi. V kapitole 6 však na rovnakej úrovni všeobecnosti podrobne preskúmame úzko súvisiaci prístup založený na lineárnej algebre.

2.3 Rady o niekoľkých komutujúcich premenných

Teóriu formálnych mocninových radov o jednej premennej z možno prirodzeným spôsobom rozšíriť na teóriu formálnych mocninových radov o m premenných $z_1, \dots, z_m$ pre ľubovoľné nenulové prirodzené číslo m . Takýmito radmi rozumieme formálne súčty typu

$$r(z_1, \dots, z_m) = \sum_{(t_1, \dots, t_m) \in \mathbb{N}^m} a_{t_1, \dots, t_m} z_1^{t_1} \dots z_m^{t_m}. \quad (2.5)$$

Pri tomto zápise sa rozumie samosebou, že premenné $z_1, \dots, z_m$ sú komutujúce: pre každú prípustnú dvojicu indexov i, j je $z_i z_j = z_j z_i$. Preto sa zvyčajne hovorí len o formálnych mocninových radoch o niekoľkých premenných a komutujúcich premenných $z_1, \dots, z_m$ sa explicitne nezdôrazňuje. My však v nasledujúcom oddiele práve od tejto zdanlivo nespochybniteľnej požiadavky upustíme a budeme pracovať s formálnymi mocninovými radmi o niekoľkých nekomutujúcich premenných – okrem iného tak získame užitočné zovšeobecnenie formálnych jazykov. Aby sme medzi oboma druhmi radov rozlíšili, budeme rady typu (2.5) nazývať formálnymi mocninovými radmi o niekoľkých *komutujúcich* premenných.

V enumeratívnej kombinatorike sa pracuje takmer výlučne s radmi o komutujúcich premenných – zavádza sa tiež pojem *vytvárajúcej funkcie niekoľkých premenných*. Tieto rady sú mimoriadne užitočné v situáciách, keď je okrem veľkosti kombinatorických objektov potrebné uvažovať ešte aspoň jeden ďalší parameter.

Na formálnych mocninových radoch o m komutujúcich premenných možno definovať operácie analogické tým z oddielu 2.1, pričom myšlienkovým východiskom sú tu opäť predovšetkým vlastnosti Maclaurinových radov funkcií viacerých premenných. My tak teraz činiť nebudeme a túto úlohu prenecháme čitateľovi ako cvičenie. Obmedzíme sa len na vyslovenie formálnej definície radov o niekoľkých komutujúcich premenných, ktorú možno – rovnako ako pre rady o jednej premennej – sformulovať pre ľubovoľnú množinu koeficientov.

Rady o jednej premennej s koeficientmi v množine S sme definovali prostredníctvom stotožnenia s postupnosťami ich koeficientov – ide teda o zobrazenia z $\mathbb{N}$ do S . Rady o m komutujúcich premenných definujeme ako zobrazenia z $\mathbb{N}^m$ do množiny S .

Definícia 2.3.1. Nech $m \geq 1$ je prirodzené číslo a S je množina. *Formálny mocninový rad o m komutujúcich premenných $z_1, \dots, z_m$ s koeficientmi v množine S je zobrazenie*

$$r: \mathbb{N}^m \rightarrow S.$$

Ak je pre všetky $t_1, \dots, t_m \in \mathbb{N}$ hodnota zobrazenia r pri argumentoch $(t_1, \dots, t_m)$ rovná $a_{t_1, \dots, t_m}$, píšeme

$$r = r(z_1, \dots, z_m) = \sum_{(t_1, \dots, t_m) \in \mathbb{N}^m} a_{t_1, \dots, t_m} z_1^{t_1} \dots z_m^{t_m}.$$

Koeficient $a_{t_1, \dots, t_m}$ pri $z_1^{t_1} \dots z_m^{t_m}$ označujeme aj

$$a_{t_1, \dots, t_m} =: [z_1^{t_1} \dots z_m^{t_m}] r(z_1, \dots, z_m) = [z_1^{t_1} \dots z_m^{t_m}] r$$

a od označenia $r(t_1, \dots, t_m)$ upúšťame.¹⁰ Koeficient $a_{0, \dots, 0}$ nazývame *konštantným*. Množinu všetkých formálnych mocninových radov o komutujúcich premenných $z_1, \dots, z_m$ s koeficientmi v množine S označujeme $S[[z_1, \dots, z_m]]$.

¹⁰Označenie takéhoto typu sa chápe, podobne ako v oddiele 2.1 pri formálnych mocninových radoch o jednej premennej, ako *dosadenie* za jednotlivé premenné $z_1, \dots, z_m$.

2.4 Rady o niekoľkých nekomutujúcich premenných

Automaty rozoznávajúce formálne jazyky nad abecedou Σ možno chápať ako „zariadenia“ realizujúce zobrazenia typu $\varphi: \Sigma^* \rightarrow \mathbb{B}$ – pre slovo $w \in \Sigma^*$ je potom $\varphi(w) = 1$ práve vtedy, keď w patrí do jazyka rozoznávaného daným automatom. Tento pohľad sa neobmedzuje na jazyky rozoznávané automatmi: *každý* jazyk $L \subseteq \Sigma^*$ je jednoznačne zadaný jeho *charakteristickou funkciou* $\chi_L: \Sigma^* \rightarrow \mathbb{B}$ danou ako $\chi_L(w) = 1$ pre všetky $w \in L$ a $\chi_L(w) = 0$ pre všetky $w \in \Sigma^* \setminus L$. Každé zobrazenie $\varphi: \Sigma^* \rightarrow \mathbb{B}$ možno naopak reprezentovať pomocou jeho *nosiča* $\text{supp}(\varphi) = \{w \in \Sigma^* \mid \varphi(w) \neq 0\}$, čo je jazyk nad Σ . Zjavne sme práve opísali bijekciu medzi jazykmi nad abecedou Σ a zobrazeniami zo Σ^* do $\mathbb{B}$.

Pri automatoch nás ale okrem príslušnosti slov do jazyka často zaujíma aj nejaká ďalšia informácia. Môžeme sa napríklad pýtať, koľko existuje behov automatu na danom slove. Užitočným modelovacím nástrojom tiež môžu byť automaty, ktorých prechodom sú priradené ceny; zaujímať nás môže napríklad cena najlacnejšieho behu na danom slove. V týchto a v mnohých ďalších prípadoch sa zdá byť prirodzené chápať automaty ako „zariadenia“ realizujúce zobrazenia typu $\varphi: \Sigma^* \rightarrow S$, kde S je nejaká nie nutne booleovská množina. Sledujúc túto myšlienku sa v nasledujúcej kapitole dostaneme k definícii pojmu *automatov s váhami*.

Zatiaľ však na automaty pozabudnime a skúmame vlastnosti samotných zobrazení $\varphi: \Sigma^* \rightarrow S$. Keby sme ako abecedu zvolili $\Sigma = \{z_1, \dots, z_m\}$, išlo by o zobrazenia priradujúce ľubovoľnému slovu nad touto abecedou – napríklad $z_1 z_1 z_2$ alebo $z_1 z_2 z_1$ – prvok množiny S . Jediným rozdielom oproti formálnym mocninovým radom z predchádzajúceho oddielu teda je, že „premenné“ $z_1, \dots, z_m$ chápeme ako *nekomutujúce* – $z_1 z_1 z_2$ a $z_1 z_2 z_1$ sú rôzne slová a nie ten istý objekt $z_1^2 z_2$. Zobrazenia $\varphi: \Sigma^* \rightarrow S$ preto budeme chápať ako *formálne mocninové rady s nekomutujúcimi premennými* – napriek tomu, že sa už v takýchto radoch kvôli nekomutujúcosti premenných žiadne mocniny nevyskytujú.

Čitateľ si na tomto mieste pravdepodobne kladie otázku, z akého dôvodu budeme namiesto dobre známeho pojmu zobrazenia pracovať s akousi čudesnou analógiou formálnych mocninových radov, v ktorej už ani žiadne mocniny prítomné nie sú. Dôvod za touto voľbou spočíva v skutočnosti, že sa snažíme nájsť vhodné zovšeobecnenie nielen pre *množinu* všetkých formálnych jazykov nad abecedou Σ , ale aj pre *algebru* týchto jazykov – to jest polokruh s operáciami zjednotenia a zreťazenia. Ak vyjadríme definíciu zreťazenia jazykov $L, K \subseteq \Sigma^*$ pomocou ich charakteristických funkcií $\chi_L, \chi_K: \Sigma^* \rightarrow \mathbb{B}$, tak pre všetky $w \in \Sigma^*$ je

$$\chi_{L \cdot K}(w) = \bigvee_{\substack{u, v \in \Sigma^* \\ uv = w}} (\chi_L(u) \wedge \chi_K(v)).$$

Vidíme, že zreťazenie jazykov nezodpovedá bežnému súčinu zobrazení, ale skôr nápadne pripomína Cauchyho súčin mocninových radov. Prirodzenou multiplikatívnou operáciou na nami uvažovaných zobrazeniach teda nie je klasické násobenie funkcií, ale *Cauchyho súčin* formálnych mocninových radov s nekomutujúcimi premennými, ktorý možno chápať ako *zovšeobecnenie zreťazenia jazykov*.

V nasledujúcom teda načneme problematiku formálnych mocninových radov o niekoľkých nekomutujúcich premenných. Hoci ich definíciu opäť sformulujeme pre koeficienty z ľubovoľnej množiny S , zaujímavý pre nás bude predovšetkým prípad, keď je S polokruh. Vtedy totiž bude možné prirodzene definovať súčet a Cauchyho súčin formálnych mocninových radov s nekomutujúcimi premennými.

Definícia 2.4.1. Nech Σ je abeceda a S je množina. *Formálny mocninový rad o nekomutujúcich premenných z abecedy Σ s koeficientmi v množine S je zobrazenie*

$$r: \Sigma^* \rightarrow S.$$

Pre $w \in \Sigma^*$ potom namiesto $r(w)$ píšeme (r, w) a tento prvok nazývame *koeficientom* slova w v rade r . Pre samotný rad r píšeme

$$r = \sum_{w \in \Sigma^*} (r, w)w.$$

Koeficient (r, ε) nazývame *konštantným*. Množinu všetkých formálnych mocninových radov o nekomutujúcich premenných zo Σ s koeficientmi v S označujeme $S\langle\langle \Sigma^* \rangle\rangle$.

Poznámka 2.4.2. Často budeme namiesto o „formálnych mocninových radoch o nekomutujúcich premenných zo Σ s koeficientmi v S “ hovoriť len o *formálnych mocninových radoch nad abecedou Σ a množinou S* , prípadne len o *formálnych mocninových radoch nad Σ a S* .

Poznámka 2.4.3. Formálne mocninové rady o niekoľkých nekomutujúcich premenných sú, rovnako ako rady o niekoľkých komutujúcich premenných, zovšeobecnením formálnych mocninových radov o jednej premennej z – tie možno považovať aj za rady nad unárnou abecedou $\Sigma = \{z\}$. Pre nás však bude omnoho podstatnejší fakt, že formálne mocninové rady nad Σ a S sú *zovšeobecnením jazykov*¹¹ nad abecedou Σ , keďže tie možno chápať ako rady nad Σ a $\mathbb{B}$.

Zamerajme sa teraz na formálne mocninové rady s *koeficientmi v polokruhu* – v takom prípade môžeme definovať súčet a Cauchyho súčin radov prirodzeným zovšeobením definícií z oddielu 2.1.

Definícia 2.4.4. Nech Σ je abeceda a S je polokruh. Nech $r = \sum_{w \in \Sigma^*} (r, w)w$ a $s = \sum_{w \in \Sigma^*} (s, w)w$ sú formálne mocninové rady z $S\langle\langle \Sigma^* \rangle\rangle$. *Súčet* radov r a s je potom formálny mocninový rad $r + s$ nad Σ a S taký, že pre všetky $w \in \Sigma^*$ je

$$(r + s, w) = (r, w) + (s, w)$$

a *Cauchyho súčin* radov r a s je formálny mocninový rad $r \cdot s$ nad Σ a S taký, že pre všetky $w \in \Sigma^*$ je

$$(r \cdot s, w) = \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} (r, u)(s, v).$$

Poznámka 2.4.5. Operácie súčtu a Cauchyho súčinu radov nad Σ a S sú zovšeobenieniami operácií zjednotenia a zretazenia jazykov nad abecedou Σ .

Nosičom formálneho mocninového radu r nad Σ a S nazveme jazyk $\text{supp}(r) = \{w \in \Sigma^* \mid (r, w) \neq 0\}$. Prvok a polokruhu S budeme stotožňovať s formálnym mocninovým radom r takým, že $(r, \varepsilon) = a$ a $(r, w) = 0$ pre všetky $w \in \Sigma^+$. Podobne každé slovo $w \in \Sigma^*$ budeme stotožňovať s radom r takým, že $(r, w) = 1$ a $(r, x) = 0$ pre všetky $x \in \Sigma^* \setminus \{w\}$. Ľahko potom vidieť, že konečným počtom aplikácií súčtu a Cauchyho súčinu možno z radov a pre $a \in S$ a w pre $w \in \Sigma^*$ získať práve všetky formálne mocninové rady r nad Σ a S také, že $\text{supp}(r)$ je konečná množina. Takéto rady nazveme *polynómami* nad Σ a S ; množinu všetkých polynómov nad Σ a S označíme $S\langle \Sigma^* \rangle$. Pre ľubovoľný jazyk $L \subseteq \Sigma^*$ navyše označíme ako $S\langle\langle L \rangle\rangle$ množinu všetkých formálnych mocninových radov $r \in S\langle\langle \Sigma^* \rangle\rangle$ takých, že $\text{supp}(r) \subseteq L$ a ako $S\langle L \rangle$ množinu všetkých polynómov $p \in S\langle \Sigma^* \rangle$ takých, že $\text{supp}(p) \subseteq L$. Špeciálne napríklad $S\langle\langle \Sigma^+ \rangle\rangle$ označuje množinu všetkých formálnych mocninových radov nad Σ a S s nulovým konštantným koeficientom, $S\langle \Sigma \rangle$ je množina všetkých polynómov typu $a_1c_1 + \dots + a_kc_k$ pre $k \in \mathbb{N}$, $a_1, \dots, a_k \in S$ a $c_1, \dots, c_k \in \Sigma$, a podobne. Keďže prvky $S\langle\langle \{\varepsilon\} \rangle\rangle = S\langle \{\varepsilon\} \rangle$ stotožňujeme s prvkami S , budeme stotožňovať aj samotné tieto množiny. *Umocňovanie* radov nad Σ a S definujeme štandardne: pre $r \in S\langle\langle \Sigma^* \rangle\rangle$ kladieme $r^0 = 1$ a $r^{t+1} = r^t \cdot r$ pre všetky $t \in \mathbb{N}$.

Je jednoduchým cvičením dokázať nasledujúce tvrdenie, podľa ktorého množiny $S\langle\langle \Sigma^* \rangle\rangle$ a $S\langle \Sigma^* \rangle$ tvoria s operáciami súčtu a Cauchyho súčinu opäť polokruh. Ľahko tiež vidieť, že pre unárnou abecedu $\Sigma = \{z\}$ sú tieto polokruhy izomorfné polokruhom $S[[z]]$ resp. $S[z]$.

Tvrdenie 2.4.6. Nech Σ je abeceda a S je polokruh. Algebry $(S\langle\langle \Sigma^* \rangle\rangle, +, \cdot, 0, 1)$ a $(S\langle \Sigma^* \rangle, +, \cdot, 0, 1)$ sú potom tiež polokruhy.

Nekonečné súčty formálnych mocninových radov nad Σ a S vo všeobecnosti definovať nemožno. Vždy ale možno definovať súčty cez *lokálne konečné* systémy radov, ktoré definujeme rovnakým spôsobom ako pri radoch o jednej premennej: systém radov nad Σ a S je lokálne konečný, ak je v tomto systéme pre každé $w \in \Sigma^*$ konečne veľa radov s nenulovým koeficientom pri w . Nekonečný súčet cez lokálne konečný systém radov je možné definovať po zložkách – pre každé $w \in \Sigma^*$ je koeficient nekonečného súčtu pri w daný konečným súčtom.

¹¹Formálne mocninové rady o niekoľkých nekomutujúcich premenných sa preto občas v literatúre nazývajú aj *jazykmi s váhami* [80].

Definícia 2.4.7. Nech Σ je abeceda a S je polokruh. Systém $(r_i \mid i \in I)$ formálnych mocninových radov z $S\langle\langle\Sigma^*\rangle\rangle$ je *lokálne konečný*, ak je pre každé $w \in \Sigma^*$ konečná množina $I(w) := \{i \in I \mid (r_i, w) \neq 0\}$.

Definícia 2.4.8. Nech Σ je abeceda, S je polokruh a $(r_i \mid i \in I)$ je lokálne konečný systém formálnych mocninových radov z $S\langle\langle\Sigma^*\rangle\rangle$. *Súčet systému* radov $(r_i \mid i \in I)$ potom definujeme ako

$$\sum_{i \in I} r_i = r,$$

kde r je formálny mocninový rad v $S\langle\langle\Sigma^*\rangle\rangle$ taký, že pre všetky $w \in \Sigma^*$ je

$$(r, w) = \sum_{i \in I(w)} (r_i, w).$$

Po zložkách tiež možno definovať nekonečné súčty cez ľubovoľný (resp. ľubovoľný konečný alebo spočítateľne nekonečný) systém formálnych mocninových radov nad úplným (resp. spočítateľne úplným) polokruhom.

Definícia 2.4.9. Nech Σ je abeceda, S je úplný (resp. spočítateľne úplný) polokruh a $(r_i \mid i \in I)$ je ľubovoľný (resp. ľubovoľný konečný alebo spočítateľne nekonečný) systém formálnych mocninových radov z $S\langle\langle\Sigma^*\rangle\rangle$. *Súčet systému* radov $(r_i \mid i \in I)$ potom definujeme ako

$$\sum_{i \in I} r_i = r,$$

kde r je formálny mocninový rad v $S\langle\langle\Sigma^*\rangle\rangle$ taký, že pre všetky $w \in \Sigma^*$ je

$$(r, w) = \sum_{i \in I} (r_i, w).$$

Poznámka 2.4.10. Definície 2.4.8 a 2.4.9 sú konzistentné v nasledujúcom zmysle: ak $(r_i \mid i \in I)$ je (spočítateľný) lokálne konečný systém radov nad (spočítateľne) úplným polokruhom – to je jediný prípad, kde možno aplikovať obe definície – tak pre všetky $w \in \Sigma^*$ je

$$\sum_{i \in I(w)} (r_i, w) = \sum_{i \in I} (r_i, w);$$

podľa oboch definícií teda dostaneme rovnaký súčet.¹² Postupným použitím podmienok (ii) a (iii) definície 1.4.2 totiž skutočne dostávame

$$\begin{aligned} \sum_{i \in I} (r_i, w) &= \sum_{i \in I(w) \cup (I \setminus I(w))} (r_i, w) = \sum_{i \in I(w)} (r_i, w) + \sum_{i \in I \setminus I(w)} 0 = \sum_{i \in I(w)} (r_i, w) + \sum_{i \in I \setminus I(w)} 0 \cdot 0 = \\ &= \sum_{i \in I(w)} (r_i, w) + 0 \cdot \sum_{i \in I \setminus I(w)} 0 = \sum_{i \in I(w)} (r_i, w). \end{aligned}$$

Dôkaz nasledujúceho tvrdenia prenechávame čitateľovi ako jednoduché cvičenie.

Tvrdenie 2.4.11. Nech Σ je abeceda a S je úplný (resp. spočítateľne úplný) polokruh. S nekonečnými súčtami z definície 2.4.9 potom $(S\langle\langle\Sigma^*\rangle\rangle, +, \cdot, 0, 1)$ tiež tvorí úplný (resp. spočítateľne úplný) polokruh.

¹²Vďaka podmienke (i) definície 1.4.2 si môžeme dovoliť nerozlišovať notáciu pre súčty cez konečné indexové množiny v ľubovoľnom polokruhu (kde ide iba o skrátený zápis pre niekoľko aplikácií operácie $+$) a v spočítateľne úplnom polokruhu (kde ide o hodnotu zobrazenia Φ na konečnom systéme).

Nekonečné súčty možno za určitých podmienok využiť na definíciu *iterácie* formálneho mocninového radu zovšeobecňujúcej iteráciu jazykov. Iteráciu radu r nad Σ a S by sme chceli definovať ako

$$r^* = \sum_{t \in \mathbb{N}} r^t.$$

Súčet na pravej strane nemusí byť vždy dobre definovaný. Ukazuje sa však, že pre rady z množiny $S\langle\langle\Sigma^+\rangle\rangle$ – t. j. rady s nulovým konštantným koeficientom – ide zakaždým o súčet cez lokálne konečný systém radov, a teda ho možno interpretovať v zmysle definície 2.4.8 (to dokážeme v tvrdení 2.4.12, ktoré je analógiou tvrdenia 2.1.7). Podobne je tento súčet dobre definovaný pre všetky rady z $S\langle\langle\Sigma^*\rangle\rangle$, kde S je úplný alebo aspoň spočítateľne úplný polokruh – v tomto prípade ho možno interpretovať v zmysle definície 2.4.9. Vďaka poznámke 2.4.10 sú navyše obidve definície navzájom konzistentné.

Tvrdenie 2.4.12. *Nech Σ je abeceda, S je polokruh a $r \in S\langle\langle\Sigma^+\rangle\rangle$. Potom $(r^t, w) = 0$ pre každé $t \in \mathbb{N}$ a všetky $w \in \Sigma^*$ také, že $|w| < t$; systém radov $(r^t \mid t \in \mathbb{N})$ je teda lokálne konečný.*

Dôkaz. Tvrdenie je triviálne pravdivé pre $t = 0$. Nech teraz tvrdenie platí pre $t = q$ a uvažujme $t = q + 1$. Ak $|w| < q + 1$, tak

$$(r^{q+1}, w) = (r^q \cdot r, w) = \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} (r^q, u)(r, v) = (r^q, w)(r, \varepsilon) + \sum_{\substack{u \in \Sigma^*, v \in \Sigma^+ \\ uv = w}} (r^q, u)(r, v) = 0,$$

keďže $(r, \varepsilon) = 0$ a v súčte

$$\sum_{\substack{u \in \Sigma^*, v \in \Sigma^+ \\ uv = w}} (r^q, u)(r, v)$$

musí pre každé u byť $|u| < q$, vďaka čomu z indukčného predpokladu dostávame $(r^q, u) = 0$. $\square$

Definícia 2.4.13. Nech Σ je abeceda, S je polokruh a $r \in S\langle\langle\Sigma^+\rangle\rangle$. *Iteráciou radu r nazveme rad r^* nad Σ a S definovaný predpisom*

$$r^* = \sum_{t \in \mathbb{N}} r^t.$$

Definícia 2.4.14. Nech Σ je abeceda, S je spočítateľne úplný polokruh a $r \in S\langle\langle\Sigma^*\rangle\rangle$. *Iteráciou radu r nazveme rad r^* nad Σ a S definovaný predpisom*

$$r^* = \sum_{t \in \mathbb{N}} r^t.$$

Nasledujúce málo prekvapivé tvrdenie ukazuje, že aj pre súčty cez lokálne konečné systémy radov možno sformulovať obdoby vlastností (i) až (iii) z definície spočítateľne úplných polokruhových.

Tvrdenie 2.4.15. *Nech Σ je abeceda a S je polokruh. Potom:*

(i) *Lubovoľný konečný systém $(r_i \mid i \in I)$ radov z $S\langle\langle\Sigma^*\rangle\rangle$ s $I = \{i_1, \dots, i_k\}$ je lokálne konečný, pričom $\sum_{i \in I} r_i = r_{i_1} + \dots + r_{i_k}$.*

(ii) *Ak $(r_i \mid i \in I)$ je lokálne konečný systém radov z $S\langle\langle\Sigma^*\rangle\rangle$ a $(I_j \mid j \in J)$ je zovšeobecnený rozklad množiny I , sú lokálne konečné aj systémy $(r_i \mid i \in I_j)$ pre všetky $j \in J$, ako aj systém $(\sum_{i \in I_j} r_i \mid j \in J)$, pričom*

$$\sum_{j \in J} \left(\sum_{i \in I_j} r_i \right) = \sum_{i \in I} r_i.$$

(iii) *Ak $(r_i \mid i \in I)$ je lokálne konečný systém radov z $S\langle\langle\Sigma^*\rangle\rangle$ a $r \in S\langle\langle\Sigma^*\rangle\rangle$, sú lokálne konečné aj systémy $(rr_i \mid i \in I)$ a $(r_i r \mid i \in I)$, pričom*

$$r \left(\sum_{i \in I} r_i \right) = \sum_{i \in I} rr_i \quad \text{a} \quad \left(\sum_{i \in I} r_i \right) r = \sum_{i \in I} r_i r.$$

Dôkaz. Jednoduché dôkazy tvrdení (i) a (ii) prenechávame čitateľovi ako cvičenie. Dokážme zostávajúce tvrdenie (iii).

Označme pre každé $x \in \Sigma^*$ ako $I_1(x)$ konečnú množinu $I_1(x) = \{i \in I \mid (r_i, x) \neq 0\}$ a ako $I_2(x)$ množinu $I_2(x) = \{i \in I \mid (rr_i, x) \neq 0\}$; samotné označenie $I(x)$ používané vyššie by v tomto prípade bolo nejednoznačné. Zrejme

$$I_2(x) \subseteq \{i \in I \mid \exists u, v \in \Sigma^* : uv = x \wedge (r, u) \neq 0 \wedge i \in I_1(v)\} =: I_3(x). \quad (2.6)$$

Množina $I_2(x)$ je teda konečná a systém $(rr_i \mid i \in I)$ je lokálne konečný. Pre všetky $i \in I_3(x) \setminus I_2(x)$ navyše

$$\sum_{\substack{u, v \in \Sigma^* \\ uv = x}} (r, u)(r_i, v) = 0. \quad (2.7)$$

Pre každé $w \in \Sigma^*$ potom vďaka distributívnosti konečných súčtov vzhľadom na násobenie dostávame

$$\begin{aligned} \left(r \left(\sum_{i \in I} r_i \right), w \right) &= \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} (r, u) \left(\sum_{i \in I} r_i, v \right) = \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} (r, u) \sum_{i \in I_1(v)} (r_i, v) = \\ &= \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} \sum_{i \in I_1(v)} (r, u)(r_i, v) = \sum_{i \in I_3(w)} \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} (r, u)(r_i, v) \stackrel{(2.6), (2.7)}{=} \\ &\stackrel{(2.6), (2.7)}{=} \sum_{i \in I_2(w)} \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} (r, u)(r_i, v) = \sum_{i \in I_2(w)} (rr_i, w) = \\ &= \left(\sum_{i \in I} rr_i, w \right) \end{aligned}$$

a podobne možno argumentovať aj pre distributívnosť sprava. $\square$

2.5 Matice formálnych mocninových radov

Čitateľovi prenechávame ako jedno z cvičení na konci kapitoly ľahký dôkaz nasledujúceho tvrdenia: pre každý polokruh S a prirodzené číslo n tvorí množina $S^{n \times n}$ všetkých štvorcových matíc typu $n \times n$ nad S , spolu s operáciami súčtu a obvyklého súčinu matíc v zmysle definície 1.3.7, opäť polokruh. Ak je navyše polokruh S (spočítateľne) úplný, je (spočítateľne) úplným aj polokruh $S^{n \times n}$.

Pre každú abecedu Σ , polokruh S a prirodzené číslo n označujeme ako $(S\langle\langle\Sigma^*\rangle\rangle)^{n \times n}$ množinu všetkých matíc typu $n \times n$, ktorých prvkami sú formálne mocninové rady nad Σ a S . Ide o polokruh, ktorý je evidentne izomorfný s polokruhom $S^{n \times n}\langle\langle\Sigma^*\rangle\rangle$ formálnych mocninových radov nad Σ s koeficientmi v $S^{n \times n}$ – izomorfizmus týchto dvoch polokruhov priradí matici $A = (r_{i,j})_{n \times n} \in (S\langle\langle\Sigma^*\rangle\rangle)^{n \times n}$ rad $r \in S^{n \times n}\langle\langle\Sigma^*\rangle\rangle$ taký, že pre všetky $w \in \Sigma^*$ je $(r, w) = ((r_{i,j}, w))_{n \times n}$. Tieto dva polokruhy budeme stotožňovať a pre množinu $(S\langle\langle\Sigma^*\rangle\rangle)^{n \times n}$ matíc typu $n \times n$ nad $S\langle\langle\Sigma^*\rangle\rangle$ tak niekedy budeme používať kratšie označenie $S^{n \times n}\langle\langle\Sigma^*\rangle\rangle$. Rovnako nebudeme rozlišovať ani medzi množinami $(S\langle\langle\Sigma^+\rangle\rangle)^{n \times n}$ a $S^{n \times n}\langle\langle\Sigma^+\rangle\rangle$, medzi množinami $(S\langle\Sigma^*\rangle)^{n \times n}$ a $S^{n \times n}\langle\Sigma^*\rangle$, atď.

Každú maticu mocninových radov teda budeme chápať aj ako mocninový rad s maticovými koeficientmi a naopak. To nám okrem iného umožňuje hovoriť aj o *lokálne konečných systémoch matíc* formálnych mocninových radov – pôjde jednoducho o systémy, ktoré sú lokálne konečné ak ich interpretujeme ako systémy formálnych mocninových radov s maticovými koeficientmi.

Spomenuli sme, že pre (spočítateľne) úplný polokruh S je (spočítateľne) úplným aj polokruh $S\langle\langle\Sigma^*\rangle\rangle$. Pre každú maticu $A \in S^{n \times n}\langle\langle\Sigma^*\rangle\rangle$ potom možno uvažovať jej iteráciu A^* (definícia 1.4.5), pre ktorú navyše platia vzťahy dané lemov 1.5.5, dôsledkom 1.5.6 a dôsledkom 1.5.7. Teraz už len naznačíme možnosť rozvinutia analogickej teórie pre „bezepsilonové“ matice nad všeobecným polokruhom – teda pre matice $A \in S^{n \times n}\langle\langle\Sigma^+\rangle\rangle$, kde S je ľubovoľný polokruh.

Dôsledok 2.5.1. *Nech Σ je abeceda, S je polokruh, n je prirodzené číslo a $A \in S^{n \times n} \langle\langle \Sigma^+ \rangle\rangle$ je matica. Potom je systém matic $(A^t \mid t \in \mathbb{N})$ lokálne konečný.*

Dôkaz. Ako sme pozorovali vyššie – a ako aj vyplýva z použitej notácie $A \in S^{n \times n} \langle\langle \Sigma^+ \rangle\rangle$ – je každá matica mocninových radov v podstate aj mocninovým radom s maticovými koeficientmi. Vieme tiež, že $S^{n \times n}$ je polokruh. Stačí sa teda odvolať na tvrdenie 2.4.12. $\square$

Dôsledok 2.5.1 poskytuje základ pre nasledujúcu definíciu iterácie štvorcovej matice nad $S \langle\langle \Sigma^+ \rangle\rangle$.

Definícia 2.5.2. *Nech Σ je abeceda, S je polokruh, n je prirodzené číslo a $A \in S^{n \times n} \langle\langle \Sigma^+ \rangle\rangle$. Iteráciou matice A nazveme maticu A^* definovanú predpisom*

$$A^* = \sum_{t \in \mathbb{N}} A^t.$$

Nasledujúce tvrdenia možno dokázať skoro rovnako ako k nim analogické tvrdenia pre spočítateľne úplné polokruhy dokázané v prvej kapitole. Dôkazy preto prenechávame čitateľovi.

Veta 2.5.3. *Nech Σ je abeceda, S je polokruh, $n \in \mathbb{N}$ a $\mathcal{G} = ([n], \varphi)$ je ohodnotený graf nad $S \langle\langle \Sigma^+ \rangle\rangle$. Nech $\mathbf{A}(\mathcal{G})^* = (a_{i,j}[*])_{n \times n}$. Potom pre $i, j = 1, \dots, n$ je $(\varphi(\gamma) \mid \gamma \in \mathcal{W}_{i,j}(\mathcal{G}))$ lokálne konečný systém formálnych mocninových radov a*

$$a_{i,j}[*] = \sum_{\gamma \in \mathcal{W}_{i,j}(\mathcal{G})} \varphi(\gamma).$$

Dôkaz. Lokálnu konečnosť systému možno ľahko dokázať s použitím faktu, že pre všetky $t \in \mathbb{N}$ je množina sledov $\mathcal{W}_{i,j}(\mathcal{G}, t)$ konečná. Zvyšok rovnako ako pre vetu 1.4.6. $\square$

Lema 2.5.4. *Nech Σ je abeceda, S je polokruh, $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad $S \langle\langle \Sigma^+ \rangle\rangle$. Nech*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

je blokový rozklad matice A , kde $A_{1,1} = (a_{i,j})_{n_1 \times n_1}$, $A_{1,2} = (a_{i,n_1+j})_{n_1 \times n_2}$, $A_{2,1} = (a_{n_1+i,j})_{n_2 \times n_1}$ a $A_{2,2} = (a_{n_1+i,n_1+j})_{n_2 \times n_2}$ pre nejaké $n_1, n_2 \in \mathbb{N}$. Potom

$$A^* = \begin{pmatrix} A_{1,1}[*] & A_{1,2}[*] \\ A_{2,1}[*] & A_{2,2}[*] \end{pmatrix},$$

pričom jednotlivé bloky $A_{1,1}[] = (a_{i,j}[*])_{n_1 \times n_1}$, $A_{1,2}[*] = (a_{i,n_1+j}[*])_{n_1 \times n_2}$, $A_{2,1}[*] = (a_{n_1+i,j}[*])_{n_2 \times n_1}$ a $A_{2,2}[*] = (a_{n_1+i,n_1+j}[*])_{n_2 \times n_2}$ sú dané nasledujúcimi vzťahmi:*

$$\begin{aligned} A_{1,1}[*] &= (A_{1,1} + A_{1,2} A_{2,2}^* A_{2,1})^*, \\ A_{1,2}[*] &= (A_{1,1} + A_{1,2} A_{2,2}^* A_{2,1})^* A_{1,2} A_{2,2}^*, \\ A_{2,1}[*] &= (A_{2,2} + A_{2,1} A_{1,1}^* A_{1,2})^* A_{2,1} A_{1,1}^*, \\ A_{2,2}[*] &= (A_{2,2} + A_{2,1} A_{1,1}^* A_{1,2})^*. \end{aligned}$$

Dôkaz. Rovnako ako pre lemu 1.5.5. $\square$

Dôsledok 2.5.5. *Nech Σ je abeceda, S je polokruh, $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad $S \langle\langle \Sigma^+ \rangle\rangle$ s blokovým rozkladom*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ \mathbf{0} & A_{2,2} \end{pmatrix},$$

kde $A_{1,1} = (a_{i,j})_{n_1 \times n_1}$, $A_{1,2} = (a_{i,n_1+j})_{n_1 \times n_2}$ a $A_{2,2} = (a_{n_1+i,n_1+j})_{n_2 \times n_2}$ pre nejaké $n_1, n_2 \in \mathbb{N}$. Potom

$$A^* = \begin{pmatrix} A_{1,1}^* & A_{1,1}^* A_{1,2} A_{2,2}^* \\ \mathbf{0} & A_{2,2}^* \end{pmatrix}.$$

Dôkaz. Rovnako ako pre dôsledok 1.5.6. $\square$

Dôsledok 2.5.6. *Nech Σ je abeceda, S je polokruh, $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad $S\langle\langle \Sigma^+ \rangle\rangle$ s blokovým rozkladom*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ \mathbf{0} & A_{2,2} & A_{2,3} \\ \mathbf{0} & \mathbf{0} & A_{3,3} \end{pmatrix},$$

kde $A_{1,1} = (a_{i,j})_{n_1 \times n_1}$, $A_{1,2} = (a_{i,n_1+j})_{n_1 \times n_2}$, $A_{1,3} = (a_{i,n_1+n_2+j})_{n_1 \times n_3}$, $A_{2,2} = (a_{n_1+i,n_1+j})_{n_2 \times n_2}$, $A_{2,3} = (a_{n_1+i,n_1+n_2+j})_{n_2 \times n_3}$ a $A_{3,3} = (a_{n_1+n_2+i,n_1+n_2+j})_{n_3 \times n_3}$ pre nejaké $n_1, n_2, n_3 \in \mathbb{N}$. Potom

$$A^* = \begin{pmatrix} A_{1,1}^* & A_{1,1}^* A_{1,2} A_{2,2}^* & A_{1,1}^* A_{1,3} A_{3,3}^* + A_{1,1}^* A_{1,2} A_{2,2}^* A_{2,3} A_{3,3}^* \\ \mathbf{0} & A_{2,2}^* & A_{2,2}^* A_{2,3} A_{3,3}^* \\ \mathbf{0} & \mathbf{0} & A_{3,3}^* \end{pmatrix}.$$

Dôkaz. Rovnako ako pre dôsledok 1.5.7. □

Poznámka 2.5.7. Takáto prezentácia, pri ktorej sme dvakrát vyslovili „rovnakú“ postupnosť tvrdení s „rovnakými“ dôkazmi, má od ideálnej ďaleko – hoci ide o viac-menej štandardný prístup. V oddiele 3.13 uvidíme, že v skutočnosti možno obidve paralelné teórie získať ako špeciálne prípady jedinej zjednocujúcej teórie, ktorá je však o poznanie menej intuitívna. Toto pozorovanie odkladáme na koniec tretej kapitoly úmyselne: čitateľovi, ktorý sa nevyžíva v prílišnej abstrakcii, bude umožnené oddiel 3.13 preskočiť; naopak čitateľ, ktorý sa rozhodne oddiel 3.13 preštudovať, bude motivovaný príslubom odbreneenia od mechanického dokazovania, ktoré mu práve bolo prenechané.

2.6 Iterácia súčtu a iterácia súčinu

Než prejdeme k nasledujúcej kapitole a teórii automatov s váhami, dokážeme ešte identity pre iterácie súčtu a súčinu, ktoré sa nám neskôr zídu. Urobíme tak ako pre prvky spočítateľne úplných polokruhov, tak aj pre formálne mocninové rady o nekomutujúcich premenných s nulovým konštantným koeficientom – teda v obidvoch kontextoch, v ktorých zatiaľ máme iteráciu definovanú. Dokážeme teda dvakrát po dve identity; začneme s identitami platnými v *spočítateľne úplných* (a teda aj úplných) polokruhoch.

Veta 2.6.1. *Nech S je spočítateľne úplný polokruh a nech $a, b \in S$. Potom $(a + b)^* = (a^* b)^* a^*$.*

Dôkaz. Je jednoduchým cvičením dokázať, že pre ľubovoľnú dvojicu formálnych jazykov L_1, L_2 platí

$$(L_1 \cup L_2)^* = (L_1^* L_2)^* L_1^*. \quad (2.8)$$

Uvažujme jazyky $A = \{\alpha\}$ a $B = \{\beta\}$ a homomorfizmus monoidov¹³ $h: (\{\alpha, \beta\}^*, \cdot, \varepsilon) \rightarrow (S, \cdot, 1)$ taký, že $h(\alpha) = a$ a $h(\beta) = b$. Platnosť nasledujúcich vzťahov pre všetky $t \in \mathbb{N}$ možno ľahko dokázať indukciou:

$$(a + b)^t = \sum_{w \in (A \cup B)^t} h(w), \quad (2.9)$$

$$a^t = h(\alpha^t) = \sum_{w \in A^t} h(w). \quad (2.10)$$

¹³Pripomeňme, že *monoid* je trojica $(M, \cdot, 1)$, kde M je množina, $\cdot$ je asociatívna binárna operácia na M a 1 je neutrálny prvok vzhľadom na $\cdot$; namiesto $(M, \cdot, 1)$ často píšeme len $(M, \cdot)$ alebo M . *Homomorfizmus monoidov* $(M, \cdot, 1)$ a $(N, \circ, e)$ je zobrazenie $h: M \rightarrow N$ také, že $h(1) = e$ a pre všetky $x, y \in M$ je $h(x \cdot y) = h(x) \circ h(y)$. Nie je ťažké dokázať, že homomorfizmus $h: (\Sigma^*, \cdot, \varepsilon) \rightarrow (M, \cdot, 1)$, kde Σ je abeceda a M je monoid, je jednoznačne určený obrazmi jednotlivých písmen abecedy Σ ; každé zobrazenie písmen zo Σ do M naopak možno rozšíriť na homomorfizmus. To je charakteristická vlastnosť *voľného monoidu* nad množinou generátorov Σ .

Triviálne tiež

$$b = h(\beta) = \sum_{w \in B} h(w). \quad (2.11)$$

Môžeme teraz pristúpiť k dôkazu samotnej identity $(a+b)^* = (a^*b)^*a^*$. Netriviálne kroky nasledujúceho odvodenia využívajú identity (2.8) až (2.11) a axiómy (i) až (iii) definície 1.4.2. To bude pre lepšiu čitateľnosť zakaždým zaznačené nad príslušným znamienkom rovnosti.

$$\begin{aligned} (a+b)^* &= \sum_{t \in \mathbb{N}} (a+b)^t \stackrel{(2.9)}{=} \sum_{t \in \mathbb{N}} \sum_{w \in (A \cup B)^t} h(w) \stackrel{(ii)}{=} \sum_{w \in (A \cup B)^*} h(w) \stackrel{(2.8)}{=} \\ &\stackrel{(2.8)}{=} \sum_{w \in (A^*B)^*A^*} h(w) \stackrel{(i) \text{ až } (iii)}{=} \left(\sum_{u \in (A^*B)^*} h(u) \right) \left(\sum_{v \in A^*} h(v) \right) \stackrel{(ii)}{=} \\ &\stackrel{(ii)}{=} \left(\sum_{t \in \mathbb{N}} \sum_{u \in (A^*B)^t} h(u) \right) \left(\sum_{t \in \mathbb{N}} \sum_{v \in A^t} h(v) \right) \stackrel{(i) \text{ až } (iii)}{=} \\ &\stackrel{(i) \text{ až } (iii)}{=} \left(\sum_{t \in \mathbb{N}} \left[\left(\sum_{x \in A^*} h(x) \right) \left(\sum_{y \in B} h(y) \right) \right]^t \right) \left(\sum_{t \in \mathbb{N}} \sum_{v \in A^t} h(v) \right) \stackrel{(ii)}{=} \\ &\stackrel{(ii)}{=} \left(\sum_{t \in \mathbb{N}} \left[\left(\sum_{t' \in \mathbb{N}} \sum_{x \in A^{t'}} h(x) \right) \left(\sum_{y \in B} h(y) \right) \right]^t \right) \left(\sum_{t \in \mathbb{N}} \sum_{v \in A^t} h(v) \right) \stackrel{(2.10), (2.11)}{=} \\ &\stackrel{(2.10), (2.11)}{=} \left(\sum_{t \in \mathbb{N}} \left[\left(\sum_{t' \in \mathbb{N}} a^{t'} \right) b \right]^t \right) \left(\sum_{t \in \mathbb{N}} a^t \right) = \left(\sum_{t \in \mathbb{N}} [a^*b]^t \right) a^* = (a^*b)^*a^*. \end{aligned}$$

Tým je veta dokázaná. $\square$

Veta 2.6.2. *Nech S je spočítateľne úplný polokruh a nech $a, b \in S$. Potom $(ab)^* = 1 + a(ba)^*b$.*

Dôkaz. Z axióm (i) až (iii) definície spočítateľne úplných polokruhových dostávame

$$(ab)^* = \sum_{t \in \mathbb{N}} (ab)^t = 1 + \sum_{\substack{t \in \mathbb{N} \\ t \geq 1}} (ab)^t = 1 + \sum_{\substack{t \in \mathbb{N} \\ t \geq 1}} a(ba)^{t-1}b = 1 + a \left(\sum_{t \in \mathbb{N}} (ba)^t \right) b = 1 + a(ba)^*b,$$

čím je veta dokázaná. $\square$

Nášim cieľom teraz bude dokázať dvojicu analogických identít pre formálne mocninové rady, ktoré sú poväčšine prvkami $S\langle\langle \Sigma^+ \rangle\rangle$ pre ľubovoľný polokruh S a ľubovoľnú abecedu Σ . V nich sa budú vyskytovať výrazy ako $(r+s)^*$ pre $r, s \in S\langle\langle \Sigma^+ \rangle\rangle$ alebo $(rs)^*$ pre $r \in S\langle\langle \Sigma^* \rangle\rangle$ a $s \in S\langle\langle \Sigma^+ \rangle\rangle$ (alebo naopak), ako aj výraz $(r^*s)^*$ pre $r, s \in S\langle\langle \Sigma^+ \rangle\rangle$, kde rad r^* už prvkom $S\langle\langle \Sigma^+ \rangle\rangle$ nie je. Nasledujúce tvrdenie je zárukou, že všetky tieto výrazy sú iteráciami prvkov $S\langle\langle \Sigma^+ \rangle\rangle$, a teda dávajú zmysel.

Tvrdenie 2.6.3. *Nech Σ je abeceda, S je polokruh, $r, s \in S\langle\langle \Sigma^+ \rangle\rangle$ a $r_\varepsilon \in S\langle\langle \Sigma^* \rangle\rangle$. Potom rady $r+s$, $r_\varepsilon r$ a rr_ε sú prvkami $S\langle\langle \Sigma^+ \rangle\rangle$.*

Dôkaz. Keďže $r, s \in S\langle\langle \Sigma^+ \rangle\rangle$, je $(r, \varepsilon) = (s, \varepsilon) = 0$. Preto $(r+s, \varepsilon) = (r, \varepsilon) + (s, \varepsilon) = 0 + 0 = 0$. Podobne tiež $(r_\varepsilon r, \varepsilon) = (r_\varepsilon, \varepsilon)(r, \varepsilon) = (r_\varepsilon, \varepsilon)0 = 0$ a $(rr_\varepsilon, \varepsilon) = (r, \varepsilon)(r_\varepsilon, \varepsilon) = 0(r_\varepsilon, \varepsilon) = 0$. $\square$

Než prejdeme k samotným identitám pre iteráciu súčtu a iteráciu súčinu, dokážeme tri pomocné lemy o formálnych mocninových radoch s nekomutujúcimi premennými. Sčasti pritom budeme sledovať prezentáciu z [29].

Lema 2.6.4. *Nech Σ je abeceda, S je polokruh, $r \in S\langle\langle\Sigma^+\rangle\rangle$. Potom $r^* = 1 + rr^*$.*

Dôkaz. Z časti (iii) tvrdenia 2.4.15 dostávame

$$1 + rr^* = 1 + r \sum_{t \in \mathbb{N}} r^t = 1 + \sum_{t \in \mathbb{N}} rr^t = 1 + \sum_{t \in \mathbb{N}} r^{t+1} = \sum_{t \in \mathbb{N}} r^t = r^*,$$

čo bolo treba dokázať. □

Lema 2.6.5. *Nech Σ je abeceda, S je polokruh, $r, s \in S\langle\langle\Sigma^*\rangle\rangle$ a aspoň jeden z radov r, s patrí do $S\langle\langle\Sigma^+\rangle\rangle$. Potom $(rs)^*r = r(sr)^*$.*

Dôkaz. Znenie lemy dáva zmysel vďaka tvrdeniu 2.6.3. Vďaka časti (iii) tvrdenia 2.4.15 navyše

$$(rs)^*r = \left(\sum_{t \in \mathbb{N}} (rs)^t \right) r = \sum_{t \in \mathbb{N}} (rs)^t r = \sum_{t \in \mathbb{N}} r(sr)^t = r \sum_{t \in \mathbb{N}} (sr)^t = r(sr)^*,$$

čím je dôkaz lemy dokončený. □

Nasledujúce užitočné tvrdenie sa v literatúre objavuje pod názvom *Ardenova lema* [94, 95].

Lema 2.6.6 (Arden). *Nech Σ je abeceda, S je polokruh, $r \in S\langle\langle\Sigma^+\rangle\rangle$ a $s \in S\langle\langle\Sigma^*\rangle\rangle$. Potom má rovnica $X = rX + s$ o neznámej X práve jedno riešenie v $S\langle\langle\Sigma^*\rangle\rangle$, ktorým je rad $\varrho = r^*s$.*

Dôkaz. Rad $\varrho = r^*s$ je riešením $X = rX + s$, pretože podľa lemy 2.6.4 je

$$\varrho = r^*s = (1 + rr^*)s = rr^*s + s = r\varrho + s.$$

Zostáva dokázať, že ϱ je jediné riešenie. Pre ľubovoľné riešenie ϱ' a každé $n \in \mathbb{N} \setminus \{0\}$ ale musí byť

$$\varrho' = r\varrho' + s = r^2\varrho' + rs + s = \dots = r^n\varrho' + r^{n-1}s + \dots + rs + s.$$

Ak teda pre ľubovoľné pevne dané $w \in \Sigma^*$ vezmeme $n = |w| + 1$, z tvrdenia 2.4.12 dostávame

$$\begin{aligned} (\varrho', w) &= (r^{|w|+1}\varrho' + r^{|w|}s + \dots + rs + s, w) = (r^{|w|+1}\varrho', w) + (r^{|w|}s + \dots + rs + s, w) = \\ &= (r^{|w|}s + \dots + rs + s, w) + \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} (r^{|w|+1}, u)(\varrho', v) = \\ &= (r^{|w|}s + \dots + rs + s, w) + \sum_{\substack{u, v \in \Sigma^* \\ uv = w}} 0(\varrho', v) = (r^{|w|}s + \dots + rs + s, w). \end{aligned}$$

Koeficient (ϱ', w) je teda rovnaký pre všetky riešenia ϱ' ; keďže je $w \in \Sigma^*$ ľubovoľné, existuje iba jedno riešenie $\varrho' = \varrho$. □

Môžeme teraz dokázať samotné analógie viet 2.6.1 a 2.6.2 pre formálne mocninové rady z $S\langle\langle\Sigma^+\rangle\rangle$.

Veta 2.6.7. *Nech Σ je abeceda, S je polokruh a $r, s \in S\langle\langle\Sigma^+\rangle\rangle$. Potom $(r + s)^* = (r^*s)^*r^*$.*

Dôkaz. Znenie vety dáva zmysel vďaka tvrdeniu 2.6.3. Z Ardenovej lemy ďalej vyplýva, že jediným riešením rovnice

$$X = (r + s)X + 1 \tag{2.12}$$

je rad $(r + s)^*1 = (r + s)^*$. Stačí teda ukázať, že aj $(r^*s)^*r^*$ je riešením rovnice (2.12). S využitím lemy 2.6.5 a lemy 2.6.4 ale dostávame

$$\begin{aligned} (r + s)(r^*s)^*r^* + 1 &= (r + s)r^*(sr^*)^* + 1 = rr^*(sr^*)^* + sr^*(sr^*)^* + 1 = rr^*(sr^*)^* + (sr^*)^* = \\ &= (1 + rr^*)(sr^*)^* = r^*(sr^*)^* = (r^*s)^*r^*, \end{aligned}$$

čím je dôkaz vety dokončený. □

Veta 2.6.8. *Nech Σ je abeceda, S je polokruh, $r, s \in S\langle\langle\Sigma^*\rangle\rangle$ a aspoň jeden z radov r, s patrí do $S\langle\langle\Sigma^+\rangle\rangle$. Potom $(rs)^* = 1 + r(sr)^*s$.*

Dôkaz. Znenie vety dáva zmysel vďaka tvrdeniu 2.6.3. Z lemy 2.6.5 a lemy 2.6.4 ďalej dostávame

$$1 + r(sr)^*s = 1 + rs(rs)^* = (rs)^*,$$

čo dokazuje naše tvrdenie. □

Cvičenia

1. Dokážte tvrdenie 2.1.3.
2. Dokážte, že operácie n -tej mocniny pre ľubovoľné prirodzené n , kanonickej m -tej odmocniny pre ľubovoľné prirodzené $m \geq 1$ a zobrazenia na multiplikatívny inverzný prvok v $\mathbb{C}[[z]]$ navzájom „komutujú“: pre každé

$$x, y \in \mathbb{N} \cup \{1/m \mid m \in \mathbb{N}; m \geq 1\} \cup \{-1\}$$

a ľubovoľný formálny mocninový rad $r \in \mathbb{C}[[z]]$ s kladným reálnym $[z^0]r(z)$ platí $(r^x)^y = (r^y)^x$. Týmto pozorovaním bola v texte odôvodnená korektnosť definície kanonickej racionálnej mocniny formálneho mocninového radu s kladným reálnym konštantným koeficientom (definícia 2.1.22).

3. Rozšírte definície operácií z oddielu 2.1 na rady o niekoľkých komutujúcich premenných.
4. Dokážte tvrdenie 2.4.6.
5. Dokážte tvrdenie 2.4.11.
6. Dokážte časti (i) a (ii) tvrdenia 2.4.15.
7. Dokážte nasledujúce tvrdenie používané v texte: pre každý polokruh S a každé $n \in \mathbb{N}$ tvorí množina $S^{n \times n}$ všetkých štvorcových matíc typu $n \times n$ nad S , spolu s operáciami sčítania a obvyklého násobenia matíc v zmysle definície 1.3.7, opäť polokruh; ten je (spočítateľne) úplný kedykoľvek je (spočítateľne) úplný polokruh S .
8. Dokážte vetu 2.5.3, lemu 2.5.4, dôsledok 2.5.5 a dôsledok 2.5.6.
9. Využite tvrdenie 2.4.15 na alternatívny dôkaz viet 2.6.7 a 2.6.8, pri ktorom sa postupuje obdobne ako pri dôkazoch viet 2.6.1 a 2.6.2.

Poznámky a odkazy na samoštúdium

Teória formálnych mocninových radov o jednej premennej je na matematicky rigoróznejšej úrovni spracovaná v [89, 81, 53]; možno tiež odporúčať učebnicu [18]. Množstvo kombinatorických aplikácií vytvárajúcich funkcií možno nájsť v knihe [104]. Spomedzi analytickejšie orientovanej literatúry, v ktorej ide predovšetkým o súvis vytvárajúcich funkcií s analytickými funkciami komplexnej premennej, spomeňme aspoň [44].

Formálne mocninové rady o jednej premennej, formálne mocninové rady o niekoľkých komutujúcich premenných a formálne mocninové rady o niekoľkých nekomutujúcich premenných možno skúmať v jednotnom rámci formálnych mocninových radov nad *ohodnoteným monoidom* (angl. *graded monoid*). Rady o m nekomutujúcich premenných sú potom radmi nad *voľným monoidom* s m generátormi a rady o m komutujúcich premenných sú radmi nad *voľným komutatívnym monoidom* s m generátormi. Voľný monoid s jedným generátorom je komutatívny; to súhlasí s tým, že rady o jednej premennej možno chápať ako rady o niekoľkých komutujúcich premenných aj ako rady o niekoľkých nekomutujúcich premenných. Ďalším druhom ohodnotených monoidov sú napríklad súčiny voľných monoidov, na ktorých stojí teória prekladačov. Čitateľa so záujmom o takýto všeobecnejší pohľad na formálne mocninové rady možno odkázať predovšetkým na Sakarovitchovu učebnicu [94].

Ardenova lema je pomenovaná podľa D. N. Ardena, ktorý ju však dokázal len pre formálne jazyky [6]. Pomenovanie „Ardenova lema“ sa napriek tomu v literatúre používa aj pre všeobecnejšiu verziu tohto výsledku, ktorú sme sformulovali ako lemu 2.6.6 – čitateľa tu odkazujeme na [94, 95].

Vety 2.6.7 a 2.6.8 sa dajú sformulovať o niečo všeobecnejšie. Vo vete 2.6.7 napríklad nie je nutné predpokladať príslušnosť do $S\langle\langle\Sigma^+\rangle\rangle$ pre obidva rady r, s – stačí tento predpoklad urobiť pre rady r, r^*s a $r+s$. Podobne vo vete 2.6.8 je možné nahradiť predpoklad príslušnosti aspoň jedného z radov r, s do $S\langle\langle\Sigma^+\rangle\rangle$ o niečo slabším predpokladom príslušnosti rs do $S\langle\langle\Sigma^+\rangle\rangle$. Ľubovoľný predpoklad príslušnosti radu do $S\langle\langle\Sigma^+\rangle\rangle$ je navyše možné nahradiť slabším predpokladom príslušnosti nejakej jeho mocniny do $S\langle\langle\Sigma^+\rangle\rangle$. Podrobnosti možno nájsť v [29].

Kapitola 3

Automaty s váhami

Využijeme teraz aparát z predošlej kapitoly a položíme základy teórie *automatov s váhami* – presnejšie *konečných automatov s váhami nad polokruhom*. Pôjde o zovšeobecnenie klasických nedeterministických konečných automatov rozoznávajúcich formálne jazyky, v ktorom každému prechodu navyše zodpovedá váha z nejakého polokruhu. Táto váha môže reprezentovať kvantitu súvisiacu s vykonaním daného prechodu: napríklad jeho multiplicitu (resp. počet spôsobov, ktorými možno na dané písmeno prejsť medzi danou dvojicou stavov), cenu, pravdepodobnosť, spoľahlivosť, užitočnosť, atď. Váhami môžu byť aj formálne jazyky – to znamená, že aj prekladače možno chápať ako automaty s váhami.

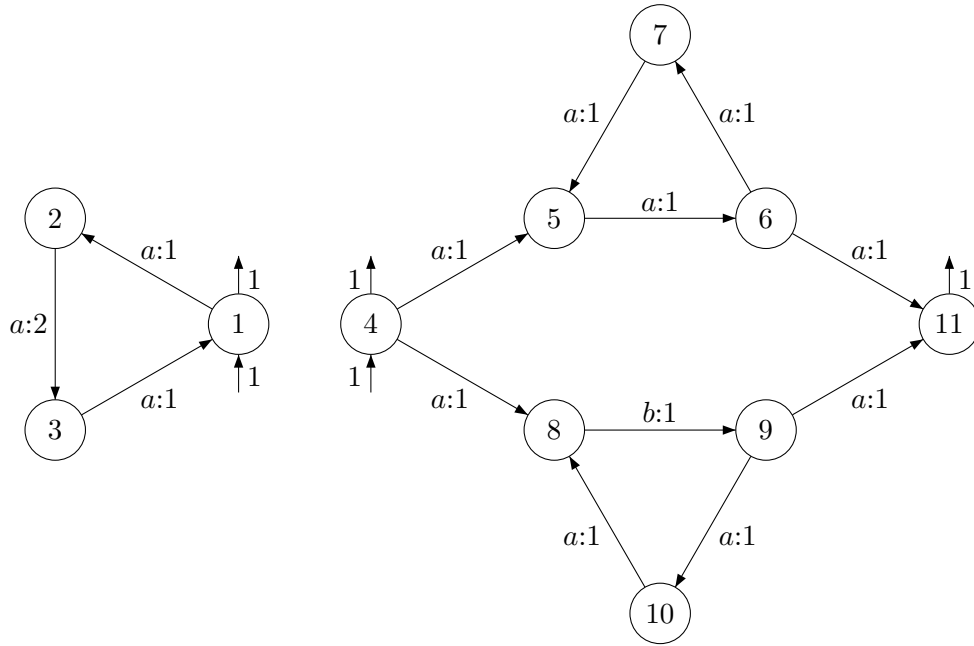
Takéto automaty už neslúžia len na rozoznávanie jazykov, ale priradujú slovám nad abecedou Σ kvantitu z polokruhu váh S . Automaty s váhami teda *realizujú formálne mocninové rady* nad Σ a S . Ku každému behu v automate $\mathcal{A}$ možno priradiť prvok polokruhu S postupným *prenásobením* počiatocnej váhy prvého stavu behu, váh všetkých prechodov behu v poradí od prvého po posledný a napokon koncovej váhy posledného stavu behu. Po *sčítaní* takto získaných hodnôt cez všetky behy na slove w dostávame koeficient slova w v rade realizovanom automatom $\mathcal{A}$.

Ak teda napríklad „obyčajný“ nedeterministický konečný automat reinterpretovejme ako automat s váhami nad polokruhom $(\mathbb{N}, +, \cdot, 0, 1)$, v ktorom má každý prechod jednotkovú váhu, je koeficientom každého slova w v rade realizovanom týmto automatom *počet behov* na slove w v pôvodnom automate. Nad polokruhom $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ je zas koeficientom slova w v realizovanom mocninovom rade *cena najlacnejšieho behu* na slove w . Automaty s váhami nad booleovským polokruhom $(\mathbb{B}, \vee, \wedge, 0, 1)$ napokon zodpovedajú – po stotožnení izomorfných polokruhov $\mathbb{B}\langle\langle\Sigma^*\rangle\rangle$ a 2^{Σ^*} – klasickým nedeterministickým konečným automatom.

3.1 Príklady na úvod

Hoci sme automaty s váhami ešte formálne nezadefinovali, v nasledujúcom tento model demonštrujeme na troch ukážkových príkladoch – tie by mali byť intuitívne pochopiteľné len na základe vysvetlenia vyššie. Čitateľ nespokojný s týmto postupom môže prejsť priamo k formálnym definíciám v oddiele 3.2 a neskôr sa prípadne k nasledujúcim príkladom vrátiť.

Príklad 3.1.1. Na obrázku 3.1 je znázornený diagram konečného automatu $\mathcal{A}$ nad abecedou $\Sigma = \{a, b\}$ s váhami nad polokruhom $(\mathbb{N}, +, \cdot, 0, 1)$; operácie sčítania a násobenia sú štandardné. Ako obyčajne sú stavy znázornené „kolečkami“ a prechody šípkami medzi nimi. Pri každom prechode je navyše značka typu $c:k$, kde $c \in \Sigma$ reprezentuje prečítaný symbol a $k \in \mathbb{N}$ váhu prechodu. Prechody s nulovou váhou v danom polokruhu – v tomto prípade ide dokonca o „zvyčajnú“ nulu – sa obvykle neuvažujú: sú totiž ekvivalentné chýbajúcemu prechodu. Šípky vedúce do jednotlivých stavov „odnikiaľ“ sú označené počiatocnou váhou daného stavu; pokiaľ sa v diagrame pri niektorom stave takáto šípka nevyskytuje, považuje sa jeho počiatocná váha za nulovú. Podobne šípky vedúce zo stavov „nikam“ sú označené koncovými váhami stavov a chýbajúca šípka tohto typu zodpovedá nulovej koncovej váhe.



Obr. 3.1: Konečný automat $\mathcal{A}$ s váhami nad polokruhom $(\mathbb{N}, +, \cdot, 0, 1)$.

Keby boli všetky prechody automatu $\mathcal{A}$ ohodnotené váhou 1, bolo by ho možné interpretovať ako bežný nedeterministický konečný automat rozoznávajúci nejaký jazyk, pre ktorý sa navyše „pýtame“ aj na počty behov na jednotlivých slovách vedúcich z počiatočného do koncového stavu. Prechod zo stavu 2 na písmeno a do stavu 3 má ale váhu 2; to sa dá interpretovať tak, že medzi touto dvojicou stavov možno na písmeno a prejsť dvoma spôsobmi.

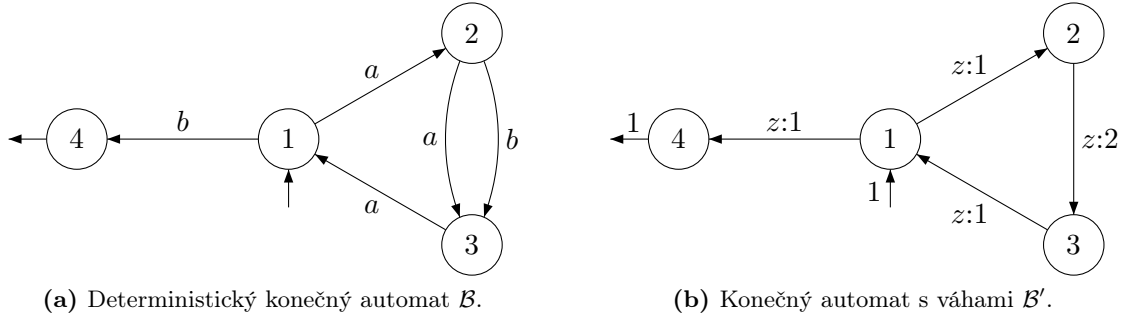
Beh automatu $\mathcal{A}$ sa musí začať, pokiaľ má prispieť nenulovou hodnotou ku koeficientu niektorého slova v realizovanom rade, v stave 1 alebo v stave 4. Behy začínajúce v stave 1 budú úspešné (skončia sa v stave s nenulovou koncovou váhou) len v prípade, že sa počas nich prečíta slovo $(aaa)^t$ pre nejaké $t \in \mathbb{N}$ – pre každé $t \in \mathbb{N}$ prispeje jediný takýto beh ku koeficientu slova $(aaa)^t$ hodnotou 2^t . Zo stavu 4 možno v rámci behu do stavu s nenulovou koncovou váhou prečítať slová typu $(aaa)^t$ a $(aba)^t$ pre $t \in \mathbb{N}$ – ku každému takémuto slovu prislúcha práve jeden beh, ktorého príspevok ku koeficientu daného slova je 1. Automat $\mathcal{A}$ teda realizuje formálny mocninový rad

$$\|\mathcal{A}\| = 2 + \sum_{t \in \mathbb{N} \setminus \{0\}} ((2^t + 1)(aaa)^t + (aba)^t).$$

Prípád $t = 0$ treba skutočne riešiť osobitne: koeficient slova $\varepsilon = (aaa)^0 = (aba)^0$ v rade $\|\mathcal{A}\|$ totiž nie je $(2^0 + 1) + 1 = 3$, ale 2. Ide tu o typickú ukážku toho, že „details“ obyčajne nehrajúce žiadnu rolu v teórii formálnych jazykov – ako napríklad započítanie slova dvakrát – sa v teórii formálnych mocninových radov môžu razom stať podstatnými.

Nasledujúci príklad – okrem toho, že ukazuje na ďalšie možnosti použitia automatov s váhami nad polokruhom prirodzených čísel – možno súčasne považovať aj za demonštráciu podstatne širšieho fenoménu z oblasti kombinatoriky. Ukazuje totiž, ako možno od špecifikácie množiny kombinatorických objektov – v našom prípade slov – jednoducho prejsť k enumerácii objektov danej veľkosti, ktorá je v našom prípade daná dĺžkou slova. Podobné postupy sú v rôznych variáciách relatívne často používané v enumeratívnej kombinatorike. Súvisu s kombinatorikou sa na však tomto mieste hlbšie venovať nebudeme – čitateľa so záujmom o túto problematiku iba odkážeme na [44] alebo na prednášky o *analytickej a enumeratívnej kombinatorike* [70].

Príklad 3.1.2. Uvažujme ľubovoľný *deterministický*¹ konečný automat $\mathcal{B}$ nad abecedou Σ , ktorý môže mať aj čiastočnú prechodovú funkciu – napríklad ten na obrázku 3.2a. Priradíme všetkým jeho prechodom jednotkové váhy z polokruhu $(\mathbb{N}, +, \cdot, 0, 1)$ a nahradíme všetky písmená z abecedy Σ písmenom z ; ak takto medzi niektorou dvojicou stavov vznikne $k \geq 2$ paralelných prechodov na z , nahradíme ich jediným prechodom s váhou k . Vo výsledku tak dostaneme konečný automat $\mathcal{B}'$ nad unárnou vstupnou abecedou $\{z\}$ s váhami nad polokruhom $(\mathbb{N}, +, \cdot, 0, 1)$; ten je znázornený na obrázku 3.2b.



Obr. 3.2: Deterministický konečný automat $\mathcal{B}$ nad vstupnou abecedou $\Sigma = \{a, b\}$ a konečný automat $\mathcal{B}'$ s váhami nad $(\mathbb{N}, +, \cdot, 0, 1)$ realizujúci vytvárajúcu funkciu pre počet slov danej dĺžky v jazyku $\|\mathcal{B}\|$.

Automat $\mathcal{B}'$ nad unárnou vstupnou abecedou $\{z\}$ teda realizuje formálny mocninový rad $\|\mathcal{B}'\|$ s koeficientmi v $\mathbb{N}$ o jednej nekomutujúcej premennej z . Keďže ide o jedinú premennú, nehrá jej nekomutujúcosť žiadnu rolu – sama so sebou totiž premenná komutuje vždy; jej nekomutujúcosť znamená len to, že nekomutuje s prípadnými inými premennými. Rovnako dobre by sme teda mohli premennú z interpretovať aj ako komutujúcu. Na rad $\|\mathcal{B}'\|$ sa teda môžeme dívať ako na prvok $\mathbb{N}\langle\langle z^* \rangle\rangle$ aj ako na prvok $\mathbb{N}[[z]]$; ide koniec koncov o izomorfné polokruhy.

Ľahko možno dokázať, že pri druhom z týchto pohľadov zodpovedá rad $\|\mathcal{B}'\|$ vytvárajúcej funkcii pre počet slov dĺžky t v jazyku $\|\mathcal{B}\|$, čiže

$$\|\mathcal{B}'\| = \sum_{t \in \mathbb{N}} |\|\mathcal{B}\| \cap \Sigma^t| z^t,$$

čo je v prípade automatov z obrázku 3.2 to isté ako

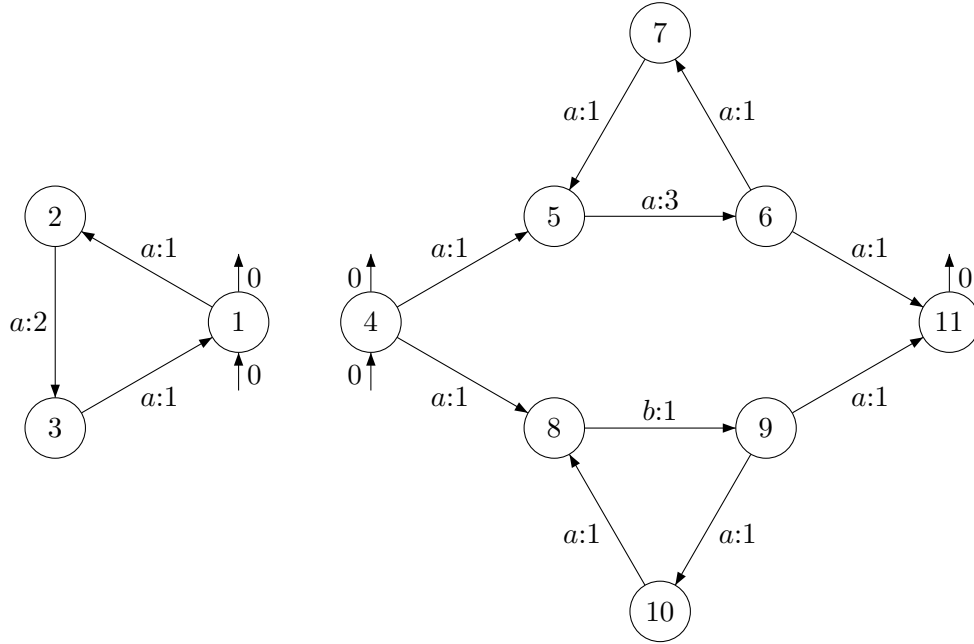
$$\|\mathcal{B}'\| = \sum_{t \in \mathbb{N}} 2^t z^{3t+1}.$$

Keby bol konečný automat $\mathcal{B}$ nedeterministický, táto vlastnosť by už vo všeobecnosti neplatila; rad $\|\mathcal{B}'\|$ by však stále bolo možné vnímať ako vytvárajúcu funkciu pre počet behov dĺžky t , ktoré v automate $\mathcal{B}$ vedú z počiatočného do koncového stavu.

Príklad 3.1.3. Na obrázku 3.3 je znázornený diagram konečného automatu $\mathcal{C}$ nad abecedou $\Sigma = \{a, b\}$ s váhami nad tropickým polokruhom $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$. Keďže nulovým prvkom tohto polokruhu je ∞ , chýbajúce šípky v diagrame treba chápať ako šípky s váhou ∞ .

Ľahko vidieť, že v rámci behu vedúceho zo stavu s konečnou počiatočnou váhou do stavu s konečnou koncovou váhou možno v automate $\mathcal{C}$ prečítať len slová typu $(aaa)^t$ a $(aba)^t$ pre $t \in \mathbb{N}$. Pre slovo $(aaa)^t$ existujú dva takéto behy: po sčítaní počiatočnej váhy prvého stavu, váh všetkých prechodov a koncovej váhy posledného stavu dostávame pri jednom z nich hodnotu $4t$ a pri druhom hodnotu $5t$. Koeficient slova $(aaa)^t$ je daný minimom týchto dvoch hodnôt; je teda rovný $4t$. Pre slovo $(aba)^t$ existuje jediný

¹Rovnako by sme mohli vziať aj ľubovoľný *jednoznačný* konečný automat bez prechodov na prázdne slovo – čiže nedeterministický konečný automat, v ktorom pre každé slovo existuje *najviac jeden* beh z počiatočného do koncového stavu a v ktorom každý prechod zodpovedá prečítaniu práve jedného písmena.



Obr. 3.3: Konečný automat $\mathcal{C}$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$.

takýto beh, pričom rovnakým postupom ako vyššie dospejeme k hodnote $3t$; koeficient slova $(aba)^t$ je teda tiež $3t$. Formálny mocninový rad $\|\mathcal{C}\|$ realizovaný automatom $\mathcal{C}$ je preto daný vzťahmi

$$\begin{aligned} (\|\mathcal{C}\|, (aaa)^t) &= 4t && \text{pre všetky } t \in \mathbb{N}, \\ (\|\mathcal{C}\|, (aba)^t) &= 3t && \text{pre všetky } t \in \mathbb{N}, \\ (\|\mathcal{C}\|, w) &= \infty && \text{pre všetky } w \in \Sigma^* \setminus ((aaa)^* \cup (aba)^*). \end{aligned}$$

Od zápisu radu $\|\mathcal{C}\|$ v podobe formálneho súčtu nad týmto polokruhom upúšťame, aby nedošlo k zámene bežného sčítania so sčítaním v tropickom polokruhu, ktorým je minimum.

3.2 Konečné automaty s váhami nad polokruhom

Model konečných automatov s váhami, ktorý sme neformálne opísali a demonštrovali na príkladoch, teraz zadefinujeme. Urobíme tak hneď dvoma ekvivalentnými spôsobmi. Kým v tomto oddiele automaty s váhami definujeme prístupom obvyklým v klasickej teórii automatov, bezprostredne sledujúcim intuitívnu predstavu z úvodu tejto kapitoly, v nasledujúcom oddiele sformulujeme alternatívnu grafovú – alebo ekvivalentne maticovú – definíciu tohto modelu.

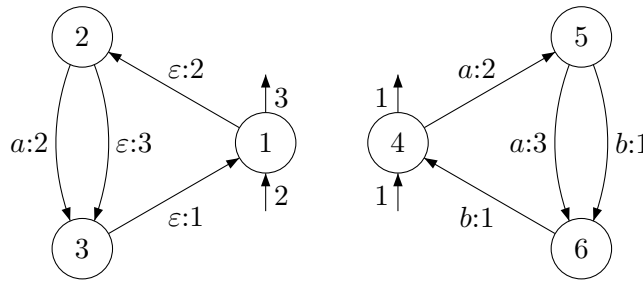
Dôvodom na takúto dvojakú definíciu rovnakého pojmu nie je, ako by sa mohlo zdať, spoľahlivé zmätenie čitateľa hneď na začiatku úvah. Je naopak možné predpokladať, že pre čitateľa dôkladne zžitého s myšlienkami z príkladu 1.2.2 bude prechod medzi oboma pohľadmi na automaty s váhami skoro samozrejímavý. Definíciu „klasického razenia“ v tomto oddiele uvádzame najmä preto, že priamo zodpovedá bežnému spôsobu kreslenia diagramov automatov s váhami, ktorý budeme používať v celom tomto texte; okrem toho tiež môže byť výhodnejšia v prípadoch, keď sa na argumentáciu o automatoch s váhami menej intenzívne využívajú matice. Definícia z oddielu 3.3 nám naopak umožní chápať automaty s váhami nad polokruhom ako špeciálny prípad konečných automatov nad polokruhmi – čiže ako ohodnotené grafy nad polokruhmi s trochou informácie navyše – a tým pádom nám aj otvorí priestor pre použitie maticových metód.

Definícia 3.2.1. Nech Σ je abeceda a S je polokruh. *Konečný automat s váhami nad polokruhom* S a abecedou Σ je štvorica $\mathcal{A} = (Q, \sigma, \iota, \tau)$, kde Q je konečná množina stavov, $\sigma: Q \times (\Sigma \cup \{\varepsilon\}) \times Q \rightarrow S$ je zobrazenie určujúce váhy prechodov, $\iota: Q \rightarrow S$ je zobrazenie určujúce počiatkové váhy stavov a $\tau: Q \rightarrow S$ je zobrazenie určujúce koncové váhy stavov.

Konečné automaty s váhami možno prirodzene reprezentovať diagramom – a to spôsobom, ktorý sme už v princípe vysvetlili v úvode tejto kapitoly.

Prechodom automatu $\mathcal{A} = (Q, \sigma, \iota, \tau)$ nazveme trojicu $(p, z, q) \in Q \times (\Sigma \cup \{\varepsilon\}) \times Q$ takú, že $\sigma(p, z, q) \neq 0$. Automat $\mathcal{A} = (Q, \sigma, \iota, \tau)$ potom možno zadať diagramom orientovaného multigrafu s množinou vrcholov Q , kde pre všetky prechody (p, z, q) automatu $\mathcal{A}$ vedie v tomto multigrafe hrana z vrcholu p do vrcholu q so „značkou“ $z: \sigma(p, z, q)$. Pokiaľ navyše pre stav $p \in Q$ platí $\iota(p) \neq 0$, je táto skutočnosť v diagrame automatu znázornená „polovičnou“ šípkou vedúcou „odnikiaľ“ do p , pri ktorej je „značka“ $\iota(p)$; podobne pri stavoch $q \in Q$ s $\tau(q) \neq 0$ kreslíme „polovičnú“ šíпку z q „nikam“, ktorá má „značku“ $\tau(q)$.

Príklad 3.2.2. Na obrázku 3.4 je znázornený diagram konečného automatu $\mathcal{D} = (Q, \sigma, \iota, \tau)$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$ a nad abecedou $\Sigma = \{a, b\}$. Keďže sa v tomto automate nevyskytuje žiadna váha s hodnotou ∞ , formálne ide aj o automat nad polokruhom $(\mathbb{N}, +, \cdot, 0, 1)$; dôvod na zahrnutie prvku ∞ však pochopíme zakrátko.



Obr. 3.4: Konečný automat $\mathcal{D}$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$.

Z diagramu automatu $\mathcal{D}$ možno vyčítať všetky zložky štvorice $\mathcal{D} = (Q, \sigma, \iota, \tau)$ – tie sú dané množinou stavov $Q = [6]$, váhami prechodov

$$\begin{aligned} \sigma(1, \varepsilon, 2) &= 2, & \sigma(2, a, 3) &= 2, & \sigma(2, \varepsilon, 3) &= 3, \\ \sigma(3, \varepsilon, 1) &= 1, & \sigma(4, a, 5) &= 2, & \sigma(5, a, 6) &= 3, \\ \sigma(5, b, 6) &= 1, & \sigma(6, b, 4) &= 1 \end{aligned}$$

a $\sigma(p, z, q) = 0$ pre všetky zvyšné trojice $(p, z, q) \in Q \times (\Sigma \cup \{\varepsilon\}) \times Q$, počiatkovými váhami stavov $\iota(1) = 2$, $\iota(4) = 1$, $\iota(2) = \iota(3) = \iota(5) = \iota(6) = 0$ a koncovými váhami $\tau(1) = 3$, $\tau(4) = 1$, $\tau(2) = \tau(3) = \tau(5) = \tau(6) = 0$.

Behom takto chápaného konečného automatu $\mathcal{A} = (Q, \sigma, \iota, \tau)$ s váhami nad polokruhom S a abecedou Σ nazveme konečnú striedavú postupnosť

$$\gamma = (q_0, z_1, q_1, z_2, q_2, \dots, q_{t-1}, z_t, q_t)$$

stavov automatu a slov zo $\Sigma \cup \{\varepsilon\}$, kde $t \in \mathbb{N}$, $q_0, \dots, q_t \in Q$ a $z_1, \dots, z_t \in \Sigma \cup \{\varepsilon\}$ sú také, že pre $i = 1, \dots, t$ je $\sigma(q_{i-1}, z_i, q_i) \neq 0$. Hovoríme tiež, že γ je behom zo stavu q_0 do stavu q_t . *Dĺžkou behu* γ nazveme prirodzené číslo $|\gamma| = t$; ďalej označme ako $\lambda(\gamma) = z_1 z_2 \dots z_t$ slovo nad abecedou Σ prečítané počas behu γ a ako $\sigma(\gamma) = \sigma(q_0, z_1, q_1) \sigma(q_1, z_2, q_2) \dots \sigma(q_{t-1}, z_t, q_t)$ hodnotu behu γ bez započítania počiatkovej váhy stavu q_0 a koncovkej váhy stavu q_t . Množinu všetkých behov v $\mathcal{A}$ označme $\mathcal{R}(\mathcal{A})$.

Dostávame sa k zaapeklitému miestu, ktorým je definícia formálneho mocninového radu realizovaného automatom $\mathcal{A} = (Q, \sigma, \iota, \tau)$ s váhami nad polokruhom S a abecedou Σ . Bez problémov ešte vieme definovať *monóm realizovaný behom* γ zo stavu p do stavu q – ten je daný ako

$$\|\gamma\| := (\iota(p)\sigma(\gamma)\tau(q))\lambda(\gamma).$$

Rad realizovaný automatom $\mathcal{A}$ by sme teraz chceli definovať ako vo všeobecnosti nekonečný súčet monómov $\|\gamma\|$ cez všetky behy $\gamma \in \mathcal{R}(\mathcal{A})$. To však nemusí byť možné: stačí napríklad automat $\mathcal{D}$ z príkladu 3.2.2 interpretovať ako automat s váhami nad polokruhom $(\mathbb{N}, +, \cdot, 0, 1)$ „bez nekonečna“; ľahko vidieť, že $\mathcal{R}(\mathcal{A})$ obsahuje pre všetky $t \in \mathbb{N}$ práve jeden beh γ s $\|\gamma\| = 6^t \varepsilon$, čo znamená, že súčet týchto monómov cez všetky behy γ neexistuje.

Tento problém súvisí s našimi ťažkosťami z oddielu 2.4, kde sme spočítateľne nekonečné súčty formálnych mocninových radov neboli schopní nedefinovať za každých okolností.

Videli sme však, že kedykoľvek je S spočítateľne úplný polokruh, je spočítateľne úplným aj polokruh $S\langle\langle\Sigma^*\rangle\rangle$; ak teda automat $\mathcal{D}$ z príkladu 3.2.2 interpretujeme ako automat s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$, bude ním realizovaný formálny mocninový rad dobre definovaný. Nekonečné súčty sme tiež v oddiele 2.4 definovali pre všetky lokálne konečné systémy mocninových radov. Ľahko pritom vidieť, že systém radov $(\|\gamma\| \mid \gamma \in \mathcal{R}(\mathcal{A}))$ je lokálne konečný kedykoľvek je automat $\mathcal{A} = (Q, \sigma, \iota, \tau)$ „bezepsilónový“ – čiže $\sigma(p, \varepsilon, q) = 0$ pre všetky $p, q \in Q$. Rady realizované „bezepsilónovými“ automaty s váhami tak budeme môcť definovať pre ľubovoľný polokruh váh S .

Formálny mocninový rad realizovaný automatom s váhami teda definujeme pre *všetky automaty s váhami nad spočítateľne úplným polokruhom* bez akéhokoľvek obmedzenia na existenciu prechodov na prázdne slovo a pre *všetky „bezepsilónové“ automaty s váhami* bez obmedzenia na polokruh váh.

Definícia 3.2.3. Nech $\mathcal{A}$ je konečný automat s váhami nad polokruhom S a vstupnou abecedou Σ . V prípade, že je polokruh váh S spočítateľne úplný *alebo* je automat $\mathcal{A}$ „bezepsilónový“, definujeme *formálny mocninový rad* $\|\mathcal{A}\|$ *realizovaný automatom* $\mathcal{A}$ ako prvok $S\langle\langle\Sigma^*\rangle\rangle$ daný predpisom

$$\|\mathcal{A}\| = \sum_{\gamma \in \mathcal{R}(\mathcal{A})} \|\gamma\|. \quad (3.1)$$

Príklad 3.2.4. Automat $\mathcal{D}$ z príkladu 3.2.2 realizuje, ako možno ľahko overiť, rad

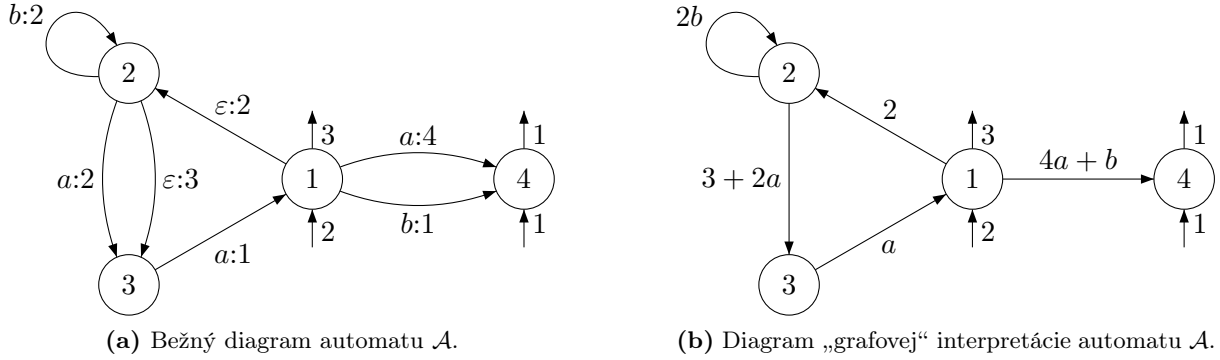
$$\|\mathcal{D}\| = \left(\sum_{t \in \mathbb{N}} \infty a^t \right) + (6aab + 2abb)^*.$$

Príklad 3.2.5. Ak $\mathcal{E} = (Q, \sigma, \iota, \tau)$ je automat s váhami taký, že $Q = \emptyset$, je $\|\mathcal{E}\| = 0$.

3.3 Alternatívny pohľad na automaty s váhami

Sformulujeme teraz alternatívnu definíciu konečných automatov s váhami, pri ktorej ich bude možné chápať ako ohodnotené grafy nad polokruhom formálnych mocninových radov spolu s dodatočnou informáciou o počiatkových a koncových váhach jednotlivých stavov. *S touto definíciou budeme pracovať po zvyšok kapitoly*, pričom jej korešpondencia s definíciou z predchádzajúceho oddielu bude zrejmá.

Od „klasického“ chápania automatov s váhami z predošlého oddielu sa k ich „grafovej“ interpretácii dostaneme rovnakou cestou, akú sme opísali v príklade 1.2.2 pre bežné nedeterministické konečné automaty. Jedinou prekážkou brániacou tomu, aby sme mohli automat s váhami v zmysle definície 3.2.1 chápať ako ohodnotený graf s dodatočnou informáciou o počiatkových a koncových váhach, je možnosť existencie násobných hrán. Tých sa však môžeme zbaviť podobne ako v príklade 1.2.2 – k paralelných prechodov na slová $z_1, \dots, z_k \in \Sigma \cup \{\varepsilon\}$ s váhami $a_1, \dots, a_k \in S$ môžeme nahradiť jediným prechodom ohodnoteným *polynómom* $p \in S\langle\Sigma \cup \{\varepsilon\}\rangle$ daným ako $p = a_1 z_1 + \dots + a_k z_k$. Tento proces je znázornený na obrázku 3.5 a jeho výsledkom je automat s váhami chápaný ako špeciálny automat nad polokruhom formálnych mocninových radov.



Obr. 3.5: Konečný automat $\mathcal{A}$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$.

Definícia 3.3.1. Nech Σ je abeceda a S je polokruh. *Konečný automat s váhami nad polokruhom S* a abecedou Σ je štvorica $\mathcal{A} = (Q, \varphi, \iota, \tau)$, kde Q je konečná množina stavov, $\varphi: Q^2 \rightarrow S \langle \Sigma \cup \{\varepsilon\} \rangle$ je zobrazenie určujúce váhy prechodov, $\iota: Q \rightarrow S$ je zobrazenie určujúce počiatočné váhy stavov a $\tau: Q \rightarrow S$ je zobrazenie určujúce koncové váhy stavov.

Podobne ako v predchádzajúcom oddiele budeme hovoriť, že konečný automat $\mathcal{A} = (Q, \varphi, \iota, \tau)$ s váhami nad polokruhom S a abecedou Σ je „bezepsilonový“, ak $\varphi(p, q) \in S \langle \Sigma \rangle$ pre všetky $p, q \in Q$.

Grafom automatu $\mathcal{A} = (Q, \varphi, \iota, \tau)$ nazveme ohodnotený graf $\mathcal{G}[\mathcal{A}] = (Q, \varphi)$ nad polokruhom $S \langle \Sigma^* \rangle$. *Behom*² automatu $\mathcal{A}$ zo stavu p do stavu q nazveme ľubovoľný sled γ z vrcholu p do vrcholu q v grafe $\mathcal{G}[\mathcal{A}]$, čo nám umožňuje hovoriť aj o dĺžke $|\gamma|$ a ohodnotení $\varphi(\gamma)$ behu γ . *Množinu všetkých behov* automatu $\mathcal{A}$ označíme $\mathcal{R}(\mathcal{A})$.

Keďže budeme pri práci s automatmi s váhami často používať maticové metódy, *budeme v nasledujúcom vždy bez ujmy na všeobecnosti predpokladať, že $Q = [n]$ pre nejaké $n \in \mathbb{N}$* . Pre takýto automat $\mathcal{A} = ([n], \varphi, \iota, \tau)$ budeme písať aj $\mathcal{A} = (n, \varphi, \iota, \tau)$. *Prechodovou maticou* automatu $\mathcal{A} = (n, \varphi, \iota, \tau)$ nazveme maticu susednosti $A[\mathcal{A}] := \mathbf{A}(\mathcal{G}[\mathcal{A}]) \in (S \langle \Sigma \cup \{\varepsilon\} \rangle)^{n \times n}$ grafu $\mathcal{G}[\mathcal{A}]$, *vektorom počiatočných váh* automatu $\mathcal{A}$ riadkový vektor $\mathbf{i}[\mathcal{A}] := (\iota(1), \dots, \iota(n)) \in S^{1 \times n}$ a *vektorom koncových váh* automatu $\mathcal{A}$ stĺpcový vektor $\mathbf{f}[\mathcal{A}] := (\tau(1), \dots, \tau(n))^T \in S^{n \times 1}$.

Poznámka 3.3.2. Prechodová matica $A[\mathcal{A}] = A$ poskytuje s vektormi $\mathbf{i}[\mathcal{A}] = \mathbf{i}$ a $\mathbf{f}[\mathcal{A}] = \mathbf{f}$ úplnú informáciu o automate $\mathcal{A}$. Pre takýto n -stavový konečný automat s váhami $\mathcal{A}$ teda budeme písať aj $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$. Takýto spôsob zadania automatu je o dosť úspornejší, než rozpisovanie definície jednotlivých zobrazení φ , ι a τ .

Všimnime si, že každý konečný automat s váhami nad polokruhom S a abecedou Σ je súčasne aj konečným $S \langle \Sigma \cup \{\varepsilon\} \rangle$ -automatom nad polokruhom $S \langle \Sigma^* \rangle$ v zmysle definície 1.6.1.³

Formálny mocninový rad realizovaný automatom s váhami nad S a abecedou Σ teraz definujeme, rovnako ako v predchádzajúcom oddiele, pre dve neporovnateľné triedy týchto automatov: jednak pre *všetky automaty s váhami nad spočítateľne úplným polokruhom* a tiež pre *všetky „bezepsilonové“ automaty s váhami nad ľubovoľným polokruhom*. Pre automaty z prvej z týchto tried sú už pritom definované, vďaka úvahám z oddielu 1.6, nimi realizované prvky polokruhu $S \langle \Sigma^* \rangle$. Naša definícia realizovaného formálneho mocninového radu bude s týmto konceptom konzistentná.

²Pojem behu je miestom, pri ktorom sa oba pohľady na automaty s váhami najväčšie rozchádzajú – beh automatu s váhami chápaného podľa definície 3.3.1 *nezodpovedá* behu automatu s váhami v zmysle definície 3.2.1.

³Tu sa dostávame k dôvodu, prečo sme konečné automaty nad polokruhmi definovali po formálnej stránke pre všetky polokruhy, hoci nimi realizovaný prvok sme doposiaľ definovali iba pre polokruhy, ktoré sú spočítateľne úplné. V nasledujúcom totiž triedu automatov s dobre definovaným realizovaným prvkom polokruhu rozšírime o „bezepsilonové“ automaty s váhami nad ľubovoľným polokruhom S . Tie sú špeciálnymi konečnými automatmi nad polokruhom $S \langle \Sigma^* \rangle$, čo nemusí byť spočítateľne úplný polokruh. Naopak v prípade, že je polokruh S spočítateľne úplný, je spočítateľne úplný aj polokruh $S \langle \Sigma^* \rangle$ a formálny mocninový rad realizovaný automatom s váhami nad S a abecedou Σ definujeme pre *všetky* takéto automaty konzistentne s definíciou prvkov realizovaných automatmi nad polokruhmi.

Polynóm realizovaný behom $\gamma \in \mathcal{R}(\mathcal{A})$ zo stavu p do stavu q definujeme ako

$$\|\gamma\| := \iota(p)\varphi(\gamma)\tau(q). \quad (3.2)$$

Prvok $\varphi(\gamma)$ tu patrí do $S\langle\Sigma^*\rangle$, kým $\iota(p)$ a $\tau(q)$ sú skaláry z S .

Definícia 3.3.3. Nech $\mathcal{A}$ je automat s váhami nad polokruhom S a abecedou Σ . V prípade, že je polokruh S spočítateľne úplný alebo je automat $\mathcal{A}$ „bezepsilonový“, definujeme *formálny mocninový rad* $\|\mathcal{A}\|$ realizovaný automatom $\mathcal{A}$ ako prvok $S\langle\Sigma^*\rangle$ daný predpisom

$$\|\mathcal{A}\| = \sum_{\gamma \in \mathcal{R}(\mathcal{A})} \|\gamma\|. \quad (3.3)$$

Korektnosť tejto definície v prípade, že S je spočítateľne úplný polokruh, je zrejma – rovnako zrejma je aj jej konzistencia s definíciou prvku realizovaného $S\langle\Sigma \cup \{\varepsilon\}\rangle$ -automatom nad $S\langle\Sigma^*\rangle$ z oddielu 1.6. V prípade, že je automat $\mathcal{A}$ „bezepsilonový“, je korektnosť uvedenej definície dôsledkom vety 2.5.3.

Z pozorovania učeného v poznámke 2.4.10 tiež vyplýva, že pri „bezepsilonových“ automatoch s váhami nad spočítateľne úplným polokruhom v definícii 3.3.3 nezáleží na tom, či ich interpretujeme ako „bezepsilonové“ automaty alebo ako automaty s váhami nad spočítateľne úplným polokruhom – rad $\|\mathcal{A}\|$ bude v oboch prípadoch rovnaký. Všetky naše doterajšie definície automatových modelov sú teda navzájom konzistentné.

Konzistencia teórie práve definovaných automatov s teóriou $S\langle\Sigma \cup \{\varepsilon\}\rangle$ -automatov nad polokruhom mocninových radov $S\langle\Sigma^*\rangle$ siaha v prípade spočítateľne úplného polokruhu S ešte o niečo ďalej, než k definícii objektov realizovaných týmito automatmi. Hoci automaty s váhami nad S a abecedou Σ tvoria iba vlastnú podtriedu týchto automatov nad polokruhmi – počiatočnými a koncovými váhami sú tu namiesto ľubovoľných prvkov $S\langle\Sigma \cup \{\varepsilon\}\rangle$ skaláry z S – z vety 1.7.2 vyplýva, že automaty s váhami nad S sú vlastne *normálnym tvarom* $S\langle\Sigma \cup \{\varepsilon\}\rangle$ -automatov nad $S\langle\Sigma^*\rangle$. To znamená, že obe triedy automatov realizujú rovnakú triedu radov.

Nasledujúca veta poskytuje maticovú charakterizáciu formálneho mocninového radu $\|\mathcal{A}\|$ realizovaného automatom s váhami $\mathcal{A}$.

Veta 3.3.4. Nech $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je automat s váhami nad polokruhom S a vstupnou abecedou Σ , pričom je polokruh S spočítateľne úplný alebo je automat $\mathcal{A}$ „bezepsilonový“. Potom $\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f}$.

Dôkaz. Pre automaty nad spočítateľne úplným polokruhom je toto tvrdenie iba špeciálnym prípadom vety 1.6.3. Pre „bezepsilonové“ automaty nad ľubovoľným polokruhom je relatívne priamočiarym dôsledkom vzťahov (3.2), (3.3), vety 2.5.3 a časti (iii) tvrdenia 2.4.15. $\square$

Čitateľovi je na tomto mieste iste zrejma vzájomná korešpondencia oboch alternatívnych definícií automatov s váhami. Pre úplnosť ju teraz ešte zachytíme formálne. Uvažujme najprv konečný automat $\mathcal{A} = (Q, \sigma, \iota, \tau)$ s váhami nad polokruhom S a vstupnou abecedou Σ , chápaný v zmysle „klasickej“ definície 3.2.1.⁴ Pri „grafovom“ pohľade z definície 3.3.1 k automatu $\mathcal{A}$ prislúcha automat $\mathcal{B} = (Q, \varphi, \iota, \tau)$, kde polynóm $\varphi(p, q) \in S\langle\Sigma \cup \{\varepsilon\}\rangle$ je pre všetky $p, q \in Q$ daný ako

$$\varphi(p, q) = \sum_{z \in \Sigma \cup \{\varepsilon\}} \sigma(p, z, q)z.$$

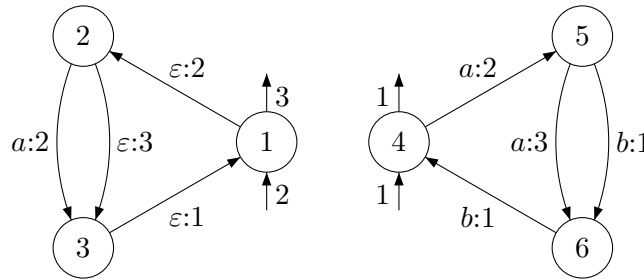
Je evidentné, že automat $\mathcal{B}$ je „bezepsilonový“ práve vtedy, keď je „bezepsilonový“ automat $\mathcal{A}$. Zostáva teda dokázať, že v prípade spočítateľnej úplnosti polokruhu S alebo „bezepsilonovosti“ oboch uvažovaných automatov je $\|\mathcal{A}\| = \|\mathcal{B}\|$. Túto úlohu prenechávame čitateľovi, ktorý pri argumentácii môže využiť vhodný rozklad množiny behov $\mathcal{R}(\mathcal{A})$ podľa postupností „navštívených“ stavov.

⁴V prípade potreby môžeme bez ujmy na všeobecnosti predpokladať, že $Q = [n]$ pre nejaké $n \in \mathbb{N}$.

Ku každému konečnému automatu $\mathcal{B} = (Q, \varphi, \iota, \tau)$ s váhami nad polokruhom S a abecedou Σ , chápanému v zmysle definície 3.3.1, naopak pri ponímaní vyplývajúcom z definície 3.2.1 zodpovedá automat $\mathcal{A} = (Q, \sigma, \iota, \tau)$ taký, že pre všetky $(p, z, q) \in Q \times (\Sigma \cup \{\varepsilon\}) \times Q$ je $\sigma(p, z, q) = (\varphi(p, q), z)$. Táto konštrukcia je zrejme inverzná ku konštrukcii opísanej v predchádzajúcom odstavci.

Po zvyšok tohto textu budeme na formálnej úrovni pracovať s definíciou 3.3.1. Diagramy automatov s váhami ale budeme naďalej kresliť spôsobom založeným na definícii 3.2.1 a vyššie opísanej korešpondencii medzi oboma definíciami.

Príklad 3.3.5. Na obrázku 3.6 je opätovne znázornený diagram konečného automatu $\mathcal{D}$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$ a vstupnou abecedou $\Sigma = \{a, b\}$, s ktorým sme sa už stretli v rámci príkladu 3.2.2.



Obr. 3.6: Konečný automat $\mathcal{D}$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$.

V zmysle definície 3.3.1 je automat $\mathcal{D}$ daný ako $\mathcal{D} = (Q, \varphi, \iota, \tau)$, kde $Q = [6]$,

$$\begin{aligned} \varphi(1, 2) &= 2, & \varphi(2, 3) &= 3 + 2a, & \varphi(3, 1) &= 1, \\ \varphi(4, 5) &= 2a, & \varphi(5, 6) &= 3a + b, & \varphi(6, 4) &= b, \end{aligned}$$

$\varphi(p, q) = 0$ pre všetky ostatné dvojice $(p, q) \in Q^2$, $\iota(1) = 2$, $\iota(4) = 1$, $\iota(2) = \iota(3) = \iota(5) = \iota(6) = 0$, $\tau(1) = 3$, $\tau(4) = 1$ a $\tau(2) = \tau(3) = \tau(5) = \tau(6) = 0$.

Automat $\mathcal{D}$ tiež možno zadať pomocou jeho vektora počiatkových váh $\mathbf{i} = \mathbf{i}[\mathcal{D}] = (2, 0, 0, 1, 0, 0)$, prechodovej matice

$$A = A[\mathcal{D}] = \begin{pmatrix} 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 + 2a & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2a & 0 \\ 0 & 0 & 0 & 0 & 0 & 3a + b \\ 0 & 0 & 0 & b & 0 & 0 \end{pmatrix}$$

a vektoru koncových váh $\mathbf{f} = \mathbf{f}[\mathcal{D}] = (3, 0, 0, 1, 0, 0)^T$.

3.4 Normálne tvary automatov s váhami

V krátkosti sa teraz pristavíme pri dvoch vetách o normálnych tvaroch konečných automatov s váhami. Prvá z nich hovorí o normálnom tvare automatov s váhami *nad spočítateľne úplným polokruhom*, ktorý je špeciálnym prípadom normálneho tvaru pre automaty nad polokruhmi opísaného v oddiele 1.7. Automat s váhami nad spočítateľne úplným polokruhom je v tomto normálnom tvare, ak obsahuje práve jeden stav s nenulovou počiatkovou váhou, ktorá je pre tento stav rovná jednej a práve jeden stav s nenulovou koncovou váhou, ktorá je tiež rovná jednej; tieto dva stavy musia byť rôzne. V automate navyše nesmú viesť žiadne prechody *do* stavu s nenulovou počiatkovou váhou, ani *zo* stavu s nenulovou koncovou váhou.

Veta 3.4.1. *Nech $\mathcal{A} = (Q, \varphi, \iota, \tau)$ je konečný automat s váhami nad spočítateľne úplným polokruhom S a abecedou Σ . Potom existuje konečný automat $\mathcal{A}' = (Q', \varphi', \iota', \tau')$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}'\| = \|\mathcal{A}\|$ a existujú stavy $q_i, q_f \in Q'$, pre ktoré sú splnené nasledujúce podmienky: $q_i \neq q_f$, $\iota'(q_i) = \tau'(q_f) = 1$, $\iota'(q) = 0$ pre všetky $q \in Q' \setminus \{q_i\}$, $\tau'(q) = 0$ pre všetky $q \in Q' \setminus \{q_f\}$ a pre všetky $p, q \in Q'$ s $\varphi'(p, q) \neq 0$ je $p \neq q_f$ a $q \neq q_i$.*

Dôkaz. Ide o špeciálny prípad vety 1.7.2. □

Pre „bezepsilonové“ automaty s váhami nad všeobecným polokruhom sa presná obdoba vety 3.4.1 zo zrejmých príčin sformulovať *nedá*: pre „bezepsilonový“ automat $\mathcal{A}$ v tvare z vety 3.4.1 totiž musí byť $(\|\mathcal{A}\|, \varepsilon) = 0$. Možnými zoslabeniami vety 3.4.1, platnými aj pre „bezepsilonové“ automaty s váhami, sa v tomto texte zaoberať nebudeme.

V automatoch s váhami nad spočítateľne úplným polokruhom sú, na rozdiel od „bezepsilonových“ automatov s váhami nad všeobecným polokruhom, povolené aj prechody na prázdne slovo ε . Dokážeme teraz, že táto „vymoženosť“ nie je nijak kľúčová: ku každému automatu s váhami nad spočítateľne úplným polokruhom S existuje k nemu ekvivalentný „bezepsilonový“ automat s váhami nad S . Prechody na prázdne slovo teda síce často uľahčia konštrukcie automatov s váhami nad spočítateľne úplnými polokruhmi, no nijako neovplyvňujú opisnú silu modelu. Na preskúmanie všetkých formálnych mocninových radov realizovateľných automatmi s váhami sa tak stačí obmedziť na „bezepsilonové“ automaty s váhami nad všeobecným polokruhom.

Veta 3.4.2. *Nech $n \in \mathbb{N}$ a $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad spočítateľne úplným polokruhom S a abecedou Σ . Potom existuje „bezepsilonový“ konečný automat $\mathcal{A}'$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}'\| = \|\mathcal{A}\|$.*

Dôkaz. Maticu⁵ $A \in S^{n \times n} \langle \Sigma \cup \{\varepsilon\} \rangle$ možno vyjadriť ako súčet jej „epsilonovej“ a jej „neepsilonovej“ zložky:

$$A = A_\varepsilon + A_\Sigma, \quad (3.4)$$

kde $(A_\varepsilon, \varepsilon) = (A, \varepsilon)$, $(A_\Sigma, \varepsilon) = \mathbf{0}$ a kde pre všetky $c \in \Sigma$ je $(A_\varepsilon, c) = \mathbf{0}$ a $(A_\Sigma, c) = (A, c)$.

Nech teraz $\mathcal{A}'$ je automat s váhami taký, že $\mathbf{i}[\mathcal{A}'] = \mathbf{i}$,

$$A[\mathcal{A}'] = A_\varepsilon^* A_\Sigma$$

a $\mathbf{f}[\mathcal{A}'] = A_\varepsilon^* \mathbf{f}$. Automat $\mathcal{A}'$ je „bezepsilonový“ vďaka tvrdeniu 2.6.3,⁶ keďže jeho prechodová matica je daná ako $A_\varepsilon^* A_\Sigma$, kde matica A_Σ je „bezepsilonová“. Zostáva teda dokázať, že $\|\mathcal{A}'\| = \|\mathcal{A}\|$. Z vety 3.3.4, vety 2.6.1 pre spočítateľne úplný polokruh $S^{n \times n} \langle \Sigma^* \rangle$ a rovnosti (3.4) ale dostávame

$$\|\mathcal{A}'\| = \mathbf{i} (A_\varepsilon^* A_\Sigma)^* (A_\varepsilon^* \mathbf{f}) = \mathbf{i} ((A_\varepsilon^* A_\Sigma)^* A_\varepsilon^*) \mathbf{f} = \mathbf{i} (A_\varepsilon + A_\Sigma)^* \mathbf{f} = \mathbf{i} A^* \mathbf{f} = \|\mathcal{A}\|$$

a dôkaz vety je tým pádom dokončený. □

Poznámka 3.4.3. Automaty s váhami nad polokruhom $\mathbb{B}$ môžeme chápať, ako sme už spomínali, aj ako klasické nedeterministické konečné automaty rozoznávajúce formálne jazyky. Pre takéto automaty je konštrukcia z predchádzajúceho dôkazu len maticovým vyjadrením jednej z klasických konštrukcií na odstránenie prechodov na prázdne slovo v nedeterministických konečných automatoch [56]. Duálnu konštrukciu možno navyše realizovať nasledovne: $\mathbf{i}[\mathcal{A}'] = \mathbf{i} A_\varepsilon^*$, $A[\mathcal{A}'] = A_\Sigma A_\varepsilon^*$ a $\mathbf{f}[\mathcal{A}'] = \mathbf{f}$.

⁵Opäť tu stotožňujeme maticu formálnych mocninových radov a k nej prislúchajúci formálny mocninový rad s maticovými koeficientmi.

⁶Pri zneužití nášho stotožňovania množín $(S \langle \Sigma^* \rangle)^{n \times n}$ a $S^{n \times n} \langle \Sigma^* \rangle$.

3.5 Racionálne mocninové rady

Definujeme teraz *racionálne výrazy s váhami*.⁷ Dokážeme ich ekvivalenciu s konečnými automatmi s váhami a formálne mocninové rady realizované týmito ekvivalentnými modelmi nazveme *racionálnymi*. Tentokrát už aj prívlastok „racionálny“ zdôvodníme: ukážeme, že množina racionálnych mocninových radov v $\mathbb{C}\langle\langle z^* \rangle\rangle$ – takéto rady možno interpretovať aj ako rady v $\mathbb{C}[[z]]$ – pozostáva z práve všetkých formálnych Maclaurinových radov racionálnych funkcií analytických v bode $z = 0$.⁸

Podobne ako pri automatoch s váhami budeme aj teraz pracovať s dvoma variantmi racionálnych výrazov: s racionálnymi výrazmi s váhami nad *spočítateľne úplným polokruhom* a s určitým spôsobom syntakticky obmedzenými racionálnymi výrazmi s váhami nad všeobecným polokruhom – v druhom prípade budeme hovoriť o *platných* racionálnych výrazoch. Prvý z týchto variantov racionálnych výrazov s váhami pritom bude ekvivalentný $S\langle\Sigma \cup \{\varepsilon\}\rangle$ -racionálnym výrazom nad spočítateľne úplným polokruhom $S\langle\Sigma^*\rangle$. Dodatočne tak odôvodníme aj používanie termínu „racionálny prvok spočítateľne úplného polokruhu“ zavedeného v oddiele 1.8.

Definícia 3.5.1 (Syntax racionálnych výrazov). Nech Σ je abeceda a S je polokruh. Množina *všetkých racionálnych výrazov s váhami* nad polokruhom S a abecedou Σ je daná jazykom $S\text{-RatE}(\Sigma^*)$ nad (vo všeobecnosti aj nekonečnou) abecedou $\{r \mid r \in S\langle\Sigma \cup \{\varepsilon\}\rangle\} \cup \{+, \cdot, *, (,)\}$ takým, že:

- (i) Jazyk $S\text{-RatE}(\Sigma^*)$ obsahuje pre každé $r \in S\langle\Sigma \cup \{\varepsilon\}\rangle$ slovo r .
- (ii) Pre všetky $E, F \in S\text{-RatE}(\Sigma^*)$ obsahuje jazyk $S\text{-RatE}(\Sigma^*)$ slová $(E + F)$, $(E \cdot F)$ a (E^*) .
- (iii) Jazyk $S\text{-RatE}(\Sigma^*)$ neobsahuje žiadne ďalšie slová.

Slová z jazyka $S\text{-RatE}(\Sigma^*)$ nazývame *racionálnymi výrazmi s váhami* nad S a abecedou Σ .

Podobne ako pre konečné automaty s váhami sa dá bez problémov definovať formálny mocninový rad realizovaný ľubovoľným racionálnym výrazom s váhami nad *spočítateľne úplným* polokruhom. Pre racionálne výrazy s váhami nad *všeobecným* polokruhom však nastanú komplikácie pri aplikácii iterácie na podvýraz realizujúci rad s nenulovým konštantným koeficientom. Preto v prípade váh zo všeobecného polokruhu zavedieme sémantiku len pre výrazy, v ktorých sa iterácia aplikuje výhradne na podvýrazy realizujúce „bezepsilonové“ mocninové rady. Keďže je však aj táto podmienka sama o sebe vecou sémantiky, nahradíme ju *syntaktickou* definíciou množiny racionálnych výrazov, ktoré ju zaručene spĺňajú (hoci nemusí ísť nad každým polokruhom o *všetky* takéto výrazy) – výrazy z tejto množiny nazveme *platnými*.

Definícia 3.5.2 (Platné racionálne výrazy). Nech Σ je abeceda a S je polokruh. Množina „*zaručene bezepsilonových*“ racionálnych výrazov s váhami nad S a abecedou Σ je daná nasledovne:

- (i) Pre každé $r \in S\langle\Sigma\rangle$ je „zaručene bezepsilonový“ výraz r .
- (ii) Pre každé dva „zaručene bezepsilonové“ výrazy E, F a pre každý racionálny výraz G sú „zaručene bezepsilonové“ výrazy $(E + F)$, $(G \cdot E)$ a $(E \cdot G)$.
- (iii) Žiaden iný racionálny výraz z $S\text{-RatE}(\Sigma^*)$ nie je „zaručene bezepsilonový“.

Množina *platných* racionálnych výrazov s váhami nad S a abecedou Σ je potom daná takto:

- (i) Pre každé $r \in S\langle\Sigma \cup \{\varepsilon\}\rangle$ je platný výraz r .
- (ii) Pre každé dva platné výrazy E, F a pre každý platný a súčasne „zaručene bezepsilonový“ výraz Z sú platné výrazy $(E + F)$, $(E \cdot F)$ a (Z^*) .
- (iii) Žiaden iný racionálny výraz z $S\text{-RatE}(\Sigma^*)$ nie je platný.

⁷Ako sme už viackrát v texte spomenuli, ide o výstižnejšie pomenovanie pre výrazy nazývané aj „regulárnymi“.

⁸Komplexné koeficienty tu nie sú nijak dôležité. Analogická vlastnosť platí aj pre racionálne rady v $\mathbb{R}\langle\langle z^* \rangle\rangle$.

Môžeme teraz pristúpiť k definícii sémantiky racionálnych výrazov s váhami. Pre racionálne výrazy s váhami nad *spočítateľne úplným* polokruhom bude korektnosť tejto definície zrejmá. Pre *platné* výrazy nad *všeobecným* polokruhom bude korektnosť definície jednoduchým dôsledkom tvrdenia 2.6.3.

Definícia 3.5.3 (Sémantika racionálnych výrazov). Nech Σ je abeceda, S je polokruh a G je racionálny výraz s váhami nad S a abecedou Σ . Predpokladajme navyše, že je polokruh S *spočítateľne úplný* alebo je racionálny výraz G platný. *Formálny mocninový rad* $\|G\|$ *realizovaný racionálnym výrazom* G je potom daný nasledujúcou induktívnou definíciou:

- (i) Ak $G = r$ pre nejaké $r \in S\langle\Sigma \cup \{\varepsilon\}\rangle$, tak $\|G\| = r$.
- (ii) Nech E, F sú racionálne výrazy z $S\text{-RatE}(\Sigma^*)$. Ak $G = (E+F)$, tak $\|G\| = \|E\| + \|F\|$; ak $G = (E \cdot F)$, tak $\|G\| = \|E\| \cdot \|F\|$; ak $G = (E^*)$, tak $\|G\| = \|E\|^*$.

Poznámka 3.5.4. Z tvrdenia 2.6.3 vyplýva, že pre platné „zaručene bezepsilonové“ racionálne výrazy E je $(\|E\|, \varepsilon) = 0$ – čo je fakt úzko previazaný s korektnosťou predošlej definície.

Poznámka 3.5.5. Takto definovaná sémantika nám umožňuje pri zápise racionálnych výrazov s váhami vynechávať zátvorky rovnako, ako sme to opísali v poznámke 1.8.3 vzťahujúcej sa na racionálne výrazy nad *spočítateľne úplnými* polokruhmi.

Dokážeme teraz ekvivalenciu racionálnych výrazov s váhami a konečných automatov s váhami. Presnejšie dokážeme nasledujúce dve tvrdenia:

1. Ku každému „bezepsilonovému“ konečnému automatu s váhami nad *všeobecným* polokruhom S a abecedou Σ existuje ekvivalentný *platný* racionálny výraz s váhami nad S a abecedou Σ .
2. Ku každému *platnému* racionálnemu výrazu s váhami nad *všeobecným* polokruhom S a abecedou Σ existuje ekvivalentný „bezepsilonový“ konečný automat s váhami nad S a abecedou Σ .

Pri racionálnych výrazoch a konečných automatoch s váhami nad *spočítateľne úplnými* polokruhmi je ich ekvivalencia dôsledkom vety 1.8.9. Keďže navyše ku každému racionálnemu výrazu s váhami nad *spočítateľne úplným* polokruhom S existuje ekvivalentný konečný automat s váhami nad S a ten ďalej možno podľa vety 3.4.2 „odepsilonať“, zisťujeme, že *ku každému racionálnemu výrazu s váhami nad spočítateľne úplným polokruhom* existuje ekvivalentný „bezepsilonový“ konečný automat s váhami, a teda aj ekvivalentný *platný racionálny výraz s váhami*. Nad *spočítateľne úplnými* polokruhmi sú teda všeobecné racionálne výrazy s váhami rovnako silné ako platné racionálne výrazy (a nad ostatnými polokruhmi majú definovanú sémantiku iba platné výrazy). Na úrovni realizovaných radov preto nie je nutné medzi oboma variantmi výrazov rozlišovať.

Dokážme teda najprv, že platné racionálne výrazy s váhami nad ľubovoľným polokruhom S sú *aspoň také silné* ako „bezepsilonové“ konečné automaty s váhami nad S .

Lema 3.5.6. Nech $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je „bezepsilonový“ konečný automat s váhami nad polokruhom S a abecedou Σ . Potom existuje *platný výraz* $E \in S\text{-RatE}(\Sigma^*)$ taký, že $\|E\| = \|\mathcal{A}\|$.

Dôkaz. Nech $A^* = (r_{i,j})_{n \times n}$. Indukciou vzhľadom na n možno napríklad⁹ s využitím lemy 2.5.4 dokázať, že pre $i, j = 1, \dots, n$ existuje platný racionálny výraz $F_{i,j}$ s váhami nad S a abecedou Σ taký, že $\|F_{i,j}\| = r_{i,j}$. Na završenie dôkazu lemy už len potom stačí využiť skutočnosť, že $\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f}$. Detaily tohto induktívneho dôkazu prenechávame čitateľovi ako jedno z cvičení v závere kapitoly. $\square$

⁹Alternatívne sa dá dokázať, že aj iteráciu „bezepsilonovej“ matice nad ľubovoľným polokruhom formálnych mocninových radov možno vypočítať obdobou algoritmu 1.5.1. Ďalej potom možno postupovať skoro rovnako ako v dôkaze lemy 1.8.4 (treba akurát argumentovať, že výsledný racionálny výraz bude platný).

V nasledujúcom budeme dokazovať, že ku každému platnému racionálnemu výrazu s váhami existuje ekvivalentný „bezepsilonový“ konečný automat s váhami. Konštrukciu rozdelíme do nasledujúcich štyroch lemm. Najprv ukážeme, že existujú konečné automaty ekvivalentné atomickému výrazu r pre $r \in S\langle \Sigma \cup \{\varepsilon\} \rangle$. V ďalších troch lemmách potom postupne zostrojíme konečné automaty realizujúce súčet, súčin a iteráciu.

Lema 3.5.7. *Nech Σ je abeceda, S je polokruh a $r \in S\langle \Sigma \cup \{\varepsilon\} \rangle$. Potom existuje „bezepsilonový“ konečný automat $\mathcal{A}$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}\| = r$.*

Dôkaz. Nech $r_\Sigma = r - (r, \varepsilon)$ a $\mathcal{A} = (2, \mathbf{i}, A, \mathbf{f})$, kde $\mathbf{i} = (1, 0)$,

$$A = \begin{pmatrix} 0 & r_\Sigma \\ 0 & 0 \end{pmatrix}$$

a $\mathbf{f} = ((r, \varepsilon), 1)^T$. Pre realizovaný rad $\|\mathcal{A}\|$ potom vďaka dôsledku 2.5.5 dostávame

$$\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f} = (1, 0) \begin{pmatrix} 0 & r_\Sigma \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} (r, \varepsilon) \\ 1 \end{pmatrix} = (1, 0) \begin{pmatrix} 1 & r_\Sigma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (r, \varepsilon) \\ 1 \end{pmatrix} = (r, \varepsilon) + r_\Sigma = r,$$

čo dokazuje správnosť konštrukcie. □

Lema 3.5.8. *Nech Σ je abeceda, S je polokruh a $\mathcal{A}_1, \mathcal{A}_2$ sú „bezepsilonové“ konečné automaty s váhami nad S a abecedou Σ . Potom existuje „bezepsilonový“ konečný automat $\mathcal{A}$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}\| = \|\mathcal{A}_1\| + \|\mathcal{A}_2\|$.*

Dôkaz. Nech $\mathcal{A}_1 = (n_1, \mathbf{i}_1, A_1, \mathbf{f}_1)$ a $\mathcal{A}_2 = (n_2, \mathbf{i}_2, A_2, \mathbf{f}_2)$. Položme $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$, kde $n = n_1 + n_2$, $\mathbf{i} = (\mathbf{i}_1, \mathbf{i}_2)$,

$$A = \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & A_2 \end{pmatrix}$$

a

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix}.$$

Z dôsledku 2.5.5 potom dostávame

$$\begin{aligned} \|\mathcal{A}\| &= \mathbf{i}A^*\mathbf{f} = (\mathbf{i}_1 \quad \mathbf{i}_2) \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & A_2 \end{pmatrix}^* \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix} = (\mathbf{i}_1 \quad \mathbf{i}_2) \begin{pmatrix} A_1^* & \mathbf{0} \\ \mathbf{0} & A_2^* \end{pmatrix} \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix} = \\ &= \mathbf{i}_1 A_1^* \mathbf{f}_1 + \mathbf{i}_2 A_2^* \mathbf{f}_2 = \|\mathcal{A}_1\| + \|\mathcal{A}_2\|, \end{aligned}$$

čo dokazuje správnosť konštrukcie. □

Lema 3.5.9. *Nech Σ je abeceda, S je polokruh a $\mathcal{A}_1, \mathcal{A}_2$ sú „bezepsilonové“ konečné automaty s váhami nad S a abecedou Σ . Potom existuje „bezepsilonový“ konečný automat $\mathcal{A}$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}\| = \|\mathcal{A}_1\| \cdot \|\mathcal{A}_2\|$.*

Dôkaz. Nech $\mathcal{A}_1 = (n_1, \mathbf{i}_1, A_1, \mathbf{f}_1)$ a $\mathcal{A}_2 = (n_2, \mathbf{i}_2, A_2, \mathbf{f}_2)$. Položme $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$, kde $n = n_1 + n_2$, $\mathbf{i} = (\mathbf{i}_1, \mathbf{0})$,

$$A = \begin{pmatrix} A_1 & \mathbf{f}_1 \mathbf{i}_2 A_2 \\ \mathbf{0} & A_2 \end{pmatrix}$$

a

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \mathbf{i}_2 \mathbf{f}_2 \\ \mathbf{f}_2 \end{pmatrix}.$$

Z dôsledku 2.5.5 a lemy 2.6.4 potom dostávame

$$\begin{aligned}
 \|\mathcal{A}\| &= \mathbf{i}A^*\mathbf{f} = \begin{pmatrix} \mathbf{i}_1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} A_1 & \mathbf{f}_1\mathbf{i}_2A_2 \\ \mathbf{0} & A_2 \end{pmatrix}^* \begin{pmatrix} \mathbf{f}_1\mathbf{i}_2\mathbf{f}_2 \\ \mathbf{f}_2 \end{pmatrix} = \\
 &= \begin{pmatrix} \mathbf{i}_1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} A_1^* & A_1^*\mathbf{f}_1\mathbf{i}_2A_2A_2^* \\ \mathbf{0} & A_2^* \end{pmatrix} \begin{pmatrix} \mathbf{f}_1\mathbf{i}_2\mathbf{f}_2 \\ \mathbf{f}_2 \end{pmatrix} = \\
 &= \mathbf{i}_1A_1^*\mathbf{f}_1\mathbf{i}_2\mathbf{f}_2 + \mathbf{i}_1A_1^*\mathbf{f}_1\mathbf{i}_2A_2A_2^*\mathbf{f}_2 = \mathbf{i}_1A_1^*\mathbf{f}_1\mathbf{i}_2(1 + A_2A_2^*)\mathbf{f}_2 = \\
 &= \mathbf{i}_1A_1^*\mathbf{f}_1\mathbf{i}_2A_2^*\mathbf{f}_2 = \|\mathcal{A}_1\| \cdot \|\mathcal{A}_2\|,
 \end{aligned}$$

čím je dokázaná správnosť konštrukcie automatu $\mathcal{A}$. $\square$

Lema 3.5.10. *Nech Σ je abeceda, S je polokruh a $\mathcal{B}$ je „bezepsilonový“ konečný automat s váhami nad S a abecedou Σ taký, že $(\|\mathcal{B}\|, \varepsilon) = 0$. Potom existuje „bezepsilonový“ konečný automat $\mathcal{A}$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}\| = \|\mathcal{B}\|^*$.*

Dôkaz. Nech $\mathcal{B} = (m, \mathbf{j}, B, \mathbf{g})$. Potom $\mathbf{jg} = (\|\mathcal{B}\|, \varepsilon) = 0$, v dôsledku čoho je $\|\mathcal{B}\| = \mathbf{j}B^+\mathbf{g}$, kde $B^+ = BB^*$. Položme $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$, kde $n = m + 1$, $\mathbf{i} = (1, \mathbf{0})$,

$$A = \begin{pmatrix} 0 & \mathbf{j}B \\ \mathbf{0} & B + \mathbf{g}\mathbf{j}B \end{pmatrix}$$

a

$$\mathbf{f} = \begin{pmatrix} 1 \\ \mathbf{g} \end{pmatrix}.$$

Takto skonštruovaný automat $\mathcal{A}$ je „bezepsilonový“, pretože všetky nenulové členy matice A možno vyjadriť aj ako súčin matíc obsahujúci faktor B ; keďže $(B, \varepsilon) = 0$, stačí sa odvolať na tvrdenie 2.6.3. Z dôsledku 2.5.5, vety 2.6.7, lemy 2.6.5, jej ľahko dokázateľného „maticovo-vektorového“ variantu a lemy 2.6.4 následne vyplýva

$$\begin{aligned}
 \|\mathcal{A}\| &= \mathbf{i}A^*\mathbf{f} = \begin{pmatrix} 1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} 0 & \mathbf{j}B \\ \mathbf{0} & B + \mathbf{g}\mathbf{j}B \end{pmatrix}^* \begin{pmatrix} 1 \\ \mathbf{g} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{j}B(B + \mathbf{g}\mathbf{j}B)^* \\ \mathbf{0} & (B + \mathbf{g}\mathbf{j}B)^* \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{g} \end{pmatrix} = \\
 &= 1 + \mathbf{j}B(B + \mathbf{g}\mathbf{j}B)^*\mathbf{g} = 1 + \mathbf{j}B(B^*\mathbf{g}\mathbf{j}B)^*B^*\mathbf{g} = 1 + \mathbf{j}BB^*(\mathbf{g}\mathbf{j}BB^*)^*\mathbf{g} = 1 + \mathbf{j}BB^*\mathbf{g}(\mathbf{j}BB^*\mathbf{g})^* = \\
 &= 1 + \mathbf{j}B^+\mathbf{g}(\mathbf{j}B^+\mathbf{g})^* = 1 + \|\mathcal{B}\| \cdot \|\mathcal{B}\|^* = \|\mathcal{B}\|^*,
 \end{aligned}$$

čo dokazuje správnosť konštrukcie. $\square$

Môžeme už teraz zozbierať skutočnosti dokázané vyššie a vysloviť samotnú vetu o ekvivalencii konečných automatov a racionálnych výrazov s váhami.

Nad všeobecným polokruhom bude táto veta hovoriť o ekvivalencii „bezepsilonových“ automatov a platných racionálnych výrazov. Nad spočítateľne úplným polokruhom bude hovoriť o ekvivalencii hneď štyroch modelov: automatov s váhami, „bezepsilonových“ automatov s váhami, racionálnych výrazov s váhami a platných racionálnych výrazov s váhami.

Veta 3.5.11. *Nech Σ je abeceda, S je polokruh a $r \in S\langle\langle\Sigma^*\rangle\rangle$ je formálny mocninový rad. Potom sú nasledujúce tvrdenia ekvivalentné:*

- (i) *Existuje „bezepsilonový“ konečný automat $\mathcal{A}$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}\| = r$.*
- (ii) *Existuje platný racionálny výraz G s váhami nad S a abecedou Σ taký, že $\|G\| = r$.*

Ak je navyše polokruh S spočítateľne úplný, sú tvrdenia (i) a (ii) ekvivalentné nasledujúcim dvom tvrdeniam:

- (iii) *Existuje konečný automat $\mathcal{A}$ s váhami nad S a abecedou Σ taký, že $\|\mathcal{A}\| = r$.*
- (iv) *Existuje racionálny výraz G s váhami nad S a abecedou Σ taký, že $\|G\| = r$.*

Dôkaz. Pri implikácii „ak (i), tak (ii)“ sa stačí odvolať na lemu 3.5.6. Opačnú implikáciu možno ľahko dokázať štruktúrnou indukciou: ak $G = r$ pre nejaké $r \in S\langle \Sigma \cup \{\varepsilon\} \rangle$, existuje „bezepsilonový“ automat $\mathcal{A}$ spĺňajúci $\|\mathcal{A}\| = \|r\| = r$ vďaka leme 3.5.7. Predpokladajme teraz, že E, F sú platné výrazy, pre ktoré existujú „bezepsilonové“ automaty $\mathcal{A}_1, \mathcal{A}_2$ také, že $\|\mathcal{A}_1\| = \|E\|$ a $\|\mathcal{A}_2\| = \|F\|$. Ak $G = E + F$, existuje „bezepsilonový“ automat $\mathcal{A}$ spĺňajúci $\|\mathcal{A}\| = \|E\| + \|F\| = \|\mathcal{A}_1\| + \|\mathcal{A}_2\|$ podľa lemy 3.5.8. Ak $G = E \cdot F$, „bezepsilonový“ automat $\mathcal{A}$ spĺňajúci $\|\mathcal{A}\| = \|E\| \cdot \|F\| = \|\mathcal{A}_1\| \cdot \|\mathcal{A}_2\|$ existuje podľa lemy 3.5.9. Ak je napokon výraz E navyše „zaručene bezepsilonový“ a $G = E^*$, nutne $(\|E\|, \varepsilon) = 0$ (poznámka 3.5.4) a automat $\mathcal{A}$ spĺňajúci $\|\mathcal{A}\| = \|E\|^* = \|\mathcal{A}_1\|^*$ existuje podľa lemy 3.5.9.

V prípade, že je S spočítateľne úplný polokruh, vyplýva ekvivalencia tvrdení (i) a (iii) z vety 3.4.2. Ekvivalencia tvrdení (iii) a (iv) nakoniec vyplýva bezprostredne z vety 1.8.9. $\square$

Definícia 3.5.12. Nech Σ je abeceda a S je polokruh. Formálny mocninový rad $r \in S\langle\langle \Sigma^* \rangle\rangle$ nazveme *racionálnym nad S* , ak existuje platný racionálny výraz E s váhami nad S a Σ taký, že $\|E\| = r$. Množinu všetkých racionálnych radov nad S označíme $S\text{-Rat}(\Sigma^*)$.

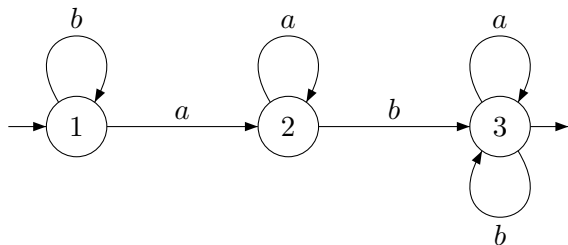
Poznámka 3.5.13. Priamo z definícií 3.5.2 a 3.5.3 vyplýva, že množina $S\text{-Rat}(\Sigma^*)$ je pre všetky abecedy Σ a polokruhy S uzavretá na súčet a na Cauchyho súčin. Z vety 3.5.11 a lemy 3.5.10 ďalej vyplýva, že množina $S\text{-Rat}(\Sigma^*)$ je uzavretá aj na iteráciu aplikovanú na rady s nulovým konštantným koeficientom. Pre spočítateľne úplné polokruhy S je množina $S\text{-Rat}(\Sigma^*)$ uzavretá, vďaka vete 3.5.11 a definíciám 3.5.1 a 3.5.3 (alebo prípadne len vďaka uzavretosti $\text{Rat}(S\langle \Sigma \cup \{\varepsilon\} \rangle, S\langle\langle \Sigma^* \rangle\rangle)$ na iteráciu), na iteráciu bez ďalšieho obmedzenia.

Príklad 3.5.14. Skôr, než budeme pokračovať ďalej, ukážeme si jednoduchú kombinatorickú aplikáciu automatov a racionálnych výrazov s váhami. Predpokladajme, že sa zaujímate o počet všetkých slov dĺžky t nad abecedou $\Sigma = \{a, b\}$, ktoré obsahujú ako faktor slovo ab . Túto úlohu je samozrejme možné ľahko vyriešiť aj s použitím elementárnej kombinatoriky: keďže faktor ab obsahujú práve všetky slová zo Σ^t , ktoré *nie sú* tvaru $b^k a^{t-k}$ pre nejaké $k \in \{0, \dots, t\}$, je hľadaný počet slov rovný

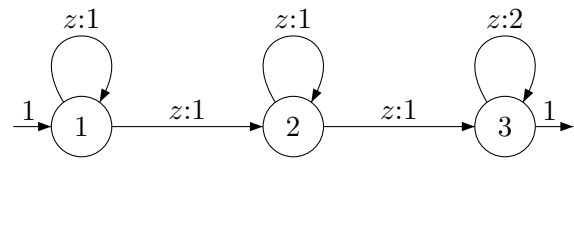
$$\left| \Sigma^t \setminus \left\{ b^k a^{t-k} \mid k \in \{0, \dots, t\} \right\} \right| = |\Sigma^t| - \left| \left\{ b^k a^{t-k} \mid k \in \{0, \dots, t\} \right\} \right| = 2^t - (t + 1).$$

V nasledujúcom prideme k rovnakému výsledku s použitím automatov a racionálnych výrazov s váhami. Výhodou tohto prístupu bude jeho väčšia systematickosť, vďaka ktorej je aplikovateľný aj na podstatne komplikovanejšie úlohy, pri ktorých elementárne metódy zlyhávajú.

Nech $L = \Sigma^* ab \Sigma^*$ je jazyk všetkých slov nad abecedou Σ obsahujúcich faktor ab . V úlohe sa teda pýtame na počet slov dĺžky t v jazyku L . Skonstruujme *deterministický* konečný automat $\mathcal{A}$ rozoznávajúci jazyk L a ten prerobme, rovnako ako v príklade 3.1.2, na konečný automat $\mathcal{A}'$ s váhami nad polokruhom prirodzených čísel a unárnou vstupnou abecedou $\{z\}$, ktorý realizuje vytvárajúcu funkciu pre počet slov danej dĺžky v jazyku $\|\mathcal{A}\| = L$. Automat $\mathcal{A}$ je znázornený na obrázku 3.7a a príslušný automat $\mathcal{A}'$ je na obrázku 3.7b.



(a) Deterministický konečný automat $\mathcal{A}$.



(b) Konečný automat s váhami $\mathcal{A}'$.

Obr. 3.7: Deterministický konečný automat $\mathcal{A}$ rozoznávajúci jazyk $L = \Sigma^* ab \Sigma^*$ a konečný automat $\mathcal{A}'$ s váhami nad $(\mathbb{N}, +, \cdot, 0, 1)$ realizujúci vytvárajúcu funkciu pre počet slov danej dĺžky v jazyku L .

Môžeme teraz spočítať formálny mocninový rad realizovaný automatom $\mathcal{A}'$:

$$\|\mathcal{A}'\| = (1, 0, 0) \begin{pmatrix} z & z & 0 \\ 0 & z & z \\ 0 & 0 & 2z \end{pmatrix}^* \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = z^* z z^* z (2z)^*;$$

pri výpočte sme použili dôsledok 2.5.6 spoločne s pozorovaním, že vďaka tvaru vektorov počiatkových a koncových váh je výsledný formálny mocninový rad rovný radu v prvom riadku a treťom stĺpci iterácie prechodovej matice. Rad $\|\mathcal{A}'\|$ teraz môžeme stotožniť s radom v $\mathbb{C}\langle\langle z \rangle\rangle$, vďaka čomu dostávame

$$\|\mathcal{A}'\| = z^* z z^* z (2z)^* = \frac{z^2}{(1-z)^2(1-2z)} = \frac{1}{1-2z} - \frac{1}{(1-z)^2},$$

kde poslednú rovnosť možno odvodiť pomocou rozkladu na parciálne zlomky. Ak nás teda zaujíma počet slov dĺžky t v jazyku L , stačí sa pozrieť na koeficient pri z^t vo vytvárajúcej funkcii $\|\mathcal{A}'\|$:

$$\begin{aligned} [z^t]\|\mathcal{A}'\| &= [z^t] \left(\frac{1}{1-2z} - \frac{1}{(1-z)^2} \right) = [z^t] \frac{1}{1-2z} - [z^t] \frac{1}{(1-z)^2} = \\ &= 2^t - \sum_{k=0}^t \left([z^k] \frac{1}{1-z} \right) \left([z^{t-k}] \frac{1}{1-z} \right) = 2^t - \sum_{k=0}^t 1 = 2^t - (t+1). \end{aligned}$$

Všetkých slov dĺžky t nad abecedou Σ obsahujúcich faktor ab je teda skutočne $2^t - (t+1)$.

Zvyšok oddielu venujeme vysvetleniu pomenovaní ako „racionálny výraz“, „racionálny prvok“, alebo „racionálny rad“. Nasledujúci príklad ukazuje, že množina *radov* v $\mathbb{C}\langle\langle z^* \rangle\rangle$ *racionálnych* nad $\mathbb{C}$ – tie možno chápať aj ako prvky $\mathbb{C}\langle\langle z \rangle\rangle$ – pozostáva z práve všetkých formálnych Maclaurinových rozvojev *racionálnych funkcií* analytických v bode $z = 0$. Rovnako sa dá argumentovať aj nad $\mathbb{R}\langle\langle z^* \rangle\rangle$ resp. $\mathbb{R}\langle\langle z \rangle\rangle$.

Príklad 3.5.15. Uvažujme polokruh $\mathbb{F}\langle\langle z^* \rangle\rangle$, kde $\mathbb{F}$ je pole. Keďže je tu abeceda unárna, možno rady z tohto polokruhu chápať aj ako prvky $\mathbb{F}\langle\langle z \rangle\rangle$. Budeme teda tieto dva izomorfné polokruhy – v tomto prípade ide dokonca o obory integrity – a aj ich vzájomne zodpovedajúce prvky stotožňovať.

Komplexná *funkcia* f komplexnej premennej z je *racionálna*, ak $f(z) = p(z)/q(z)$ pre nejaké polynomicke funkcie $p(z)$ a $q(z)$, kde $q(z)$ nie je konštantne nulová. Takáto funkcia je analytická v bode $z = 0$ práve vtedy, keď je rád α koreňa funkcie $q(z)$ v bode 0 menší alebo rovný rádu koreňa funkcie $p(z)$ v tomto bode. V takom prípade možno bez ujmy na všeobecnosti predpokladať, že $q(0) \neq 0$; pôvodný zlomok totiž môžeme vykrátiť funkciou z^α . Maclaurinovým rozvojom polynomickej funkcie je vždy príslušný polynóm – v analýze sa koniec koncov často ani medzi polynómami a polynomickými funkciami nerozlišuje – a požiadavku $q(0) \neq 0$ možno preformulovať ako nenulovosť konštantného koeficientu polynómu prislúchajúceho k polynomickej funkcii $q(z)$.

Prejdime teraz od funkcií k radom z $\mathbb{C}\langle\langle z \rangle\rangle$ resp. $\mathbb{C}\langle\langle z^* \rangle\rangle$. Z vyššie uvedeného vyplýva, že množina *formálnych mocninových radov* zodpovedajúca Maclaurinovým rozvojom *racionálnych funkcií* analytických v bode 0 je daná ako množina $M(\mathbb{C})$ všetkých radov $r \in \mathbb{C}\langle\langle z \rangle\rangle$ takých, že $r = p/q$ pre nejaké polynómy $p, q \in \mathbb{C}[z]$, kde q má nenulový konštantný koeficient. Dokážeme teraz nielen rovnosť $M(\mathbb{C}) = \mathbb{C}\text{-Rat}(z^*)$, ale aj všeobecnejšiu rovnosť $M(\mathbb{F}) = \mathbb{F}\text{-Rat}(z^*)$, kde $\mathbb{F}$ je ľubovoľné pole a množina $M(\mathbb{F})$ pozostáva zo všetkých radov $r \in \mathbb{F}\langle\langle z \rangle\rangle$ takých, že $r = p/q$ pre nejaké $p, q \in \mathbb{F}[z]$, kde konštantný koeficient polynómu q je nenulový.

$\subseteq$: Nech $r = p/q$ pre nejaké polynómy $p, q \in \mathbb{F}[z]$, kde q má nenulový konštantný koeficient. Každý polynóm z $\mathbb{F}[z]$ – alebo pri „nekomutatívnom“ pohľade: z $\mathbb{F}\langle\langle z^* \rangle\rangle$ – je očividne v $\mathbb{F}\text{-Rat}(z^*)$; racionálnymi formálnymi mocninovými radmi nad $\mathbb{F}$ sú teda aj polynómy p a q . Zostáva ukázať, že je racionálnym nad $\mathbb{F}$ aj ich podiel p/q .

Polynóm q možno práve jedným spôsobom rozložiť na rozdiel $q = c_0 - q_+$, kde $c_0 \in \mathbb{F} \setminus \{0\}$ a polynóm q_+ má nulový konštantný koeficient. Potom ale

$$\frac{1}{q} = \frac{1}{c_0 - q_+} = \sum_{t \in \mathbb{N}} \frac{q_+^t}{c_0^{t+1}} = \frac{1}{c_0} \left(\frac{q_+}{c_0} \right)^*.$$

Konštanta $c := 1/c_0$ je evidentne aj racionálnym formálnym mocninovým radom nad $\mathbb{F}$. Samotný podiel p/q je daný ako $p/q = cp(cq_+)^*$ a racionálnosť radu p/q nad $\mathbb{F}$ je tak zrejmým dôsledkom uzáverových vlastností množiny $\mathbb{F}\text{-Rat}(z^*)$.

$\supseteq$: Zjavne stačí štruktúrnou indukciou dokázať, že pre všetky platné výrazy $G \in \mathbb{F}\text{-RatE}(z^*)$ je rad $\|G\|$ daný ako $\|G\| = p/q$ pre nejaké polynómy $p, q \in \mathbb{F}[z]$ (resp. $\mathbb{F}\langle z^* \rangle$), kde q má nenulový konštantný koeficient.

Ak $G = r$ pre nejaké $r \in \mathbb{F}\langle\{z, \varepsilon\}\rangle$, je tvrdenie zrejmé. Ak $G = E + F$ pre platné výrazy E, F , je z indukčného predpokladu $\|E\| = p_1/q_1$ a $\|F\| = p_2/q_2$ pre nejaké polynómy $p_1, p_2, q_1, q_2 \in \mathbb{F}[z]$, kde q_1 a q_2 majú nenulový konštantný koeficient. V dôsledku toho

$$\|G\| = \frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1q_2 + p_2q_1}{q_1q_2},$$

kde polynóm q_1q_2 má takisto nenulový konštantný koeficient. Ak $G = E \cdot F$ pre platné výrazy E, F , z indukčného predpokladu opäť dostávame $\|E\| = p_1/q_1$ a $\|F\| = p_2/q_2$ pre nejaké polynómy $p_1, p_2, q_1, q_2 \in \mathbb{F}[z]$, kde q_1 a q_2 majú nenulový konštantný koeficient. Potom

$$\|G\| = \frac{p_1}{q_1} \cdot \frac{p_2}{q_2} = \frac{p_1p_2}{q_1q_2},$$

kde q_1q_2 má nenulový konštantný koeficient. Ak nakoniec $G = E^*$ pre nejaký platný a „zaručene bezepsilonový“ výraz E , z indukčného predpokladu máme $\|E\| = p/q$ pre nejaké $p, q \in \mathbb{F}[z]$, kde polynóm q má nenulový konštantný koeficient. Navyše¹⁰

$$(\|E\|, \varepsilon) = (p/q, \varepsilon) = 0;$$

keďže $(1/q, \varepsilon) \neq 0$, nutne musí byť $(p, \varepsilon) = 0$. Preto

$$\|G\| = \left(\frac{p}{q} \right)^* = \frac{1}{1 - \frac{p}{q}} = \frac{q}{q - p},$$

kde $q - p$ má nenulový konštantný koeficient vďaka tomu, že $(q, \varepsilon) \neq 0$ a $(p, \varepsilon) = 0$.

3.6 Lineárne reprezentácie

V tomto a aj v niekoľkých nasledujúcich oddieloch budeme pod automatom s váhami vždy rozumieť „bezepsilonový“ automat. Videli sme, že každý takýto končený automat $\mathcal{A}$ s váhami nad polokruhom S a vstupnou abecedou Σ možno zadať vektorom počiatočných váh $\mathbf{i} \in S^{1 \times n}$, prechodovou maticou $A \in (S\langle\Sigma\rangle)^{n \times n}$ a vektorom koncových váh $\mathbf{f} \in S^{n \times 1}$; číslo n tu označuje počet stavov automatu. Platí potom $\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f}$.

Pomocou maticového násobenia teda vieme vyjadriť *formálny mocninový rad* realizovaný automatom s váhami. Občas sa však môže zísť vyjadriť v reči matic aj koeficient jedného konkrétneho slova w v tomto rade – čiže hodnotu $(\|\mathcal{A}\|, w)$. Na to možno použiť *lineárne S -reprezentácie*, ktoré teraz definujeme; S -reprezentáciu získame z automatu s váhami tak, že jeho prechodovú maticu $A \in (S\langle\Sigma\rangle)^{n \times n}$ „rozbijeme“ na niekoľko matic v $S^{n \times n}$ obsahujúcich koeficienty jednotlivých písmen v matici A .

¹⁰V nasledujúcom budeme $\mathbb{F}\langle z^* \rangle$ a $\mathbb{F}[z]$ stotožňovať snáď až príliš arogantne. Čitateľovi zdieľajúcemu tento názor by však nemalo robiť väčšie problémy prísť s korektnejšou – hoci nutne zdĺhavejšou – formalizáciou tejto časti dôkazu.

Definícia 3.6.1. Nech Σ je abeceda a S je polokruh. *Lineárnou S -reprezentáciou* nad Σ nazveme štvoricu $\mathcal{P} = (n, \mathbf{i}, \mu, \mathbf{f})$, kde $n \in \mathbb{N}$ je rád S -reprezentácie, $\mathbf{i} \in S^{1 \times n}$ je vektor počiatkových váh, $\mu: \Sigma^* \rightarrow S^{n \times n}$ je prechodový homomorfizmus¹¹ a $\mathbf{f} \in S^{n \times 1}$ je vektor koncových váh.

Definícia 3.6.2. Nech Σ je abeceda, S je polokruh a $\mathcal{P} = (n, \mathbf{i}, \mu, \mathbf{f})$ je S -reprezentácia nad Σ . *Formálny mocninový rad realizovaný reprezentáciou $\mathcal{P}$* je daný ako

$$\|\mathcal{P}\| = \sum_{w \in \Sigma^*} (\mathbf{i}\mu(w)\mathbf{f})w.$$

Čitateľovi s dostatočnou intuíciou pre maticové násobenie je už na tomto mieste pravdepodobne zrejmé, že S -reprezentácie nad Σ sú len odlišným pohľadom na automaty s váhami nad polokruhom S a abecedou Σ . Dôvod na zavedenie nového pojmu tu pochádza z teórie rozšírených automatov s váhami, ktoré realizujú formálne mocninové rady nad monoidmi (takýmito automatmi sú napríklad prekladače s váhami, ktoré možno chápať ako automaty s váhami nad monoidom $\Sigma^* \times \Sigma^*$); tam sa automaty s váhami a S -reprezentácie razom stávajú fundamentálne odlišnými objektmi. My sa však do takýchto zovšeobecnení púšťať nebudeme a S -reprezentácie pre nás budú naozaj len odlišným spôsobom, ako zadať konečný automat s váhami. To formalizujeme v nasledujúcej definícii a vete.

Definícia 3.6.3. Nech Σ je abeceda, S je polokruh a $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad S a abecedou Σ . *Reprezentáciou zodpovedajúcou automatu $\mathcal{A}$* potom nazveme S -reprezentáciu $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$, kde homomorfizmus $\mu: \Sigma^* \rightarrow S^{n \times n}$ je daný pre všetky $c \in \Sigma$ ako $\mu(c) = (A, c)$.¹²

Veta 3.6.4. *Nech Σ je abeceda, S je polokruh a $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad S a Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Pre všetky $w \in \Sigma^*$ potom $(A^*, w) = \mu(w)$ a $(\|\mathcal{A}\|, w) = \mathbf{i}\mu(w)\mathbf{f}$, v dôsledku čoho $\|\mathcal{A}\| = \|\mathcal{P}_{\mathcal{A}}\|$.*

Dôkaz. Pre všetky $w \in \Sigma^*$ dokážeme indukciou vzhľadom na $|w|$ rovnosť $(A^*, w) = \mu(w)$.

Ak $|w| = 0$, nutne $w = \varepsilon$ a dokazované tvrdenie platí, pretože podľa lemy 2.6.4 a tvrdenia 2.6.3 je $(A^*, \varepsilon) = (\mathbf{I}_n + AA^*, \varepsilon) = (\mathbf{I}_n, \varepsilon) + (AA^*, \varepsilon) = \mathbf{I}_n = \mu(\varepsilon)$. Predpokladajme teda, že tvrdenie platí pre všetky $w \in \Sigma^t$ pre nejaké prirodzené číslo t a uvažujme slovo cu , kde $c \in \Sigma$ a $u \in \Sigma^t$. Z lemy 2.6.4, príslušnosti matice A do $S^{n \times n}(\Sigma)$ a indukčného predpokladu potom

$$\begin{aligned} (A^*, cu) &= (\mathbf{I}_n + AA^*, cu) = (\mathbf{I}_n, cu) + (AA^*, cu) = 0 + \sum_{\substack{x, y \in \Sigma^* \\ xy = cu}} (A, x)(A^*, y) = (A, c)(A^*, u) = \\ &= \mu(c)\mu(u) = \mu(cu), \end{aligned}$$

čím je indukčný krok dokázaný. Pre všetky $w \in \Sigma^*$ teda naozaj $(A^*, w) = \mu(w)$, z čoho dostávame aj

$$(\|\mathcal{A}\|, w) = (\mathbf{i}A^*\mathbf{f}, w) = \mathbf{i}\mu(w)\mathbf{f}.$$

Rovnosť $\|\mathcal{A}\| = \|\mathcal{P}_{\mathcal{A}}\|$ potom vyplýva priamo z definície 3.6.2. $\square$

Pre ľubovoľnú S -reprezentáciu $\mathcal{P} = (n, \mathbf{i}, \mu, \mathbf{f})$ nad abecedou Σ môžeme uvažovať konečný automat $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ s váhami nad S a Σ taký, že

$$A = \sum_{c \in \Sigma} \mu(c)c.$$

Zrejme potom $\mathcal{P}_{\mathcal{A}} = \mathcal{P}$ a z vety 3.6.4 tak dostávame $\|\mathcal{A}\| = \|\mathcal{P}\|$. *Množina radov realizovaných lineárnymi S -reprezentáciami nad Σ je teda práve $S\text{-Rat}(\Sigma^*)$.*¹³

¹¹Ide o homomorfizmus *monoidov* $(\Sigma^*, \cdot, \varepsilon)$ a $(S^{n \times n}, \cdot, \mathbf{I}_n)$. Keďže je monoid $(\Sigma^*, \cdot, \varepsilon)$ voľný, stačí každý takýto homomorfizmus zadať obrazmi $\mu(c)$ jednotlivých písmen $c \in \Sigma$.

¹²Opäť tu využívame stotožňovanie polokruhov $(S\langle\langle\Sigma^*\rangle\rangle)^{n \times n}$ a $S^{n \times n}\langle\langle\Sigma^*\rangle\rangle$.

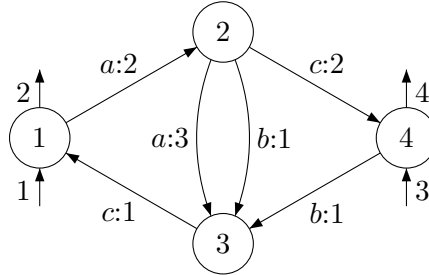
¹³Rady realizované lineárnymi reprezentáciami nad S a Σ nazývame aj *rozoznateľnými* nad S a množinu všetkých takýchto radov označujeme $S\text{-Rec}(\Sigma^*)$. Práve dokázané tvrdenie teda hovorí, že $S\text{-Rec}(\Sigma^*) = S\text{-Rat}(\Sigma^*)$; dôvodom na zavedenie pojmu rozoznateľného radu pritom je, že uvedená rovnosť prestáva platiť akonáhle prejdeme od Σ^* k všeobecnejším monoidom. (Niektorí autori definujú rozoznateľné rady ako rady realizované konečnými automatmi s váhami – takto chápané rozoznateľné rady sú ale v skutočnosti *nerozoznateľné* od racionálnych radov.)

Príklad 3.6.5. Uvažujme konečný automat $\mathcal{A}$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$ a vstupnou abecedou $\Sigma = \{a, b, c\}$, ktorý je znázornený na obrázku 3.8.

Automat $\mathcal{A}$ je daný vektorom počiatkových váh $\mathbf{i} = (1, 0, 0, 3)$, prechodovou maticou

$$A = \begin{pmatrix} 0 & 2a & 0 & 0 \\ 0 & 0 & 3a+b & 2c \\ c & 0 & 0 & 0 \\ 0 & 0 & b & 0 \end{pmatrix}$$

a vektorom koncových váh $\mathbf{f} = (2, 0, 0, 4)^T$.



Obr. 3.8: Konečný automat $\mathcal{A}$ s váhami nad polokruhom $(\mathbb{N} \cup \{\infty\}, +, \cdot, 0, 1)$.

K automatu $\mathcal{A}$ prislúcha reprezentácia $\mathcal{P}_{\mathcal{A}} = (4, \mathbf{i}, \mu, \mathbf{f})$, kde homomorfizmus $\mu: \Sigma^* \rightarrow \mathbb{N}^{4 \times 4}$ je daný nasledovne:

$$\mu(a) = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mu(b) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mu(c) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ľahko možno overiť, že platí napríklad

$$(\|\mathcal{A}\|, aac) = \mathbf{i}\mu(aac)\mathbf{f} = (1, 0, 0, 3) \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 0 \\ 4 \end{pmatrix} = 12.$$

Poznámka 3.6.6. Nech je notácia rovnaká ako vyššie. Ľahko potom vidieť, že okrem „úplných“ súčinov typu $\mathbf{i}\mu(w)\mathbf{f} = (\mathbf{i}A^*\mathbf{f}, w) = (\|\mathcal{A}\|, w)$ možno prirodzene interpretovať aj „čiastočné“ súčiny $\mathbf{i}\mu(w)$ a $\mu(w)\mathbf{f}$. V prvom prípade ide o riadkový vektor

$$\mathbf{i}\mu(w) = (d_1, \dots, d_n),$$

kde d_i je pre $i = 1, \dots, n$ rovné koeficientu slova w v rade realizovanom automatom, ktorý vznikne z automatu $\mathcal{A}$ zmenou koncovej váhy stavu i na 1 a ostatných koncových váh na 0. Intuitívne možno d_i chápať ako súčet hodnôt „neúplných“ behov na slove w končiacich v stave i ; „neúplnosťou“ behu pritom máme na mysli to, že sa do jeho hodnoty započítava počiatková váha prvého stavu a váhy jednotlivých prechodov, nie však koncová váha posledného stavu. Podobne

$$\mu(w)\mathbf{f} = (e_1, \dots, e_n)^T,$$

kde e_j je pre $j = 1, \dots, n$ rovné hodnote slova w v automate, ktorý vznikne z automatu $\mathcal{A}$ zmenou počiatkovej váhy stavu j na 1 a ostatných počiatkových váh na 0. Intuitívne možno e_j chápať ako súčet hodnôt „neúplných“ behov na slove w začínajúcich v stave j ; v tomto prípade máme „neúplnosťou“ behu na mysli to, že sa do jeho hodnoty započítavajú váhy jednotlivých prechodov a koncová váha posledného stavu, nie však počiatková váha prvého stavu.

3.7 Automaty s váhami nad poľom

V nasledujúcom budeme skúmať konečné automaty s váhami nad veľmi špeciálnou triedou polokruhov: nad poľami. Pre každé pole $\mathbb{F}$ tvorí polokruh $\mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ súčasne aj *vektorový priestor* nad poľom $\mathbb{F}$. Automaty s váhami nad $\mathbb{F}$ dáme do súvisu s konečnorozmernými podpriestormi tohto priestoru, čo neskôr využijeme na rozpracovanie teórie *minimalizácie* automatov s váhami nad poľom.¹⁴

Poznámka 3.7.1. Stále sa budeme držať dohovoru z predchádzajúceho oddielu, podľa ktorého máme pod automatom s váhami nad polokruhom S vždy na mysli „bezepsilonový“ automat; polokruh S tak *nemusi* byť spočítateľne úplný. Budeme uvažovať prípad, keď je týmto polokruhom S nejaké pole $\mathbb{F}$.¹⁵

Teraz položíme *teoretické základy* pre štúdium minimalizácie automatov s váhami nad poľom. Pre každý racionálny formálny mocninový rad $r \in \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ opíšeme *nejaký*¹⁶ automat $\mathcal{A}$ s váhami nad $\mathbb{F}$ taký, že $\|\mathcal{A}\| = r$ a automat $\mathcal{A}$ má spomedzi všetkých automatov realizujúcich rad r zaručene *minimálny počet stavov*. Sformulujeme tiež kritérium minimálnosti automatu s váhami nad poľom, ktoré následne pretavíme aj do *algoritmu* rozhodujúceho, či je daný automat s váhami nad poľom $\mathbb{F}$ minimálny. Na tieto úvahy neskôr – po krátkej odbočke do lineárnej algebry – naviažeme v oddiele 3.9, kde opíšeme *minimalizačný algoritmus* pre automaty s váhami nad poľom $\mathbb{F}$.

Tvrdenie 3.7.2. *Nech Σ je abeceda a $\mathbb{F}$ je pole. Potom $\mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ tvorí, spoločne s operáciami súčtu radov a súčinu prvku $\mathbb{F}$ s radom,¹⁷ vektorový priestor nad poľom $\mathbb{F}$.*

Dôkaz. Zrejmé. □

Ku každému konečnému automatu $\mathcal{A}$ s váhami nad poľom $\mathbb{F}$ teraz priradíme podpriestor $\mathcal{V}(\mathcal{A})$ vektorového priestoru $\mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ pozostávajúci zo všetkých radov $\|\mathcal{A}'\|$, kde $\mathcal{A}'$ je automat líšiaci sa od $\mathcal{A}$ nanajvýš v počiatočných váhach jednotlivých stavov.

Definícia 3.7.3. Nech $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Ako $\mathcal{V}(\mathcal{A})$ označíme množinu radov

$$\mathcal{V}(\mathcal{A}) := \{\|\mathcal{A}'\| \mid \mathcal{A}' = (n, \mathbf{j}, A, \mathbf{f}) \text{ je konečný automat s váhami nad } \mathbb{F} \text{ a abecedou } \Sigma\},$$

pričom zrejme

$$\mathcal{V}(\mathcal{A}) = \{\mathbf{j}A^*\mathbf{f} \mid \mathbf{j} \in \mathbb{F}^n\} = \left\{ \sum_{w \in \Sigma^*} (\mathbf{j}\mu(w)\mathbf{f})w \mid \mathbf{j} \in \mathbb{F}^n \right\}. \quad (3.5)$$

Tvrdenie 3.7.4. *Nech $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ . Potom $\mathcal{V}(\mathcal{A})$ tvorí vektorový podpriestor priestoru $\mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ a $\dim(\mathcal{V}(\mathcal{A})) \leq n$.*

Dôkaz. Treba dokázať, že pre ľubovoľnú dvojicu radov $r, s \in \mathcal{V}(\mathcal{A})$ a ľubovoľnú dvojicu skalárov $a, b \in \mathbb{F}$ je $ar + bs \in \mathcal{V}(\mathcal{A})$. Keďže $r, s \in \mathcal{V}(\mathcal{A})$, existujú vektory $\mathbf{j}, \mathbf{k} \in \mathbb{F}^n$ také, že $r = \mathbf{j}A^*\mathbf{f}$ a $s = \mathbf{k}A^*\mathbf{f}$. Potom ale

$$ar + bs = a\mathbf{j}A^*\mathbf{f} + b\mathbf{k}A^*\mathbf{f} = (a\mathbf{j} + b\mathbf{k})A^*\mathbf{f} \in \mathcal{V}(\mathcal{A}),$$

čím je dokázané, že $\mathcal{V}(\mathcal{A})$ tvorí vektorový podpriestor. Ten je navyše zjavne generovaný ľubovoľnou množinou $\{\mathbf{i}_1 A^* \mathbf{f}, \dots, \mathbf{i}_n A^* \mathbf{f}\}$, kde $(\mathbf{i}_1, \dots, \mathbf{i}_n)$ je báza priestoru $\mathbb{F}^n$; nutne teda $\dim(\mathcal{V}(\mathcal{A})) \leq n$. □

¹⁴V skutočnosti ale budeme namiesto automatov pracovať hlavne s príslušnými $\mathbb{F}$ -reprezentáciami; v literatúre sa často namiesto o *minimálnych automatoch* hovorí o *minimálnych* alebo *redukovaných reprezentáciách* [11, 94, 95].

¹⁵Čitateľ, ktorý v kapitole 1 vyriešil cvičenie 12, pravdepodobne chápe, že uvažovať automaty s váhami nad spočítateľne úplným poľom by veľmi zmysluplné nebolo.

¹⁶Taký automat totiž obyčajne nebýva určený jednoznačne.

¹⁷Druhú z týchto operácií sme síce explicitne nezaviedli, stotožňovali sme však prvky $\mathbb{F}\langle\{\varepsilon\}\rangle$ s prvkami $\mathbb{F}$. Ide tu teda v skutočnosti o Cauchyho súčin radu z $\mathbb{F}\langle\{\varepsilon\}\rangle$ s radom z $\mathbb{F}\langle\langle\Sigma^*\rangle\rangle$.

V nasledujúcom sa nám zíše operácia *ľavého kvocientu* formálneho mocninového radu z $S\langle\langle\Sigma^*\rangle\rangle$ podľa slova nad abecedou Σ . Pôjde o zovšeobecnenie štandardnej operácie ľavého kvocientu formálneho jazyka podľa slova. Pre formálny mocninový rad r a slovo x bude ľavým kvocientom radu r podľa slova x rad $x^{-1}r$ taký, že koeficient slova w v $x^{-1}r$ je rovný koeficientu slova xw v r . Akoby sme teda z každého člena radu r „odtrhli“ prefix x , pričom pokiaľ rad r obsahuje nenulový člen $(r, w)w$ taký, že x nie je prefixom slova w , stratí sa tento člen pri aplikovaní operácie kvocientu úplne.

Definícia 3.7.5. Nech Σ je abeceda, S je polokruh, $r \in S\langle\langle\Sigma^*\rangle\rangle$ a $x \in \Sigma^*$. Ľavým kvocientom radu r podľa slova x nazveme rad $x^{-1}r$ taký, že pre všetky $w \in \Sigma^*$ je $(x^{-1}r, w) = (r, xw)$. To znamená:

$$x^{-1}r = \sum_{w \in \Sigma^*} (r, xw)w.$$

Tvrdenie 3.7.6. Nech Σ je abeceda a S je polokruh. Potom:

a) Pre všetky $r \in S\langle\langle\Sigma^*\rangle\rangle$, $x \in \Sigma^*$ a $c \in \Sigma$ je

$$(xc)^{-1}r = c^{-1}x^{-1}r.$$

b) Pre všetky $n \in \mathbb{N}$, $r_1, \dots, r_n \in S\langle\langle\Sigma^*\rangle\rangle$, $a_1, \dots, a_n \in S$ a $x \in \Sigma^*$ je

$$x^{-1}(a_1r_1 + \dots + a_nr_n) = a_1x^{-1}r_1 + \dots + a_nx^{-1}r_n.$$

Dôkaz. Jednoduché cvičenie prenechané čitateľovi. □

Tvrdenie 3.7.7. Nech $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$ a $x \in \Sigma^*$. Potom

$$x^{-1}\|\mathcal{A}\| = \sum_{w \in \Sigma^*} ((\mathbf{i}\mu(x))\mu(w)\mathbf{f})w = (\mathbf{i}\mu(x))A^*\mathbf{f}$$

a rad $x^{-1}\|\mathcal{A}\|$ preto patrí do $\mathcal{V}(\mathcal{A})$.

Dôkaz. Jednoduchými úpravami dostávame

$$x^{-1}\|\mathcal{A}\| = \sum_{w \in \Sigma^*} (\|\mathcal{A}\|, xw)w = \sum_{w \in \Sigma^*} (\mathbf{i}\mu(xw)\mathbf{f})w = \sum_{w \in \Sigma^*} ((\mathbf{i}\mu(x))\mu(w)\mathbf{f})w = (\mathbf{i}\mu(x))A^*\mathbf{f}.$$

Rad $x^{-1}\|\mathcal{A}\|$ potom patrí do $\mathcal{V}(\mathcal{A})$ podľa (3.5). □

Ku každému radu $r \in \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ teraz kanonicky priradíme vektorový podpriestor priestoru $\mathbb{F}\langle\langle\Sigma^*\rangle\rangle$, ktorým bude priestor $\mathcal{Q}(r)$ generovaný ľavými kvocientmi radu r podľa slov nad abecedou Σ . Ako je dobre známe, v klasickej teórii automatov zodpovedajú ľavé kvocienty jazyka L stavom minimálneho deterministického konečného automatu rozoznávajúceho jazyk L ; počet rôznych kvocientov jazyka teda určuje počet stavov minimálneho automatu, ktorý ho rozoznáva. Uvidíme, že priestor $\mathcal{Q}(r)$ zohráva podobnú úlohu v teórii automatov s váhami nad poľom: jeho *dimenzia* určuje počet stavov minimálneho automatu realizujúceho rad r .¹⁸ Dimenziu priestoru $\mathcal{Q}(r)$ teda bude možné chápať ako „mieru stavovej zložitosti“ radu r .

Definícia 3.7.8. Nech Σ je abeceda, $\mathbb{F}$ je pole a $r \in \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$. Priestorom kvocientov radu r nazveme podpriestor $\mathcal{Q}(r)$ vektorového priestoru $\mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ generovaný množinou $\{x^{-1}r \mid x \in \Sigma^*\}$.

Priestor $\mathcal{Q}(r)$ je evidentne podpriestorom všetkých vektorových priestorov $\mathcal{V}(\mathcal{A})$, kde $\mathcal{A}$ je automat s váhami nad $\mathbb{F}$ realizujúci rad r – podľa tvrdenia 3.7.7 sú totiž všetky ľavé kvocienty radu r , čiže všetky generátory priestoru $\mathcal{Q}(r)$, prvkami $\mathcal{V}(\mathcal{A})$.

¹⁸Ak bude priestor $\mathcal{Q}(r)$ nekonečnorozmerný, pôjde o známku toho, že rad r nie je racionálny nad $\mathbb{F}$ – nie je teda realizovaný žiadnym konečným automatom s váhami nad $\mathbb{F}$.

Poznámka 3.7.9. Je zrejmé, že ak $r \neq 0$, nutne $\dim(\mathcal{Q}(r)) \geq 1$. Pre $r = 0$ je $\dim(\mathcal{Q}(r)) = 0$. Jediným radom realizovaným nulastavovým automatom s váhami je pritom tiež $r = 0$.

Budeme teraz dokazovať, že dimenzia priestoru $\mathcal{Q}(r)$, pokiaľ je konečná, skutočne určuje počet stavov minimálneho automatu s váhami nad $\mathbb{F}$ realizujúceho rad r . Začnime triviálnym pozorovaním, že priestor $\mathcal{Q}(r)$ nemôže mať dimenziu väčšiu, než je počet stavov hociktorého konečného automatu $\mathcal{A}$ realizujúceho rad r – priestor $\mathcal{Q}(r)$ je totiž podpriestorom $\mathcal{V}(\mathcal{A})$.

Lema 3.7.10. *Nech Σ je abeceda, $\mathbb{F}$ je pole a $r \in \mathbb{F}\langle\Sigma^*\rangle$. Ak $r = \|\mathcal{A}\|$ pre konečný automat $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ s váhami nad $\mathbb{F}$ a abecedou Σ , tak $\dim(\mathcal{Q}(r)) \leq \dim(\mathcal{V}(\mathcal{A})) \leq n$.*

Dôkaz. Ide o bezprostredný dôsledok tvrdenia 3.7.4 a skutočnosti, že $\mathcal{Q}(r)$ je podpriestorom $\mathcal{V}(\mathcal{A})$. $\square$

Lemu 3.7.10 teraz doplníme pozorovaním, podľa ktorého pre každý formálny mocninový rad r s $\dim(\mathcal{Q}(r)) = n$ existuje n -stavový konečný automat s váhami $\mathcal{A}$ taký, že $\|\mathcal{A}\| = r$. Dimenzia priestoru $\mathcal{Q}(r)$ je tak rovná *minimálnemu počtu stavov* konečného automatu s váhami realizujúceho rad r .

Lema 3.7.11. *Nech Σ je abeceda, $\mathbb{F}$ je pole a $r \in \mathbb{F}\langle\Sigma^*\rangle$. Ak $\dim(\mathcal{Q}(r)) = n$ pre nejaké $n \in \mathbb{N}$, tak existuje konečný automat $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ s váhami nad $\mathbb{F}$ a abecedou Σ taký, že $\|\mathcal{A}\| = r$.*

Dôkaz. Ak $n = 0$, nutne $r = 0$ a automat $\mathcal{A}$ existuje. Predpokladajme teda, že $n \geq 1$. Nech $\{g_1, \dots, g_n\}$ je množina *kvocientov radu r* generujúca priestor $\mathcal{Q}(r)$. Skonstruujeme $\mathbb{F}$ -reprezentáciu $\mathcal{P} = (n, \mathbf{i}, \mu, \mathbf{f})$ takú, že $\|\mathcal{P}\| = r$; hľadaným automatom s váhami bude automat $\mathcal{A}$ spĺňajúci $\mathcal{P}_{\mathcal{A}} = \mathcal{P}$.

Konstruáciu reprezentácie $\mathcal{P}$ založíme na pozorovaní, že pre všetky $x \in \Sigma^*$ je $(r, x) = (x^{-1}r, \varepsilon)$. Rad $x^{-1}r$ patrí do $\mathcal{Q}(r)$ – existujú teda koeficienty $a_{x,1}, \dots, a_{x,n} \in \mathbb{F}$ také, že

$$x^{-1}r = a_{x,1}g_1 + \dots + a_{x,n}g_n, \quad (3.6)$$

z čoho

$$(r, x) = (x^{-1}r, \varepsilon) = a_{x,1}(g_1, \varepsilon) + \dots + a_{x,n}(g_n, \varepsilon).$$

Stačí teda vziať $\mathbf{f} = ((g_1, \varepsilon), \dots, (g_n, \varepsilon))^T$ a zabezpečiť, aby pre každé $x \in \Sigma^*$ bolo

$$\mathbf{i}\mu(x) = (a_{x,1}, \dots, a_{x,n}). \quad (3.7)$$

V takom prípade totiž

$$(\|\mathcal{P}\|, x) = \mathbf{i}\mu(x)\mathbf{f} = (a_{x,1}, \dots, a_{x,n}) \begin{pmatrix} (g_1, \varepsilon) \\ \vdots \\ (g_n, \varepsilon) \end{pmatrix} = a_{x,1}(g_1, \varepsilon) + \dots + a_{x,n}(g_n, \varepsilon) = (r, x),$$

a teda $\|\mathcal{P}\| = r$, keďže uvedená rovnosť platí pre každé $x \in \Sigma^*$.

Zostáva špecifikovať vektor $\mathbf{i}$ a homomorfizmus $\mu: \Sigma^* \rightarrow \mathbb{F}^{n \times n}$ tak, aby bola pre každé $x \in \Sigma^*$ splnená rovnosť (3.7). Keďže má byť μ homomorfizmus, nutne $\mu(\varepsilon) = \mathbf{I}_n$, a teda $\mathbf{i} = (a_{\varepsilon,1}, \dots, a_{\varepsilon,n})$.¹⁹ Nech sú teraz $y \in \Sigma^*$ a $c \in \Sigma$ ľubovoľné. Ak má byť rovnosť (3.7) splnená pre $x = y$ a pre $x = yc$, musí byť

$$(a_{yc,1}, \dots, a_{yc,n}) = \mathbf{i}\mu(yc) = \mathbf{i}\mu(y)\mu(c) = (a_{y,1}, \dots, a_{y,n})\mu(c)$$

Je navyše zrejmé, že pre homomorfizmus $\mu: \Sigma^* \rightarrow \mathbb{F}^{n \times n}$ spĺňajúci pre všetky $y \in \Sigma^*$ a $c \in \Sigma$ rovnosť

$$(a_{yc,1}, \dots, a_{yc,n}) = (a_{y,1}, \dots, a_{y,n})\mu(c) \quad (3.8)$$

je pre všetky $x \in \Sigma^*$ splnená aj rovnosť (3.7). *Dôkaz teda bude hotový, keď sa nám podarí pre každé $c \in \Sigma$ nájsť maticu $\mu(c)$ vyhovujúcu rovnici (3.8) pre všetky $y \in \Sigma^*$.*

¹⁹Z rovnosti (3.6) vidieť, že takto definované $\mathbf{i}$ je vektorom koeficientov radu r v báze $(g_1, \dots, g_n)$.

Nech je $c \in \Sigma$ dané pevne. Z rovnosti (3.6) a tvrdenia 3.7.6 potom pre všetky $y \in \Sigma^*$ dostávame

$$\begin{aligned} a_{yc,1}g_1 + \dots + a_{yc,n}g_n &= (yc)^{-1}r = c^{-1}y^{-1}r = c^{-1}(a_{y,1}g_1 + \dots + a_{y,n}g_n) = \\ &= a_{y,1}c^{-1}g_1 + \dots + a_{y,n}c^{-1}g_n. \end{aligned} \quad (3.9)$$

Keďže sú uvažované generátory $g_1, \dots, g_n$ priestoru $\mathcal{Q}(r)$ kvocientmi radu r ,²⁰ musia do $\mathcal{Q}(r)$ patriť aj rady $c^{-1}g_1, \dots, c^{-1}g_n$. Pre $i = 1, \dots, n$ teda existujú koeficienty $b_{i,1}, \dots, b_{i,n} \in \mathbb{F}$ také, že

$$c^{-1}g_i = b_{i,1}g_1 + \dots + b_{i,n}g_n.$$

Pre maticu B definovanú ako

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \dots & b_{n,n} \end{pmatrix}$$

tak možno rovnosť (3.9) prepísať ako

$$(a_{yc,1}, \dots, a_{yc,n}) \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} = (a_{y,1}, \dots, a_{y,n}) \begin{pmatrix} c^{-1}g_1 \\ \vdots \\ c^{-1}g_n \end{pmatrix} = (a_{y,1}, \dots, a_{y,n})B \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix},$$

z čoho – keďže sú rady $g_1, \dots, g_n$ lineárne nezávislé – vyplýva

$$(a_{yc,1}, \dots, a_{yc,n}) = (a_{y,1}, \dots, a_{y,n})B.$$

Maticu $\mu(c)$ teda možno definovať predpisom $\mu(c) = B$, čím je konštrukcia reprezentácie $\mathcal{P}$ dokončená. Správnosť tejto konštrukcie je takisto zaručená úvahami vyššie. $\square$

Nasledujúca veta sumarizuje naše doterajšie zistenia o význame dimenzie vektorového priestoru kvocientov $\mathcal{Q}(r)$, ktorá sa niekedy nazýva aj *hodnotou* radu r .

Veta 3.7.12. *Nech Σ je abeceda, $\mathbb{F}$ je pole a $r \in \mathbb{F}\langle\Sigma^*\rangle$. Ak $r \in \mathbb{F}\text{-Rat}(\Sigma^*)$, je dimenzia priestoru $\mathcal{Q}(r)$ rovná minimálnemu počtu stavov konečného automatu $\mathcal{A}$ s váhami nad poľom $\mathbb{F}$ takého, že $\|\mathcal{A}\| = r$. Ak $r \notin \mathbb{F}\text{-Rat}(\Sigma^*)$, je priestor $\mathcal{Q}(r)$ nekonečnorozmerný.*

Dôkaz. Ide o bezprostredný dôsledok lemy 3.7.10 a lemy 3.7.11. $\square$

Minimálnym automatom realizujúcim formálny mocninový rad $r \in \mathbb{F}\text{-Rat}(\Sigma^*)$ nazveme ľubovoľný automat $\mathcal{A} = (Q, \varphi, \iota, \tau)$ s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\|\mathcal{A}\| = r$ a počet stavov automatu $\mathcal{A}$ je minimálny spomedzi všetkých konečných automatov s váhami nad $\mathbb{F}$ realizujúcich rad r . Z vyššie uvedeného vyplýva, že automat $\mathcal{A}$ je minimálny práve vtedy, keď $|Q| = \dim(\mathcal{Q}(r))$. Pre každý rad $r \in \mathbb{F}\text{-Rat}(\Sigma^*)$ môže byť minimálny automat daný ako v dôkaze lemy 3.7.11.

V dôkaze lemy 3.7.11 však ide takpovediac o „nekonštruktívnu konštrukciu“, ktorá neposkytuje žiaden návod, ako k danému automatu $\mathcal{B}$ s $\|\mathcal{B}\| = r$ *algoritmicke* zostrojiť minimálny automat $\mathcal{A}$ taký, že $\|\mathcal{A}\| = r$. Spolieha sa totiž na existenciu *nejakých* generátorov priestoru $\mathcal{Q}(r)$, o ktorých toho len na základe automatu $\mathcal{B}$ zatiaľ veľa povedať nevieme. Ďalej predpokladá vyjadriteľnosť určitých radov v podobe lineárnej kombinácie týchto generátorov; opäť však nie jasné, ako s využitím informácie o automate $\mathcal{B}$ nájsť koeficienty takýchto lineárnych kombinácií.

²⁰Táto voľba generátorov v skutočnosti nie je nijak zásadná. S využitím tvrdenia 3.7.6 sa totiž dá ľahko dokázať, že priestor $\mathcal{Q}(r)$ je uzavretý na ľavé kvocienty podľa slov a nasledujúca argumentácia by tak „prešla“ pre ľubovoľných n generátorov $g_1, \dots, g_n$ priestoru $\mathcal{Q}(r)$.

Budeme sa teraz snažiť tento nedostatok napraviť. Ako prvý krok pritom dokážeme vetu 3.7.15 poskytujúcu kritérium minimality založené *iba na vlastnostiach uvažovaného automatu*. Toto kritérium nám v konečnom dôsledku umožní aj *algoritmicky rozhodovať minimalitu automatov*, čo je problém, ktorým sa budeme zaoberať v samom závere tohto oddielu. K samotnému minimalizačnému algoritmu pre automaty s váhami nad poľom sa dostaneme v oddiele 3.9.

Najskôr ale budeme potrebovať ďalšie dva vektorové priestory – tentokrát pôjde o podpriestory priestoru $\mathbb{F}^n$ – ktoré kanonicky priradíme každému n -stavovému automatu s váhami nad poľom $\mathbb{F}$. Priestor $\mathbb{F}^{n \times 1}$ v nasledujúcom stotožňujeme s priestorom $\mathbb{F}^{1 \times n}$, pričom z kontextu bude vždy zrejmé, či vektory z $\mathbb{F}^n$ chápeme ako riadkové alebo ako stĺpcové.

Definícia 3.7.13. Nech $\mathcal{A}$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ a $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Potom:

- a) *Ľavým vektorovým priestorom* automatu $\mathcal{A}$ nazveme podpriestor $\text{Left}(\mathcal{A})$ priestoru $\mathbb{F}^n$ generovaný množinou $\{\mathbf{i}\mu(w) \mid w \in \Sigma^*\}$.
- b) *Pravým vektorovým priestorom* automatu $\mathcal{A}$ nazveme podpriestor $\text{Right}(\mathcal{A})$ priestoru $\mathbb{F}^n$ generovaný množinou $\{\mu(w)\mathbf{f} \mid w \in \Sigma^*\}$.

Pre n -stavový automat $\mathcal{A}$ s váhami nad $\mathbb{F}$ je zrejmé $\dim(\text{Left}(\mathcal{A})) \leq n$, ako aj $\dim(\text{Right}(\mathcal{A})) \leq n$. Dokážeme teraz vetu, podľa ktorej je automat $\mathcal{A}$ minimálny práve vtedy, keď *v oboch týchto prípadoch nastane rovnosť*; to vyjadríme aj ďalšou podmienkou využívajúcou nasledujúce označenie.

Označenie 3.7.14. Nech $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Ako $\Lambda[\mathcal{A}]$ označíme lineárne zobrazenie $\Lambda[\mathcal{A}]: \text{Left}(\mathcal{A}) \rightarrow \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ dané pre všetky $x \in \Sigma^*$ predpisom $\Lambda[\mathcal{A}](\mathbf{i}\mu(x)) := (\mathbf{i}\mu(x))A^*\mathbf{f} = x^{-1}\|\mathcal{A}\|$.

Uvedenú rovnosť $(\mathbf{i}\mu(x))A^*\mathbf{f} = x^{-1}\|\mathcal{A}\|$ sme pozorovali v rámci tvrdenia 3.7.7. Vďaka nej je zrejmé, že obrazom zobrazenia $\Lambda[\mathcal{A}]: \text{Left}(\mathcal{A}) \rightarrow \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ je vždy priestor kvocientov $\mathcal{Q}(\|\mathcal{A}\|)$.

Veta 3.7.15. Nech $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ . Potom sú nasledujúce tvrdenia ekvivalentné:

- (i) Automat $\mathcal{A}$ minimálny – čiže $n = \dim(\mathcal{Q}(\|\mathcal{A}\|))$.
- (ii) Platí $\dim(\text{Left}(\mathcal{A})) = n$ a zároveň $\dim(\text{Right}(\mathcal{A})) = n$.
- (iii) Platí $\dim(\text{Left}(\mathcal{A})) = n$ a lineárne zobrazenie $\Lambda[\mathcal{A}]$ je injektívne.

Zobrazenie $\Lambda[\mathcal{A}]$ je navyše injektívne, kedykoľvek $\dim(\text{Right}(\mathcal{A})) = n$ (bez ohľadu na dimenziu $\text{Left}(\mathcal{A})$).

Dôkaz. Pre $n = 0$ je tvrdenie triviálne; predpokladajme teda, že $n \geq 1$. Nech $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$ je lineárna $\mathbb{F}$ -reprezentácia zodpovedajúca konečnému automatu $\mathcal{A}$.

Začnime dôkazom ekvivalencie podmienok (i) a (iii). Keďže je priestor kvocientov $\mathcal{Q}(\|\mathcal{A}\|)$ obrazom lineárneho zobrazenia $\Lambda[\mathcal{A}]$, je

$$\dim(\mathcal{Q}(\|\mathcal{A}\|)) \leq \dim(\text{Left}(\mathcal{A})) \leq n.$$

Ak teda $\dim(\mathcal{Q}(\|\mathcal{A}\|)) = n$, nutne aj $\dim(\text{Left}(\mathcal{A})) = n$ a zobrazenie $\Lambda[\mathcal{A}]: \text{Left}(\mathcal{A}) \rightarrow \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ s obrazom $\mathcal{Q}(\|\mathcal{A}\|)$ musí byť injektívne. Ak naopak $\dim(\text{Left}(\mathcal{A})) = n$, vyplýva z injektívnosti zobrazenia $\Lambda[\mathcal{A}]$, že aj jeho obraz – čiže priestor $\mathcal{Q}(\|\mathcal{A}\|)$ – musí byť dimenzie n ; automat $\mathcal{A}$ je teda minimálny.

Dokážeme teraz implikáciu „ak (iii), tak (ii)“. Nech teda $\dim(\text{Left}(\mathcal{A})) = n$ a nech je zobrazenie $\Lambda[\mathcal{A}]$ injektívne; ukážeme, že $\dim(\text{Right}(\mathcal{A})) = n$. Keďže $\dim(\text{Left}(\mathcal{A})) = n$, nutne $\text{Left}(\mathcal{A}) = \mathbb{F}^n$. Z injektívnosti zobrazenia $\Lambda[\mathcal{A}]$ navyše vyplýva $\text{Ker}(\Lambda[\mathcal{A}]) = \{\mathbf{0}\}$. To znamená, že pre $\mathbf{y} \in \text{Left}(\mathcal{A}) = \mathbb{F}^n$ je $\Lambda[\mathcal{A}](\mathbf{y}) = \mathbf{y}A^*\mathbf{f} = \mathbf{0}$ iba ak $\mathbf{y} = \mathbf{0}$.

Rovnosť $\mathbf{y}A^*\mathbf{f} = 0$ je ale splnená práve vtedy, keď pre všetky $w \in \Sigma^*$ je $(\mathbf{y}A^*\mathbf{f}, w) = \mathbf{y}\mu(w)\mathbf{f} = 0$. Ak teda $w_0 \preceq w_1 \preceq w_2 \preceq \dots$ je (napríklad) koreňové usporiadanie slov zo Σ^* , je jediným riešením nekonečného systému rovníc²¹

$$\mathbf{y} \begin{pmatrix} \uparrow & \uparrow & & \\ \mu(w_0)\mathbf{f} & \mu(w_1)\mathbf{f} & \dots & \\ \downarrow & \downarrow & & \end{pmatrix} = (0, 0, \dots)$$

resp.

$$\begin{pmatrix} \leftarrow & (\mu(w_0)\mathbf{f})^T & \rightarrow \\ \leftarrow & (\mu(w_1)\mathbf{f})^T & \rightarrow \\ & \vdots & \end{pmatrix} \mathbf{y}^T = \begin{pmatrix} 0 \\ 0 \\ \vdots \end{pmatrix}$$

v prípade platnosti podmienky (iii) vektor $\mathbf{y} = \mathbf{0}$. Gaussovou elimináciou²² možno tento nekonečný systém upraviť na konečný systém

$$\begin{pmatrix} \leftarrow & \mathbf{v}_1^T & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}_m^T & \rightarrow \end{pmatrix} \mathbf{y}^T = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

s rovnakou množinou riešení, kde $m \leq n$ je prirodzené číslo a $\mathbf{v}_1^T, \dots, \mathbf{v}_m^T$ – a teda aj $\mathbf{v}_1, \dots, \mathbf{v}_m$ – sú lineárne nezávislé vektory, pričom $\mathbf{v}_1, \dots, \mathbf{v}_m \in \text{Right}(\mathcal{A})$. Keďže má byť jediným riešením tohto systému vektor $\mathbf{y} = \mathbf{0}$, nevyhnutne $m = n$. Priestor $\text{Right}(\mathcal{A})$ tak obsahuje n lineárne nezávislých vektorov $\mathbf{v}_1, \dots, \mathbf{v}_n$, v dôsledku čoho $\dim(\text{Right}(\mathcal{A})) = n$.

Zostáva dokázať implikáciu „ak $\dim(\text{Right}(\mathcal{A})) = n$, tak je $\Lambda[\mathcal{A}]$ injektívne“, ktorej dôsledkom je aj implikácia „ak (ii), tak (iii)“. Nech teda $\dim(\text{Right}(\mathcal{A})) = n$. Potom musia existovať $j_1, \dots, j_n \in \mathbb{N}$ také, že vektory $\mu(w_{j_1})\mathbf{f}, \dots, \mu(w_{j_n})\mathbf{f}$ sú lineárne nezávislé. Ak má nejaké $\mathbf{y} \in \text{Left}(\mathcal{A})$ vyhovovať rovnosti $\Lambda[\mathcal{A}](\mathbf{y}) = \mathbf{y}A^*\mathbf{f} = 0$, musí pre $\ell = 1, \dots, n$ byť $(\mathbf{y}A^*\mathbf{f}, w_{j_\ell}) = \mathbf{y}\mu(w_{j_\ell})\mathbf{f} = 0$ – to jest

$$\mathbf{y} \begin{pmatrix} \uparrow & & \uparrow \\ \mu(w_{j_1})\mathbf{f} & \dots & \mu(w_{j_n})\mathbf{f} \\ \downarrow & & \downarrow \end{pmatrix} = (0, \dots, 0)$$

resp.

$$\begin{pmatrix} \leftarrow & (\mu(w_{j_1})\mathbf{f})^T & \rightarrow \\ & \vdots & \\ \leftarrow & (\mu(w_{j_n})\mathbf{f})^T & \rightarrow \end{pmatrix} \mathbf{y}^T = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

To je však možné len pre $\mathbf{y} = \mathbf{0}$ – jadrom $\text{Ker}(\Lambda[\mathcal{A}])$ lineárneho zobrazenia $\Lambda[\mathcal{A}]$ je teda množina $\{\mathbf{0}\}$, čo znamená, že je zobrazenie $\Lambda[\mathcal{A}]$ injektívne. $\square$

Dokázali sme, že automat $\mathcal{A}$ s váhami nad poľom $\mathbb{F}$ je minimálny práve vtedy, keď $\dim(\text{Left}(\mathcal{A})) = n$ a $\dim(\text{Right}(\mathcal{A})) = n$. Aby sme toto kritérium mohli pretaviť do algoritmu rozhodujúceho minimalitu automatov, potrebujeme algoritmus na výpočet dimenzie $\text{Left}(\mathcal{A})$ a $\text{Right}(\mathcal{A})$. Bázy oboch týchto priestorov možno vybrať z ich generujúcich množín

$$\{\mathbf{i}\mu(w) \mid w \in \Sigma^*\} \quad \text{resp.} \quad \{\mu(w)\mathbf{f} \mid w \in \Sigma^*\},$$

kde $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$ je $\mathbb{F}$ -reprezentácia zodpovedajúca automatu $\mathcal{A}$. Takéto bázy sú nutne konečné, keďže dimenzia oboch uvažovaných priestorov je zhora ohraničená počtom stavov automatu $\mathcal{A}$. To znamená, že existujú konečné jazyky $L, R \subseteq \Sigma^*$ také, že $(\mathbf{i}\mu(w) \mid w \in L)$ je báza priestoru $\text{Left}(\mathcal{A})$ a $(\mu(w)\mathbf{f} \mid w \in R)$ je báza priestoru $\text{Right}(\mathcal{A})$.

²¹Šípky v nasledujúcich maticových zápisoch slúžia len na vizuálne zdôraznenie toho, že $\mu(w_0)\mathbf{f}, \mu(w_1)\mathbf{f}, \dots$ sú stĺpcové vektory a $(\mu(w_0)\mathbf{f})^T, (\mu(w_1)\mathbf{f})^T, \dots$ sú riadkové vektory. Takýmto spôsobom budeme občas matice zapisovať aj neskôr.

²²Nie je ťažké ukázať, že klasickú metódu Gaussovej eliminácie možno rozšíriť aj na prípad matíc s konečným počtom stĺpcov a nekonečne veľa riadkami.

V nasledujúcom ukážeme, že pre každý automat $\mathcal{A}$ možno tieto konečné jazyky L a R zvoliť tak, aby L bol uzavretý na prefixy a R bol uzavretý na sufixy – v takom prípade budeme hovoriť o *ľavom* resp. *pravom* *bázickom jazyku* automatu $\mathcal{A}$. Odtiaľ bude už len krok k algoritmu na výpočet ľavého a pravého bázického jazyka, a teda aj dimenzie priestorov $\text{Left}(\mathcal{A})$ a $\text{Right}(\mathcal{A})$.

Definícia 3.7.16. Nech $\mathcal{A}$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ s $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Potom:

- Ľavým bázickým jazykom* automatu $\mathcal{A}$ nazveme konečný jazyk $L \subseteq \Sigma^*$ uzavretý na prefixy²³ taký, že $(\mathbf{i}\mu(w) \mid w \in L)$ je báza priestoru $\text{Left}(\mathcal{A})$.
- Pravým bázickým jazykom* automatu $\mathcal{A}$ nazveme konečný jazyk $R \subseteq \Sigma^*$ uzavretý na sufixy taký, že $(\mu(w)\mathbf{f} \mid w \in R)$ je báza priestoru $\text{Right}(\mathcal{A})$.

Pre „neúplné“ obdoby bázických jazykov budeme navyše používať nasledujúcu terminológiu:

- Ľavým predbázickým jazykom* automatu $\mathcal{A}$ nazveme konečný jazyk $L \subseteq \Sigma^*$ uzavretý na prefixy taký, že $(\mathbf{i}\mu(w) \mid w \in L)$ je systém lineárne nezávislých vektorov v $\text{Left}(\mathcal{A})$.
- Pravým predbázickým jazykom* automatu $\mathcal{A}$ nazveme konečný jazyk $R \subseteq \Sigma^*$ uzavretý na sufixy taký, že $(\mu(w)\mathbf{f} \mid w \in R)$ je systém lineárne nezávislých vektorov v $\text{Right}(\mathcal{A})$.

Lema 3.7.17. Nech $\mathcal{A}$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ . Potom existuje aspoň jeden ľavý bázický jazyk a aspoň jeden pravý bázický jazyk automatu $\mathcal{A}$. Každý ľavý predbázický jazyk automatu $\mathcal{A}$ je navyše podjazykom nejakého ľavého bázického jazyka automatu $\mathcal{A}$ a každý pravý predbázický jazyk automatu $\mathcal{A}$ je podjazykom nejakého pravého bázického jazyka automatu $\mathcal{A}$.

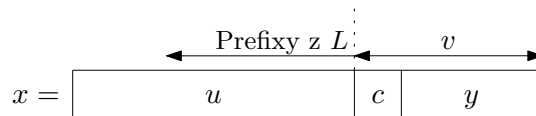
Dôkaz. Nech $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Dokážeme tvrdenia o ľavých bázických a predbázických jazykoch; pre pravé bázické a predbázické jazyky možno argumentovať symetricky.

Ak $\mathbf{i} = \mathbf{0}$, tak $\text{Left}(\mathcal{A}) = \{\mathbf{0}\}$, v dôsledku čoho $\dim(\text{Left}(\mathcal{A})) = 0$ a $L = \emptyset$ je ľavý bázický jazyk a zároveň jediný ľavý predbázický jazyk. Predpokladajme teda, že $\mathbf{i} \neq \mathbf{0}$. Potom $1 \leq \dim(\text{Left}(\mathcal{A})) \leq n$ a je zrejmé, že:

- Jazyk $\{\varepsilon\}$ je ľavý predbázický jazyk automatu $\mathcal{A}$.
- Každý ľavý predbázický jazyk automatu $\mathcal{A}$ môže mať najviac n prvkov.

Vďaka týmto dvom vlastnostiam existuje aspoň jeden *maximálny* ľavý predbázický jazyk L – čiže ľavý predbázický jazyk L taký, že pre žiaden ľavý predbázický jazyk L' nie je $L \subsetneq L'$. Každý ľavý predbázický jazyk navyše musí byť podjazykom nejakého takéhoto maximálneho predbázického jazyka. Na završenie argumentácie tak stačí ukázať, že *ľubovoľný* takýto jazyk L je ľavý bázický.

Sporom. Predpokladajme, že L má uvedenú vlastnosť a súčasne nie je ľavý bázický. Ak je potom $\mathcal{L}$ vektorový priestor generovaný množinou $\{\mathbf{i}\mu(w) \mid w \in L\}$, musí existovať $x \in \Sigma^*$ také, že $\mathbf{i}\mu(x) \notin \mathcal{L}$. Zrejme $x \notin L$. Nech $u \in \Sigma^*$ je *najdlhší* prefix slova x taký, že $u \in L$ – pre nejaké $v \in \Sigma^+$ potom $x = uv$.²⁴ Predpokladajme ďalej, že slovo x bolo zvolené tak, aby sa *minimalizovala dĺžka slova* v . Nech $v = cy$ pre nejaké $c \in \Sigma$ a $y \in \Sigma^*$ – teda $x = ucy$. Táto situácia je znázornená na obrázku 3.9.



Obr. 3.9: Faktorizácia slova $x \in \Sigma^*$ s $\mathbf{i}\mu(x) \notin \mathcal{L}$, ktoré je zvolené tak, aby sa minimalizovala dĺžka slova v .

²³Jazyk L je uzavretý na prefixy (resp. na sufixy), ak pre všetky $u, v \in \Sigma^*$ je $u \in L$ (resp. $v \in L$) kedykoľvek $uv \in L$.

²⁴Takýto prefix u je vždy dobre definovaný, keďže evidentne $\varepsilon \in L$.

Keby bolo $\mathbf{i}\mu(uc) \notin \mathcal{L}$, bol by jazyk $L \cup \{uc\} \supsetneq L$ ľavý predbázický, čo by bol spor s naším predpokladom o jazyku L . Preto $\mathbf{i}\mu(uc) \in \mathcal{L}$, a teda

$$\mathbf{i}\mu(uc) = \sum_{w \in L} a_w \mathbf{i}\mu(w),$$

kde pre každé $w \in L$ je a_w nejaký prvok poľa $\mathbb{F}$. Z toho

$$\mathbf{i}\mu(x) = \mathbf{i}\mu(ucy) = \mathbf{i}\mu(uc)\mu(y) = \left(\sum_{w \in L} a_w \mathbf{i}\mu(w) \right) \mu(y) = \sum_{w \in L} a_w \mathbf{i}\mu(w)\mu(y) = \sum_{w \in L} a_w \mathbf{i}\mu(wy). \quad (3.10)$$

Pre každé $w \in L$ ale musí byť $\mathbf{i}\mu(wy)$ prvkom vektorového priestoru $\mathcal{L}$ – keďže totiž $|y| < |v|$, išlo by v opačnom prípade o spor s voľbou x . Preto musí do $\mathcal{L}$ patriť ľubovoľná lineárna kombinácia takýchto vektorov a podľa (3.10) aj vektor $\mathbf{i}\mu(x)$: spor. $\square$

Na leme 3.7.17 teraz založíme algoritmy hľadajúce, pre daný automat $\mathcal{A}$ s váhami nad poľom $\mathbb{F}$, jeho ľavý resp. pravý bázičný jazyk. Keďže sú $\text{Left}(\mathcal{A})$ a $\text{Right}(\mathcal{A})$ podpriestormi $\mathbb{F}^n$, možno pre n -stavový automat $\mathcal{A}$, vektor $\mathbf{u} \in \text{Left}(\mathcal{A})$ a maticu v redukovanom stupňovitom tvare s nenulovými riadkami $\mathbf{v}_1, \dots, \mathbf{v}_m \in \text{Left}(\mathcal{A})$ s vykonaním $O(mn)$ skalárnych operácií $+$ a $\cdot$ pomocou Gaussovej eliminácie rozhodnúť, či $\mathbf{u}$ patrí do priestoru generovaného množinou $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$. Vtedy možno s vykonaním $O(mn)$ skalárnych operácií vypočítať aj súradnice vektora $\mathbf{u}$ vzhľadom na bázu $(\mathbf{v}_1, \dots, \mathbf{v}_m)$. V opačnom prípade možno s vykonaním $O(mn)$ operácií nájsť vektor $\mathbf{v}_{m+1}$ taký, že matica s nenulovými riadkami $\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}$ (nie nutne v tomto poradí) je redukovaná stupňovitá a priestor generovaný vektormi $\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{v}_{m+1}$ je rovný priestoru generovanému vektormi $\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{u}$. Rovnaké pozorovania možno urobiť aj pre systémy vektorov z $\text{Right}(\mathcal{A})$.

Algoritmus 3.7.18 (Nájdenie ľavého bázičného jazyka).

Vstup: Konečný automat $\mathcal{A}$ s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$.

Výstup: Konečný jazyk $L \subseteq \Sigma^*$.

1. Ak $\mathbf{i} = \mathbf{0}$, polož $L := \emptyset$ a skonči. Inak pokračuj krokom 2.
2. Polož $L := \{\varepsilon\}$, $L_? := \Sigma$ a $L_{\times} := \emptyset$.
3. Opakuj, kým $L_? \neq \emptyset$:
 - 3.1. Nech x je ľubovoľné slovo z $L_?$ (napríklad prvé v koreňovom usporiadaní).
 - 3.2. Over, či $\mathbf{i}\mu(x)$ patrí do priestoru generovaného množinou $\{\mathbf{i}\mu(w) \mid w \in L\}$.
 - 3.2.a. Ak nie, polož $L := L \cup \{x\}$ a $L_? := (L\Sigma \setminus L) \setminus L_{\times}$.
 - 3.2.b. Ak áno, polož $L_? := L_? \setminus \{x\}$ a $L_{\times} := L_{\times} \cup \{x\}$.

Algoritmus 3.7.19 (Nájdenie pravého bázičného jazyka).

Vstup: Konečný automat $\mathcal{A}$ s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$.

Výstup: Konečný jazyk $R \subseteq \Sigma^*$.

1. Ak $\mathbf{f} = \mathbf{0}$, polož $R := \emptyset$ a skonči. Inak pokračuj krokom 2.
2. Polož $R := \{\varepsilon\}$, $R_? := \Sigma$ a $R_{\times} := \emptyset$.
3. Opakuj, kým $R_? \neq \emptyset$:
 - 3.1. Nech x je ľubovoľné slovo z $R_?$ (napríklad prvé v koreňovom usporiadaní).
 - 3.2. Over, či $\mu(x)\mathbf{f}$ patrí do priestoru generovaného množinou $\{\mu(w)\mathbf{f} \mid w \in R\}$.
 - 3.2.a. Ak nie, polož $R := R \cup \{x\}$ a $R_? := (\Sigma R \setminus R) \setminus R_{\times}$.
 - 3.2.b. Ak áno, polož $R_? := R_? \setminus \{x\}$ a $R_{\times} := R_{\times} \cup \{x\}$.

Veta 3.7.20. Algoritmy 3.7.18 a 3.7.19 sa pre každý vstupný automat $\mathcal{A} = (n, \mathbf{i}, A, \mathbf{f})$ zastavia. Výstupom algoritmu 3.7.18 je pritom niektorý ľavý bázičský jazyk L automatu $\mathcal{A}$ a výstupom algoritmu 3.7.19 je pravý bázičský jazyk R automatu $\mathcal{A}$. Počas behu jedného aj druhého algoritmu sa vykoná $O(|\Sigma|n^3)$ elementárnych operácií násobenia a sčítania v poli $\mathbb{F}$.

Dôkaz. Vetu dokážeme len pre algoritmus 3.7.18; v prípade algoritmu 3.7.19 možno argumentovať analogicky. Pred vykonaním prvej iterácie cyklu v kroku 3 sú v algoritme 3.7.18 zjavne splnené nasledujúce invarianty: L je ľavý predbázičský jazyk, $L_\times$ je jazyk „doposiaľ objavených“ slov w takých, že $L \cup \{w\}$ nie je ľavý predbázičský jazyk a $L_\circ = (L\Sigma \setminus L) \setminus L_\times$. Iterácia cyklu v kroku 3 tieto invarianty zachováva.

Je evidentné, že sa počas vykonávania algoritmu z jazyka L neodoberajú žiadne slová. Každý ľavý predbázičský jazyk L je navyše podľa lemy 3.7.17 buď ľavým bázičským jazykom, alebo existuje $x \in \Sigma^* \setminus L$ také, že $L \cup \{x\}$ je ľavý predbázičský jazyk. Z uzavretosti ľavých predbázičských jazykov na prefixy ďalej vyplýva, že sa slovo x dá vybrať tak, aby bolo $x \in L\Sigma \setminus L$. V každej iterácii cyklu musí byť každé slovo x s touto vlastnosťou prvkom jazyka $L_\circ = (L\Sigma \setminus L) \setminus L_\times$. Inak by totiž muselo patriť do $L_\times$; tam sa však v kroku 3.2.b pridávajú iba slová x , ktorú túto vlastnosť *nevratne*²⁵ nemajú. Každé slovo $x \in \Sigma^*$ sa navyše v kroku 3.1 vyberie nanajvýš raz, pretože následne sa buď v kroku 3.2.a pridá do L , alebo sa v kroku 3.2.b pridá do $L_\times$.

Zisťujeme teda napokon, že ak jazyk L nie je ľavý bázičský, musí sa po každom vykonaní kroku 3.2.b niekedy vykonať krok 3.2.a, čím sa počet prvkov v L zvýši o 1. Ľavý predbázičský jazyk však môže mať najviac n prvkov, čo znamená, že v určitom bode vykonávania algoritmu sa jazyk L ľavým bázičským jazykom stane. Potom sa už len pre všetky $x \in L_\circ$ vykoná krok 3.2.b, čím sa jazyk $L_\circ$ vyprázdni a beh algoritmu 3.7.18 sa skončí s ľavým bázičským jazykom na výstupe.

Na časovú zložitosť meranú skalárnymi operáciami poľa $\mathbb{F}$ má vplyv iba overenie podmienky v kroku 3.2.²⁶ Výsledný jazyk L – označme ho L_{fin} – má najviac n prvkov a každé slovo x vybrané v kroku 3.1 musí patriť do L_{fin} alebo do $L_{\text{fin}}\Sigma \setminus L_{\text{fin}}$; každé slovo x sa navyše v kroku 3.1 vyberie najviac raz. Dohromady je preto všetkých potenciálnych slov x najviac $O(n + |\Sigma|n) = O(|\Sigma|n)$ a najviac $O(|\Sigma|n)$ -krát sa tak vykoná aj krok 3.2, ktorý zakaždým prebehne s vykonaním $O(n^2)$ skalárnych operácií poľa $\mathbb{F}$.²⁷ Výsledný počet vykonaných skalárnych operácií je teda $O(|\Sigma|n^3)$. $\square$

Príklad použitia algoritmov 3.7.18 a 3.7.19 možno nájsť v oddiele 3.9. Po nájdení bázičských jazykov automatu $\mathcal{A}$ týmito algoritmami je už triviálne vypočítať dimenzie priestorov $\text{Left}(\mathcal{A})$ a $\text{Right}(\mathcal{A})$ – ide totiž o počty prvkov ľavého resp. pravého bázičského jazyka. Je tak možné algoritmicke overiť platnosť podmienky (ii) lemy 3.7.15, a teda aj *algoritmicky rozhodnúť minimalitu* automatu $\mathcal{A}$.

3.8 Odbočka do lineárnej algebry

Na čo možno najpriamočiarejší opis minimalizačného algoritmu pre automaty s váhami nad poľom budeme potrebovať jednu myšlienku z lineárnej algebry, s ktorou sa predpokladaný čitateľ tohto textu doposiaľ nemusel stretnúť. Je známe, že ku každej *nesingulárnej* (to jest *regulárnej*) štvorcovej matici $A \in \mathbb{F}^{n \times n}$ nad poľom $\mathbb{F}$ – čiže matici $A \in \mathbb{F}^{n \times n}$ s hodnotou n – existuje práve jedna inverzná matica A^{-1} taká, že $A^{-1}A = AA^{-1} = \mathbf{I}_n$. V tomto oddiele ukážeme, že pojem inverznej matice možno čiastočne rozšíriť aj na obdĺžnikové matice.

Nech $A = (a_{i,j})_{m \times n}$ je matica typu $m \times n$ nad poľom $\mathbb{F}$ s *riadkovou hodnotou* m . Riadková hodnota matice je definovaná ako dimenzia vektorového priestoru generovaného jej riadkami; preto nutne $m \leq n$. Matica A je maticou lineárneho zobrazenia $f: \mathbb{F}^m \rightarrow \mathbb{F}^n$ daného pre $i = 1, \dots, m$ ako

$$f(\mathbf{e}_i) = (a_{i,1}, \dots, a_{i,n}),$$

²⁵Ak $L \cup \{x\}$ nie je ľavý predbázičský jazyk, tak neexistuje ani žiaden jazyk $L' \supseteq L$ taký, že $L' \cup \{x\}$ je ľavý predbázičský jazyk.

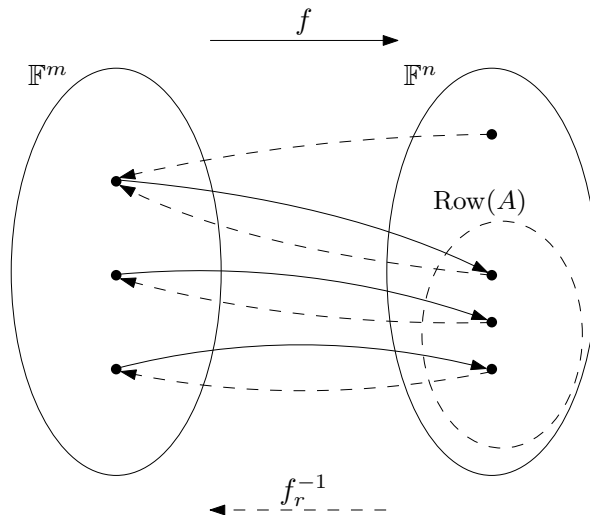
²⁶Napriek tomu ale ide o „rozumnú“ mieru časovej zložitosti – aj pri reálnej implementácii je overenie podmienky v kroku 3.2 zďaleka časovo najnáročnejšou časťou celého výpočtu.

²⁷Pri implementácii s *postupným* budovaním redukovanej stupňovitej matice.

kde $\mathbf{e}_i$ označuje vektor z $\mathbb{F}^m$ s jednotkou na i -tej pozícii a s nulami na zvyšných pozíciách. Keďže sú riadky $(a_{i,1}, \dots, a_{i,n})$ matice A pre $i = 1, \dots, m$ lineárne nezávislé, je zobrazenie f injektívne; systém vektorov $((a_{i,1}, \dots, a_{i,n}) \mid i \in [m])$ navyše možno doplniť na bázu priestoru $\mathbb{F}^n$. Musí teda existovať aspoň jedno lineárne zobrazenie $f_r^{-1}: \mathbb{F}^n \rightarrow \mathbb{F}^m$ také, že pre $i = 1, \dots, m$ je

$$f_r^{-1}(a_{i,1}, \dots, a_{i,n}) = \mathbf{e}_i,$$

pričom možno ľahko dokázať, že zobrazenie $f_r^{-1} \circ f: \mathbb{F}^m \rightarrow \mathbb{F}^m$ je identické.²⁸ Maticou lineárneho zobrazenia f_r^{-1} je potom matica A_r^{-1} typu $n \times m$, pre ktorú $AA_r^{-1} = \mathbf{I}_m$. Ľubovoľnú takúto maticu A_r^{-1} nazveme *pravou inverznou maticou* matice A . Situácia je znázornená na obrázku 3.10.



Obr. 3.10: Lineárne zobrazenia f a f_r^{-1} prislúchajúce k maticiam A a A_r^{-1} .

Je navyše zrejmá ďalšia vlastnosť lineárnych zobrazení f a f_r^{-1} : ak $\text{Row}(A)$ je *riadkový priestor* matice A – čiže vektorový priestor generovaný riadkami matice A – musí byť na tomto podpriestore priestoru $\mathbb{F}^n$ identické aj zobrazenie $f \circ f_r^{-1}$. Pre všetky $(c_1, \dots, c_n) \in \text{Row}(A)$ totiž existujú skaláry $d_1, \dots, d_m \in \mathbb{F}$ také, že $(c_1, \dots, c_n) = d_1(a_{1,1}, \dots, a_{1,n}) + \dots + d_m(a_{m,1}, \dots, a_{m,n})$. Potom

$$f_r^{-1}(c_1, \dots, c_n) = d_1 f_r^{-1}(a_{1,1}, \dots, a_{1,n}) + \dots + d_m f_r^{-1}(a_{m,1}, \dots, a_{m,n}) = d_1 \mathbf{e}_1 + \dots + d_m \mathbf{e}_m,$$

z čoho

$$f(f_r^{-1}(c_1, \dots, c_n)) = d_1 f(\mathbf{e}_1) + \dots + d_m f(\mathbf{e}_m) = d_1(a_{1,1}, \dots, a_{1,n}) + \dots + d_m(a_{m,1}, \dots, a_{m,n}) = (c_1, \dots, c_n).$$

Pre všetky $\mathbf{v} \in \text{Row}(A)$ tak musí byť aj $\mathbf{v} A_r^{-1} A = \mathbf{v}$.

Nech je ďalej $B \in \mathbb{F}^{n \times n}$ matica prislúchajúca k nejakému lineárnemu zobrazeniu $g: \mathbb{F}^n \rightarrow \mathbb{F}^n$. Hovoríme, že podpriestor $\mathcal{V}$ vektorového priestoru $\mathbb{F}^n$ je *invariantný* vzhľadom na zobrazenie g alebo vzhľadom na maticu B , ak pre všetky $\mathbf{v} \in \mathcal{V}$ je $g(\mathbf{v}) = \mathbf{v}B \in \mathcal{V}$ – prvky podpriestoru $\mathcal{V}$ sa teda zobrazia späť do $\mathcal{V}$. Ak je teraz vektorový priestor $\text{Row}(A)$ invariantný vzhľadom na zobrazenie g , obrazom zobrazenia $g \circ f: \mathbb{F}^m \rightarrow \mathbb{F}^n$ je nejaký podpriestor priestoru $\text{Row}(A)$. Z predchádzajúceho teda vidieť, že zobrazenie $f \circ f_r^{-1} \circ g \circ f$ je rovné zobrazeniu $g \circ f$. V reči matíc: ak je priestor $\text{Row}(A)$ invariantný vzhľadom na B , je $ABA_r^{-1}A = AB$.

Obdobne možno ukázať, že ku každej matici $A = (a_{i,j})_{m \times n}$ typu $m \times n$ nad poľom $\mathbb{F}$ so *stĺpcovou hodnotou* n – tu musí byť $m \geq n$ – existuje aspoň jedna matica A_ℓ^{-1} typu $n \times m$ taká, že $A_\ell^{-1}A = \mathbf{I}_n$. Ľubovoľnú takúto maticu A_ℓ^{-1} nazveme *ľavou inverznou maticou* matice A . Pre všetky stĺpcové vektory $\mathbf{v} \in \text{Col}(A)$, kde $\text{Col}(A)$ je *stĺpcový priestor* matice A – čiže vektorový priestor generovaný stĺpcami tejto matice – navyše musí byť $AA_\ell^{-1}\mathbf{v} = \mathbf{v}$. Ak je teda priestor $\text{Col}(A)$ invariantný vzhľadom na maticu $B \in \mathbb{F}^{m \times m}$ – čo v prípade stĺpcových vektorov znamená, že pre všetky $\mathbf{v} \in \text{Col}(A)$ je $B\mathbf{v} \in \text{Col}(A)$ – dostávame rovnosť $AA_\ell^{-1}BA = BA$.

²⁸Pre všetky $(b_1, \dots, b_m) \in \mathbb{F}^m$ je $f(b_1, \dots, b_m) = b_1(a_{1,1}, \dots, a_{1,n}) + \dots + b_m(a_{m,1}, \dots, a_{m,n})$, v dôsledku čoho $f_r^{-1}(f(b_1, \dots, b_m)) = b_1 f_r^{-1}(a_{1,1}, \dots, a_{1,n}) + \dots + b_m f_r^{-1}(a_{m,1}, \dots, a_{m,n}) = b_1 \mathbf{e}_1 + \dots + b_m \mathbf{e}_m = (b_1, \dots, b_m)$.

3.9 Minimalizácia automatov s váhami nad poľom

Máme teraz dostatočný aparát na dôkaz dvojice lem, z ktorých bude evidentný postup konštrukcie minimálneho automatu $\mathcal{B}'$ ekvivalentného danému konečnému automatu $\mathcal{B}$ s váhami nad poľom $\mathbb{F}$. Podľa kritéria dokázaného vo vete 3.7.15 treba automat $\mathcal{B}$ pri minimalizácii prerobiť na automat $\mathcal{B}'$ s m stavmi taký, že *súčasne* $\dim(\text{Left}(\mathcal{B}')) = m$ a $\dim(\text{Right}(\mathcal{B}')) = m$. Lema 3.9.1 nám poskytne „návod“ na konštrukciu automatu $\mathcal{B}'$ takého, že platí prvá z týchto rovností; ak pritom bolo lineárne zobrazenie $\Lambda[\mathcal{B}]$ injektívne, bude injektívne aj zobrazenie $\Lambda[\mathcal{B}']$. Z lemy 3.9.2 bude naopak zrejmy postup konštrukcie automatu, pre ktorý je splnená druhá z uvedených rovností.

Spojením týchto dvoch procedúr napokon získame samotný minimalizačný algoritmus. Ak totiž najprv podľa lemy 3.9.2 zostrojíme k automatu $\mathcal{B}$ automat $\mathcal{B}''$ s k stavmi spĺňajúci $\dim(\text{Right}(\mathcal{B}'')) = k$, bude zobrazenie $\Lambda[\mathcal{B}'']$ injektívne podľa vety 3.7.15. Na automat $\mathcal{B}''$ tak môžeme aplikovať konštrukciu z lemy 3.9.1, ktorej výstupom je m -stavový automat $\mathcal{B}'$ taký, že $\dim(\text{Left}(\mathcal{B}')) = m$ a $\Lambda[\mathcal{B}']$ je injektívne. Z vety 3.7.15 potom vyplynie minimalita automatu $\mathcal{B}'$.²⁹

Intuitívny význam konštrukcie z nasledujúcej lemy je vysvetlený na začiatku jej dôkazu.

Lema 3.9.1. *Nech $\mathcal{A}$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Nech $\dim(\text{Left}(\mathcal{A})) = m \geq 1$ a $L = \{w_1, \dots, w_m\}$, kde $w_1 = \varepsilon$, je ľubovoľný ľavý bázičný jazyk automatu $\mathcal{A}$. Označme ako X maticu typu $m \times n$ nad $\mathbb{F}$ danú ako*

$$X = \begin{pmatrix} \leftarrow & \mathbf{i}\mu(w_1) & \rightarrow \\ \leftarrow & \mathbf{i}\mu(w_2) & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{i}\mu(w_m) & \rightarrow \end{pmatrix}.$$

Riadková hodnota tejto matice je očividne³⁰ m ; nech teda X_r^{-1} je ľubovoľná pravá inverzná matica k matici X . Nech $\mathcal{A}'$ je konečný automat s váhami nad $\mathbb{F}$ a Σ taký, že $\mathcal{P}_{\mathcal{A}'} = (m, \mathbf{i}', \mu', \mathbf{f}')$, kde

$$\mathbf{i}' = (1, 0, \dots, 0), \quad \mu'(c) = X\mu(c)X_r^{-1} \text{ pre všetky } c \in \Sigma \quad \text{a} \quad \mathbf{f}' = X\mathbf{f}.$$

Potom $\|\mathcal{A}'\| = \|\mathcal{A}\|$ a $\dim(\text{Left}(\mathcal{A}')) = m$. Ak je navyše zobrazenie $\Lambda[\mathcal{A}]$ injektívne, je injektívne aj zobrazenie $\Lambda[\mathcal{A}']$.

Dôkaz. Konštrukcia je *intuitívne* založená na pozorovaní, že vektor $\mathbf{i}'\mu'(w) \in \mathbb{F}^m$ je pre všetky $w \in \Sigma^*$ vyjadrením vektora $\mathbf{i}\mu(w) \in \mathbb{F}^n$ v súradniciach bázy $(\mathbf{i}\mu(w_1), \dots, \mathbf{i}\mu(w_m))$ priestoru $\text{Left}(\mathcal{A})$, ktorý je generovaný množinou

$$\{\mathbf{i}\mu(w) \mid w \in \Sigma^*\}.$$

Keďže totiž $w_1 = \varepsilon$, je vektor $\mathbf{i}' = (1, 0, \dots, 0)$ vyjadrením $\mathbf{i} = \mathbf{i}\mu(\varepsilon) = \mathbf{i}\mu(w_1)$ v súradniciach tejto bázy. Ak je teraz pre nejaké $u \in \Sigma^*$ vektor $\mathbf{i}'\mu'(u)$ vyjadrením $\mathbf{i}\mu(u)$ v súradniciach bázy $(\mathbf{i}\mu(w_1), \dots, \mathbf{i}\mu(w_m))$, pre každé $c \in \Sigma$ dostávame

$$\mathbf{i}'\mu'(uc) = \mathbf{i}'\mu'(u)\mu'(c) = \mathbf{i}'\mu'(u)X\mu(c)X_r^{-1}.$$

To možno chápať nasledovne: vektor $\mathbf{i}'\mu'(u)$ sa pre násobením maticou X zmení na $\mathbf{i}\mu(u)$ – matica X totiž transformuje vyjadrenia v súradniciach bázy $(\mathbf{i}\mu(w_1), \dots, \mathbf{i}\mu(w_m))$ na vyjadrenia v štandardných súradniciach priestoru $\mathbb{F}^n$. Ďalším pre násobením maticou $\mu(c)$ dostávame vektor $\mathbf{i}\mu(u)\mu(c) = \mathbf{i}\mu(uc)$.

²⁹Hoci by sme v skutočnosti mohli obidve konštrukcie aplikovať aj v opačnom poradí, z nasledujúcich dvoch lem to vyplývať nebude. Ide o dôsledok asymetrie v definícii zobrazenia $\Lambda[\mathcal{A}]$, v ktorej sa dosadzuje prvok $\text{Left}(\mathcal{A})$ za počiatočný vektor. Rovnako by sme však mohli definovať aj lineárne zobrazenie, v ktorom sa dosadzuje prvok $\text{Right}(\mathcal{A})$ za koncový vektor – preň by sme potom mohli rozvinúť teóriu do istej miery „symetrickú“ tej pre zobrazenie $\Lambda[\mathcal{A}]$ a napokon tak aj dokázať možnosť obrátenia uvedeného postupu. O niečo jednoduchší dôkaz je založený na pozorovaní z cvičenia 16 na konci tejto kapitoly.

³⁰Jej riadky totiž z definície ľavých bázičných jazykov tvoria bázu priestoru $\text{Left}(\mathcal{A})$.

Tento vektor opäť patrí do $\text{Left}(\mathcal{A})$ – priestor $\text{Left}(\mathcal{A})$ je očividne invariantný vzhľadom na $\mu(c)$. Prenásobením vektorov z $\text{Left}(\mathcal{A}) = \text{Row}(X)$ maticou X_r^{-1} sa „správajú inverzne“ k prenasobeniam vektorov z $\mathbb{F}^m$ maticou X – čo znamená, že $\mathbf{i}'\mu'(uc) = \mathbf{i}'\mu'(u)X\mu(c)X_r^{-1} = \mathbf{i}\mu(uc)X_r^{-1}$ je opäť len vyjadrením vektora $\mathbf{i}\mu(uc)$ v súradniciach bázy $(\mathbf{i}\mu(w_1), \dots, \mathbf{i}\mu(w_m))$. Na konci behu na slove w sa už len každý vektor $\mathbf{i}'\mu'(w)$ prenasobí koncovým vektorom $\mathbf{f}' = X\mathbf{f}$ – prenasobenie maticou X „vyrobí“ z vektora $\mathbf{i}'\mu'(w)$ vektor $\mathbf{i}\mu(w)$, čo znamená, že $\mathbf{i}'\mu'(w)\mathbf{f}' = \mathbf{i}\mu(w)\mathbf{f}$ – obidva automaty sú teda ekvivalentné.

Prejdime teraz k *samotnému dôkazu*. Keďže $w_1 = \varepsilon$, je $\mathbf{i}\mu(w_1) = \mathbf{i}$, a teda

$$\mathbf{i}'X = (1, 0, \dots, 0)X = \mathbf{i}.$$

Pre všetky $c \in \Sigma$ je navyše priestor $\text{Left}(\mathcal{A})$ invariantný vzhľadom na maticu $\mu(c)$: pre všetky vektory $\mathbf{v} = \mathbf{i}\mu(x)$ pre $x \in \Sigma^*$, ktorých množina generuje $\text{Left}(\mathcal{A})$, totiž $\mathbf{v}\mu(c) = \mathbf{i}\mu(x)\mu(c) = \mathbf{i}\mu(xc) \in \text{Left}(\mathcal{A})$. Keďže $\text{Left}(\mathcal{A}) = \text{Row}(X)$, z pozorovaní učených v oddiele 3.8 vyplýva

$$\mu'(c)X = X\mu(c)X_r^{-1}X = X\mu(c).$$

Priamo z konštrukcie navyše

$$\mathbf{f}' = X\mathbf{f}.$$

Dohromady teda³¹

$$\mathbf{i}'X = \mathbf{i}, \quad \mu'(c)X = X\mu(c) \text{ pre všetky } c \in \Sigma \quad \text{a} \quad \mathbf{f}' = X\mathbf{f}. \quad (3.11)$$

Nech teraz $w = c_1 \dots c_k$ pre $k \in \mathbb{N}$ a $c_1, \dots, c_k \in \Sigma$. Z (3.11) dostávame

$$\begin{aligned} \mu'(w)X &= \mu'(c_1)\mu'(c_2) \dots \mu'(c_{k-1})\mu'(c_k)X = \mu'(c_1)\mu'(c_2) \dots \mu'(c_{k-1})X\mu(c_k) = \\ &= \dots = \mu'(c_1)X\mu(c_2) \dots \mu(c_{k-1})\mu(c_k) = X\mu(c_1)\mu(c_2) \dots \mu(c_{k-1})\mu(c_k) = X\mu(w) \end{aligned} \quad (3.12)$$

a v dôsledku toho aj

$$(\|\mathcal{A}'\|, w) = \mathbf{i}'\mu'(w)\mathbf{f}' = \mathbf{i}'\mu'(w)X\mathbf{f} = \mathbf{i}'X\mu(w)\mathbf{f} = \mathbf{i}\mu(w)\mathbf{f} = (\|\mathcal{A}\|, w).$$

Ekvivalencia oboch automatov je teda dokázaná, keďže $w \in \Sigma^*$ môže byť ľubovoľné.

Dokážeme teraz, že $\dim(\text{Left}(\mathcal{A}')) = m$. Keďže je $L = \{w_1, \dots, w_m\}$ ľavým bázičným jazykom automatu $\mathcal{A}$, je systém $(\mathbf{i}\mu(w) \mid w \in L)$ bázou priestoru $\text{Left}(\mathcal{A})$. Pre $i = 1, \dots, m$ z (3.12) dostávame

$$\mathbf{i}'\mu'(w_i)X = \mathbf{i}'X\mu(w_i) = \mathbf{i}\mu(w_i).$$

Lineárnym zobrazením – prenasobením maticou X sprava – teda možno z vektorov najviac m -prvkovej množiny $\{\mathbf{i}'\mu'(w) \mid w \in L\}$ získať bázu $(\mathbf{i}\mu(w) \mid w \in L)$ priestoru $\text{Left}(\mathcal{A})$. Priestor generovaný množinou $\{\mathbf{i}'\mu'(w) \mid w \in L\}$ teda má dimenziu m . Navyše je očividne podpriestorom priestoru $\text{Left}(\mathcal{A}')$, ktorý nemôže mať dimenziu väčšiu ako m (lebo $\mathcal{A}'$ má m stavov). Preto $\dim(\text{Left}(\mathcal{A}')) = m$ a L je ľavým bázičným jazykom aj pre automat $\mathcal{A}'$. Samotné násobenie maticou X sprava je navyše nutne injektívnym lineárnym zobrazením na $\text{Left}(\mathcal{A}')$ – táto jeho vlastnosť sa nám zíde o chvíľu.

Zostáva dokázať, že $\Lambda[\mathcal{A}']$ je injektívne, kedykoľvek je injektívne zobrazenie $\Lambda[\mathcal{A}]$. Keby ale $\Lambda[\mathcal{A}']$ injektívne nebolo, muselo by jeho jadro $\text{Ker}(\Lambda[\mathcal{A}'])$ obsahovať nejaký nenulový vektor $\mathbf{v} \in \text{Left}(\mathcal{A}')$; preň však musia existovať $a_1, \dots, a_m \in \mathbb{F}$ také, že $\mathbf{0} \neq \mathbf{v} = a_1\mathbf{i}'\mu'(w_1) + \dots + a_m\mathbf{i}'\mu'(w_m)$. V dôsledku toho

$$\Lambda[\mathcal{A}'](\mathbf{v}) = \Lambda[\mathcal{A}'](a_1\mathbf{i}'\mu'(w_1) + \dots + a_m\mathbf{i}'\mu'(w_m)) = \mathbf{0}.$$

Pre všetky $w \in \Sigma^*$ teda

$$(a_1\mathbf{i}'\mu'(w_1) + \dots + a_m\mathbf{i}'\mu'(w_m))\mu'(w)\mathbf{f}' = a_1\mathbf{i}'\mu'(w_1)\mu'(w)\mathbf{f}' + \dots + a_m\mathbf{i}'\mu'(w_m)\mu'(w)\mathbf{f}' = \mathbf{0};$$

vďaka (3.12) a rovnosti $\mathbf{f}' = X\mathbf{f}$ tak pre všetky $w \in \Sigma^*$ aj

$$a_1\mathbf{i}'\mu'(w_1)X\mu(w)\mathbf{f} + \dots + a_m\mathbf{i}'\mu'(w_m)X\mu(w)\mathbf{f} = (a_1\mathbf{i}'\mu'(w_1)X + \dots + a_m\mathbf{i}'\mu'(w_m)X)\mu(w)\mathbf{f} = \mathbf{0}.$$

³¹Na nasledujúcom riadku vlastne hovoríme, že automaty $\mathcal{A}$ a $\mathcal{A}'$ sú navzájom *konjugované* [8, 95].

Z toho

$$\Lambda[\mathcal{A}](\mathbf{v}X) = \Lambda[\mathcal{A}](a_1\mathbf{i}'\mu'(w_1)X + \dots + a_m\mathbf{i}'\mu'(w_m)X) = 0$$

a vektor $\mathbf{v}X$ je prvkom $\text{Ker}(\Lambda[\mathcal{A}])$. Z injektívnosti zobrazenia $\Lambda[\mathcal{A}]$ následne vyplýva, že musí byť $\mathbf{v}X = \mathbf{0}$, hoci $\mathbf{v} \neq \mathbf{0}$. Násobenie maticou X sprava teda určite nie je injektívnym lineárnym zobrazením na $\text{Left}(\mathcal{A}')$. Vyššie sme ale dokázali, že injektívnym je: spor. $\square$

Lema 3.9.2. *Nech $\mathcal{A}$ je konečný automat s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$. Nech $\dim(\text{Right}(\mathcal{A})) = k \geq 1$ a $R = \{v_1, \dots, v_k\}$, kde $v_1 = \varepsilon$, je ľubovoľný pravý bázičný jazyk automatu $\mathcal{A}$. Označme ako Y maticu typu $n \times k$ nad $\mathbb{F}$ danú ako*

$$Y = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mu(v_1)\mathbf{f} & \mu(v_2)\mathbf{f} & \dots & \mu(v_k)\mathbf{f} \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.$$

Stĺpcová hodnosť tejto matice je zjavne³² k ; nech teda Y_ℓ^{-1} je ľubovoľná ľavá inverzná matica matice Y . Nech $\mathcal{A}'$ je konečný automat s váhami nad $\mathbb{F}$ a Σ taký, že $\mathcal{P}_{\mathcal{A}'} = (k, \mathbf{i}', \mu', \mathbf{f}')$, kde

$$\mathbf{i}' = \mathbf{i}Y, \quad \mu'(c) = Y_\ell^{-1}\mu(c)Y \text{ pre všetky } c \in \Sigma \quad \text{a} \quad \mathbf{f}' = (1, 0, \dots, 0)^T.$$

Potom $\|\mathcal{A}'\| = \|\mathcal{A}\|$ a $\dim(\text{Right}(\mathcal{A}')) = k$.

Dôkaz. Symetricky k dôkazu lemy 3.9.1 (jediná „asymetrická“ časť lemy 3.9.1 a jej dôkazu hovorí o injektívnosti zobrazení $\Lambda[\mathcal{A}]$ a $\Lambda[\mathcal{A}']$; tá je však teraz zo znenia lemy vypustená). $\square$

Následujúci minimalizačný algoritmus pre automaty s váhami nad poľom $\mathbb{F}$ na vstupný automat najprv aplikuje konštrukciu z lemy 3.9.2 a následne konštrukciu z lemy 3.9.1. Na výpočet bázičných jazykov potrebných pri týchto konštrukciách využíva algoritmy 3.7.18 a 3.7.19.

Algoritmus 3.9.3 (Minimalizácia automatu s váhami nad poľom $\mathbb{F}$).

Vstup: Konečný automat $\mathcal{B}$ s váhami nad poľom $\mathbb{F}$ a abecedou Σ taký, že $\mathcal{P}_{\mathcal{B}} = (n, \mathbf{i}, \mu, \mathbf{f})$.

Výstup: Konečný automat $\mathcal{B}'$ s váhami nad poľom $\mathbb{F}$ a abecedou Σ .

1. Nájdí pravý bázičný jazyk $R = \{v_1, \dots, v_k\}$ automatu $\mathcal{B}$, kde $v_1 = \varepsilon$ kedykoľvek $k \geq 1$.
2. Ak $k = 0$, vráť na výstupe (jediný) automat $\mathcal{B}'$ s 0 stavmi a skonči. Inak pokračuj krokom 3.
3. Nech Y je matica so stĺpcami $\mu(v_1)\mathbf{f}, \dots, \mu(v_k)\mathbf{f}$ (v tomto poradí).
4. Vypočítaj ľubovoľnú ľavú inverznú maticu Y_ℓ^{-1} matice Y .
5. Polož $\mathbf{i}'' := \mathbf{i}Y$, $\mu''(c) := Y_\ell^{-1}\mu(c)Y$ pre všetky $c \in \Sigma$ a $\mathbf{f}'' := (1, 0, \dots, 0)^T$.
6. Nech $\mathcal{B}''$ je automat, ktorému zodpovedá reprezentácia $\mathcal{P}_{\mathcal{B}''} = (k, \mathbf{i}'', \mu'', \mathbf{f}'')$.
7. Nájdí ľavý bázičný jazyk $L = \{w_1, \dots, w_m\}$ automatu $\mathcal{B}''$, kde $w_1 = \varepsilon$ kedykoľvek $m \geq 1$.
8. Ak $m = 0$, vráť na výstupe (jediný) automat $\mathcal{B}'$ s 0 stavmi a skonči. Inak pokračuj krokom 9.
9. Nech X je matica s riadkami $\mathbf{i}''\mu''(w_1), \dots, \mathbf{i}''\mu''(w_m)$ (v tomto poradí).
10. Vypočítaj ľubovoľnú pravú inverznú maticu X_r^{-1} matice X .
11. Polož $\mathbf{i}' := (1, 0, \dots, 0)$, $\mu'(c) := X_r^{-1}\mu''(c)X_r$ pre všetky $c \in \Sigma$ a $\mathbf{f}' := X\mathbf{f}''$.
12. Nech $\mathcal{B}'$ je automat, ktorému zodpovedá reprezentácia $\mathcal{P}_{\mathcal{B}'} = (m, \mathbf{i}', \mu', \mathbf{f}')$.

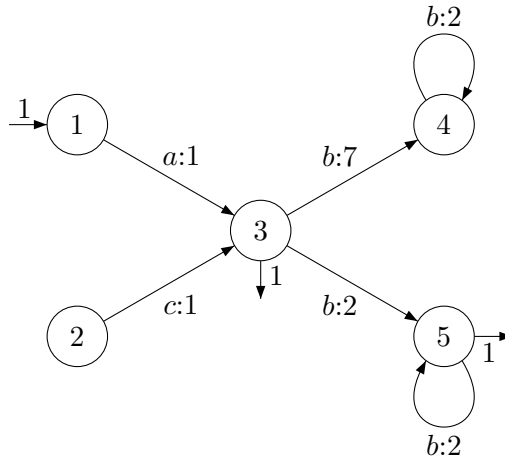
³²Jej stĺpce totiž z definície pravých bázičných množín tvoria bázu priestoru $\text{Right}(\mathcal{A})$.

Sformulujme a dokážme ešte vetu o správnosti a časovej zložitosti práve opísaného minimalizačného algoritmu.

Veta 3.9.4. *Algoritmus 3.9.3 sa pre každý vstupný automat $\mathcal{B}$ zastaví. Jeho výstupom je pritom minimálny automat $\mathcal{B}'$ s váhami nad $\mathbb{F}$ taký, že $\|\mathcal{B}'\| = \|\mathcal{B}\|$ a počas behu algoritmu sa vykoná $O(|\Sigma|n^3)$ elementárnych operácií násobenia a sčítania v poli $\mathbb{F}$.*

Dôkaz. Zastavenie algoritmu na každom vstupe je zrejmé a jeho správnosť je dôsledkom lemy 3.9.1, lemy 3.9.2 a vety 3.7.15. Z vety 3.7.20 navyše vyplýva, že v prípade realizovania kroku 1 a kroku 7 prostredníctvom algoritmov 3.7.19 a 3.7.18 sa vykoná $O(|\Sigma|n^3)$ skalárnych operácií sčítania a násobenia v poli $\mathbb{F}$. Výpočet pravej či ľavej inverznej matice k matici typu $s \times t$ možno vďaka Gaussovej eliminácii realizovať s vykonaním $O(q^3)$ skalárnych operácií, kde $q := \max\{s, t\}$.³³ Počas krokov 4 a 10 sa teda vykoná $O(n^3)$ operácií a maticové násobenie v krokoch 5 a 11 možno tiež realizovať s vykonaním $O(n^3)$ operácií. Celkovo sa teda nevykoná viac než $O(|\Sigma|n^3)$ skalárnych operácií. $\square$

Príklad 3.9.5. Proces minimalizácie automatov s váhami nad poľom prostredníctvom algoritmu 3.9.3 teraz demonštrujeme na príklade. Nech $\mathcal{B}$ je konečný automat s váhami nad poľom reálnych čísel a abecedou $\Sigma = \{a, b, c\}$ znázornený na obrázku 3.11.



Obr. 3.11: Minimalizovaný konečný automat $\mathcal{B}$ s váhami nad poľom $\mathbb{R}$.

Zodpovedajúcou $\mathbb{R}$ -reprezentáciou je teda $\mathcal{P}_{\mathcal{B}} = (5, \mathbf{i}, \mu, \mathbf{f})$, kde $\mathbf{i} = (1, 0, 0, 0, 0)$, homomorfizmus $\mu: \Sigma^* \rightarrow \mathbb{R}^{5 \times 5}$ je daný ako

$$\mu(a) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mu(b) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 2 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}, \quad \mu(c) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

a $\mathbf{f} = (0, 0, 1, 0, 1)^T$. Zrejme

$$\|\mathcal{B}\| = \sum_{t \in \mathbb{N}} 2^t ab^t; \quad (3.13)$$

čitateľ teda bude môcť ľahko overiť správnosť výstupu minimalizačného algoritmu 3.9.3. Pomocou algoritmu 3.7.19 najprv nájdeme pravý bázičný jazyk automatu $\mathcal{B}$ – beh tohto algoritmu pri vstupnom automate $\mathcal{B}$ je opísaný v tabuľke 3.1.

³³V skutočnosti to ide, podobne ako pri násobení matíc, aj rýchlejšie.

Iterácia	x	Jazyk R	Jazyk $R_?$	Jazyk $R_\times$	Báza
0.	—	$\{\varepsilon\}$	$\{a, b, c\}$	$\emptyset$	$\begin{pmatrix} 0 & 0 & 1 & 0 & 1 \end{pmatrix}^T$
1.	a	$\{\varepsilon, a\}$	$\{b, c, aa, ba, ca\}$	$\emptyset$	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}^T$
2.	b	$\{\varepsilon, a\}$	$\{c, aa, ba, ca\}$	$\{b\}$	Nezmenená.
3.	c	$\{\varepsilon, a, c\}$	$\{aa, ac, ba, bc, ca, cc\}$	$\{b\}$	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}^T$
4.	aa	$\{\varepsilon, a, c\}$	$\{ac, ba, bc, ca, cc\}$	$\{b, aa\}$	Nezmenená.
5.	ac	$\{\varepsilon, a, c\}$	$\{ba, bc, ca, cc\}$	$\{b, aa, ac\}$	Nezmenená.
6.	ba	$\{\varepsilon, a, c\}$	$\{bc, ca, cc\}$	$\{b, aa, ac, ba\}$	Nezmenená.
7.	bc	$\{\varepsilon, a, c\}$	$\{ca, cc\}$	$\{b, aa, ac, ba, bc\}$	Nezmenená.
8.	ca	$\{\varepsilon, a, c\}$	$\{cc\}$	$\{b, aa, ac, ba, bc, ca\}$	Nezmenená.
9.	cc	$\{\varepsilon, a, c\}$	$\emptyset$	$\{b, aa, ac, ba, bc, ca, cc\}$	Nezmenená.

Tabuľka 3.1: Beh algoritmu 3.7.19 hľadajúci pravý bázičský jazyk automatu $\mathcal{B}$. V tabuľke je pre každú iteráciu cyklu z kroku 3 uvedené slovo x vybrané v rámci tejto iterácie, ako aj všetky slová v jazykoch R , $R_?$ a $R_\times$ a redukovaná (stĺpcovo) stupňovitá matica, ktorej stĺpce – uvedený zápis cez transponovanú maticu teda treba čítať po riadkoch – tvoria bázu priestoru generovaného množinou $\{\mu(w)\mathbf{f} \mid w \in R\}$, po skončení danej iterácie. Pod nultou iteráciou sa chápe stav pred vykonaním prvej iterácie cyklu.

Zistili sme, že pravým bázičským jazykom automatu $\mathcal{B}$ je $R = \{\varepsilon, a, c\}$. Nech teda

$$Y = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \mu(\varepsilon)\mathbf{f} & \mu(a)\mathbf{f} & \mu(c)\mathbf{f} \\ \downarrow & \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

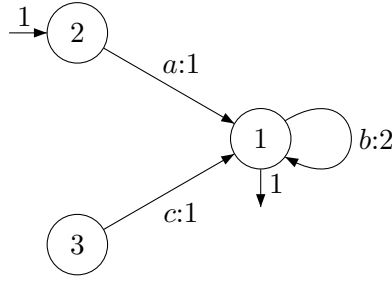
Ľavá inverzná matica k matici Y potom môže byť daná napríklad ako

$$Y_\ell^{-1} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Môžeme teda uvažovať automat $\mathcal{B}''$, ktorému zodpovedá $\mathbb{R}$ -reprezentácia $\mathcal{P}_{\mathcal{B}''} = (3, \mathbf{i}'', \mu'', \mathbf{f}'')$ daná ako $\mathbf{i}'' = \mathbf{i}Y = (0, 1, 0)$,

$$\begin{aligned} \mu''(a) &= Y_\ell^{-1} \mu(a) Y = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \mu''(b) &= Y_\ell^{-1} \mu(b) Y = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ \mu''(c) &= Y_\ell^{-1} \mu(c) Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \end{aligned}$$

a $\mathbf{f}'' = (1, 0, 0)^T$. Pre tento trojstavový automat je $\dim(\text{Right}(\mathcal{B}'')) = 3$ podľa lemy 3.9.2 a z vety 3.7.15 vyplýva aj injektívnosť lineárneho zobrazenia $\Lambda[\mathcal{B}'']$. Automat $\mathcal{B}''$ je znázornený na obrázku 3.12.

Obr. 3.12: Medzivýstup algoritmu 3.9.3 – automat $\mathcal{B}''$.

Pokračujme nájdením ľavého bázičského jazyka automatu $\mathcal{B}''$. Na realizáciu tejto úlohy použijeme algoritmus 3.7.18 – jeho beh pri vstupnom automate $\mathcal{B}''$ je opísaný v tabuľke 3.2.

Iterácia	x	Jazyk L	Jazyk $L_?$	Jazyk $L_\times$	Báza
0.	—	$\{\varepsilon\}$	$\{a, b, c\}$	$\emptyset$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$
1.	a	$\{\varepsilon, a\}$	$\{b, c, aa, ab, ac\}$	$\emptyset$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$
2.	b	$\{\varepsilon, a\}$	$\{c, aa, ab, ac\}$	$\{b\}$	Nezmenená.
3.	c	$\{\varepsilon, a\}$	$\{aa, ab, ac\}$	$\{b, c\}$	Nezmenená.
4.	aa	$\{\varepsilon, a\}$	$\{ab, ac\}$	$\{b, c, aa\}$	Nezmenená.
5.	ab	$\{\varepsilon, a\}$	$\{ac\}$	$\{b, c, aa, ab\}$	Nezmenená.
6.	ac	$\{\varepsilon, a\}$	$\emptyset$	$\{b, c, aa, ab, ac\}$	Nezmenená.

Tabuľka 3.2: Beh algoritmu 3.7.18 hľadajúceho ľavý bázičský jazyk automatu $\mathcal{B}''$. V tabuľke je pre každú iteráciu cyklu z kroku 3 uvedené slovo x vybrané v rámci tejto iterácie, ako aj všetky slová v jazykoch L , $L_?$ a $L_\times$ a redukovaná stupňovitá matica, ktorej riadky tvoria bázu priestoru generovaného množinou $\{\mathbf{i}''\mu''(w) \mid w \in L\}$, po skončení danej iterácie. Pod nultou iteráciou sa chápe stav pred vykonaním prvej iterácie cyklu.

Vidíme, že ľavým bázičským jazykom automatu $\mathcal{B}''$ je $L = \{\varepsilon, a\}$. Môžeme teda pristúpiť ku konštrukcii minimálneho automatu $\mathcal{B}'$ ekvivalentného automatu $\mathcal{B}''$ a v dôsledku toho aj automatu $\mathcal{B}$. Nech

$$X = \begin{pmatrix} \leftarrow & \mathbf{i}''\mu''(\varepsilon) & \rightarrow \\ \leftarrow & \mathbf{i}''\mu''(a) & \rightarrow \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

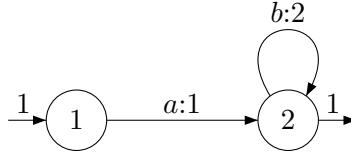
Pravá inverzná matica k matici X môže byť daná napríklad ako

$$X_r^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Automatu $\mathcal{B}'$ tak zodpovedá $\mathbb{R}$ -reprezentácia $\mathcal{P}_{\mathcal{B}'} = (2, \mathbf{i}', \mu', \mathbf{f}')$, kde $\mathbf{i}' = (1, 0)$,

$$\begin{aligned} \mu'(a) &= X\mu''(a)X_r^{-1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ \mu'(b) &= X\mu''(b)X_r^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \\ \mu'(c) &= X\mu''(c)X_r^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

a $\mathbf{f}' = X\mathbf{f}'' = (0, 1)^T$.



Obr. 3.13: Výstup algoritmu 3.9.3 – minimálny automat $\mathcal{B}'$ ekvivalentný automatu $\mathcal{B}$.

Môžeme teda uzavrieť, že minimálny automat $\mathcal{B}'$ s váhami nad poľom $\mathbb{R}$ a abecedou $\Sigma = \{a, b, c\}$ ekvivalentný automatu $\mathcal{B}$ je daný diagramom na obrázku 3.13. Je evidentné, že

$$\|\mathcal{B}'\| = \sum_{t \in \mathbb{N}} 2^t ab^t$$

a porovnanie tohto radu s radom (3.13) realizovaným automatom $\mathcal{B}$ nám tak môže poskytnúť skúšku správnosti práve vykonanej konštrukcie. Automat $\mathcal{B}'$ je tiež zrejme minimálny.

Minimálny automat s váhami nad poľom obvykle nebýva daný jednoznačne. V nasledujúcom ale ukážeme, že ľubovoľné dva minimálne automaty realizujúce ten istý rad možno jeden z druhého získať veľmi jednoduchým spôsobom, vyjadreným nasledujúcou definíciou relácie *podobnosti automatov*. Tento pojem je do veľkej miery inšpirovaný *podobnosťou matic*, s ktorou sa ešte stretneme v nasledujúcej kapitole.

Definícia 3.9.6. Nech $\mathcal{A}$ a $\mathcal{B}$ sú konečné automaty s váhami nad poľom $\mathbb{F}$ a abecedou Σ také, že $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$ a $\mathcal{P}_{\mathcal{B}} = (m, \mathbf{i}', \mu', \mathbf{f}')$. Hovoríme, že automaty $\mathcal{A}$ a $\mathcal{B}$ sú *podobné*, ak $n = m$ a existuje nesingulárna matica P taká, že $\mathbf{i}' = \mathbf{i}P$, $\mu'(c) = P^{-1}\mu(c)P$ pre všetky $c \in \Sigma$ a $\mathbf{f}' = P^{-1}\mathbf{f}$.

Tvrdenie 3.9.7. Nech $\mathcal{A}$ a $\mathcal{B}$ sú podobné konečné automaty s váhami nad poľom $\mathbb{F}$ a abecedou Σ . Potom $\|\mathcal{A}\| = \|\mathcal{B}\|$.

Dôkaz. Jednoduché cvičenie. □

Veta 3.9.8. Nech $\mathbb{F}$ je pole, Σ abeceda a $r \in \mathbb{F}\text{-Rat}(\Sigma^*)$. Ľubovoľné dva minimálne automaty $\mathcal{A}, \mathcal{B}$ nad $\mathbb{F}$ a Σ také, že $\|\mathcal{A}\| = \|\mathcal{B}\| = r$ sú podobné.

Dôkaz. Nech $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$, $\mathcal{P}_{\mathcal{B}} = (n, \mathbf{j}, \nu, \mathbf{g})$. Dokážeme najprv, že ľubovoľný ľavý bázičný jazyk $L = \{w_1, \dots, w_n\}$ automatu $\mathcal{A}$ je aj ľavým bázičným jazykom automatu $\mathcal{B}$.

Podľa vety 3.7.15 sú zobrazenia $\Lambda[\mathcal{A}]$ a $\Lambda[\mathcal{B}]$ injektívne a keďže $\dim(\mathcal{Q}(r)) = n$, je

$$(\Lambda[\mathcal{A}](\mathbf{i}\mu(w_1)), \dots, \Lambda[\mathcal{A}](\mathbf{i}\mu(w_n))) = (w_1^{-1}r, \dots, w_n^{-1}r)$$

bázou priestoru $\mathcal{Q}(r)$. Súčasne však

$$(w_1^{-1}r, \dots, w_n^{-1}r) = (\Lambda[\mathcal{B}](\mathbf{j}\nu(w_1)), \dots, \Lambda[\mathcal{B}](\mathbf{j}\nu(w_n))).$$

Keďže je podľa vety 3.7.15 dimenzia priestoru $\text{Left}(\mathcal{B})$ rovná n a keďže je $(w_1^{-1}r, \dots, w_n^{-1}r)$ báza priestoru dimenzie n , musí byť $(\mathbf{j}\nu(w_1), \dots, \mathbf{j}\nu(w_n))$ bázou priestoru $\text{Left}(\mathcal{B})$ a jazyk L tak musí byť ľavým bázičným jazykom aj pre automat $\mathcal{B}$.

Aplikujme teraz na obidva automaty $\mathcal{A}$ a $\mathcal{B}$ konštrukciu z lemy 3.9.1 pre rovnaký ľavý bázičný jazyk $L = \{w_1, \dots, w_n\}$. Označme výsledné automaty ako $\mathcal{A}'$ resp. $\mathcal{B}'$. Ukážeme, že $\mathcal{A}' = \mathcal{B}'$. Na to stačí ukázať, že sa rovnajú príslušné lineárne reprezentácie $\mathcal{P}_{\mathcal{A}'} = (n, \mathbf{i}', \mu', \mathbf{f}')$ a $\mathcal{P}_{\mathcal{B}'} = (n, \mathbf{j}', \nu', \mathbf{g}')$.

Podľa lemy 3.9.1 je $\mathbf{i}' = \mathbf{j}' = (1, 0, \dots, 0)$. Vektory $\mathbf{i}'\mu'(w_1), \dots, \mathbf{i}'\mu'(w_n)$ ďalej tvoria, v tomto poradí, štandardnú bázu priestoru $\mathbb{F}^n$ a rovnakú vlastnosť majú aj vektory $\mathbf{j}'\nu'(w_1), \dots, \mathbf{j}'\nu'(w_n)$. Ak totiž označíme maticu X z lemy 3.9.1 aplikovanej na automat $\mathcal{A}$ ako $X_{\mathcal{A}}$ a rovnakú maticu pre automat $\mathcal{B}$ ako $X_{\mathcal{B}}$, pre $i = 1, \dots, n$ je $\mathbf{i}'\mu'(w_i)X_{\mathcal{A}} = \mathbf{i}\mu(w_i)$ a $\mathbf{j}'\nu'(w_i)X_{\mathcal{B}} = \mathbf{j}\nu(w_i)$. Naozaj teda

$$(\mathbf{i}'\mu'(w_1), \dots, \mathbf{i}'\mu'(w_n)) = (\mathbf{e}_1, \dots, \mathbf{e}_n) = (\mathbf{j}'\nu'(w_1), \dots, \mathbf{j}'\nu'(w_n)).$$

Ak ďalej pre $w \in \Sigma^*$ je $\mathbf{i}'\mu'(w) = (a_1, \dots, a_n) \in \mathbb{F}^n$ a $\mathbf{j}'\nu'(w) = (b_1, \dots, b_n) \in \mathbb{F}^n$, nutne

$$\Lambda[\mathcal{A}'](a_1, \dots, a_n) = \Lambda[\mathcal{A}'](\mathbf{i}'\mu'(w)) = w^{-1}r$$

a súčasne

$$\Lambda[\mathcal{A}'](a_1, \dots, a_n) = a_1\Lambda[\mathcal{A}'](\mathbf{e}_1) + \dots + a_n\Lambda[\mathcal{A}'](\mathbf{e}_n) = a_1w_1^{-1}r + \dots + a_nw_n^{-1}r,$$

z čoho

$$w^{-1}r = a_1w_1^{-1}r + \dots + a_nw_n^{-1}r. \quad (3.14)$$

Prostredníctvom zobrazenia $\Lambda[\mathcal{B}']$ rovnakým spôsobom dostávame

$$w^{-1}r = b_1w_1^{-1}r + \dots + b_nw_n^{-1}r$$

a keďže je $(w_1^{-1}r, \dots, w_n^{-1}r)$ bázou priestoru $\mathcal{Q}(r)$, nutne

$$\mathbf{i}'\mu'(w) = (a_1, \dots, a_n) = (b_1, \dots, b_n) = \mathbf{j}'\nu'(w)$$

pre všetky $w \in \Sigma^*$. Pre všetky $c \in \Sigma$ a $i = 1, \dots, n$ teda špeciálne

$$\mathbf{e}_i\mu'(c) = \mathbf{i}'\mu'(w_i)\mu'(c) = \mathbf{i}'\mu'(w_ic) = \mathbf{j}'\nu'(w_ic) = \mathbf{j}'\nu'(w_i)\nu'(c) = \mathbf{e}_i\nu'(c),$$

z čoho vyplýva, že $\mu'(c)$ a $\nu'(c)$ sú maticami toho istého lineárneho zobrazenia, a teda $\mu'(c) = \nu'(c)$. Pre $i = 1, \dots, n$ napokon

$$\mathbf{e}_i\mathbf{f}' = \mathbf{i}'\mu'(w_i)\mathbf{f}' = (r, w_i) = \mathbf{j}'\nu'(w_i)\mathbf{g}' = \mathbf{e}_i\mathbf{g}',$$

čiže $\mathbf{f}' = \mathbf{g}'$. Naozaj teda $\mathcal{A}' = \mathcal{B}'$, z čoho

$$\mathbf{i}'X_{\mathcal{A}} = \mathbf{i}, \quad \mu'(c) = X_{\mathcal{A}}\mu(c)X_{\mathcal{A}}^{-1} \text{ pre všetky } c \in \Sigma \quad \text{a} \quad \mathbf{f}' = X_{\mathcal{A}}\mathbf{f},$$

ako aj

$$\mathbf{i}'X_{\mathcal{B}} = \mathbf{j}, \quad \mu'(c) = X_{\mathcal{B}}\nu(c)X_{\mathcal{B}}^{-1} \text{ pre všetky } c \in \Sigma \quad \text{a} \quad \mathbf{f}' = X_{\mathcal{B}}\mathbf{g}.$$

Preto

$$\mathbf{j} = \mathbf{i}(X_{\mathcal{A}}^{-1}X_{\mathcal{B}}), \quad \nu(c) = (X_{\mathcal{A}}^{-1}X_{\mathcal{B}})^{-1}\mu(c)(X_{\mathcal{A}}^{-1}X_{\mathcal{B}}) \text{ pre všetky } c \in \Sigma \quad \text{a} \quad \mathbf{g} = (X_{\mathcal{A}}^{-1}X_{\mathcal{B}})^{-1}\mathbf{f},$$

čo znamená, že automaty $\mathcal{A}$ a $\mathcal{B}$ sú podobné. □

Poznámka 3.9.9. Celú teóriu okolo minimalizácie konečných automatov s váhami nad poľom možno pri troche opatrnosti prerobiť tak, aby ostala platná aj pre konečné automaty s váhami nad telesom, čiže „poľom, ktoré nemusí byť komutatívne“. Vyžadovalo by si to však používať výsledky z nekomutatívnej lineárnej algebry, ktoré síce sú priamočiarym rozšírením ich náprotivkov z klasického komutatívneho rámca, no ťažko ich považovať za všeobecne známe. Čitateľa tu odkážeme na J. Sakarovitcha [94, 95].

3.10 Racionálne rady nad poľom a unárnou abecedou

Uvažujme teraz formálne mocninové rady nad poľom $\mathbb{F}$ a jednoprvkovou abecedou $\Sigma = \{z\}$, ktoré možno – ako sme už pozorovali – považovať za prvky hociktorého z izomorfných oborov integrity $\mathbb{F}\langle\langle z^* \rangle\rangle$ a $\mathbb{F}[[z]]$. V rámci príkladu 3.5.15 sme navyše dokázali, že takýto rad r je racionálny nad $\mathbb{F}$ – čiže $r \in \mathbb{F}\text{-Rat}(z^*)$ – práve vtedy, keď ho možno vyjadriť ako podiel p/q dvoch polynómov $p, q \in \mathbb{F}[z]$, kde q má nenulový konštantný koeficient.

Predpokladajme, že na konečný automat s váhami $\mathcal{B}$ nad $\mathbb{F}$ a $\Sigma = \{z\}$ realizujúci rad $r \in \mathbb{F}\langle\langle z^* \rangle\rangle$ aplikujeme minimalizačný algoritmus 3.9.3, na výstupe ktorého dostaneme minimálny automat $\mathcal{B}'$ pre rad r taký, že $\mathcal{P}_{\mathcal{B}'} = (m, \mathbf{i}', \mu', \mathbf{f}')$. Ak $r \neq 0$ – čiže ak $m > 0$ – musel byť v kroku 7 algoritmu 3.9.3 nájdený ako ľavý bázičský jazyk automatu $\mathcal{B}''$ s $\mathcal{P}_{\mathcal{B}''} = (k, \mathbf{i}'', \mu'', \mathbf{f}'')$ jazyk $L = \{z^0, z^1, \dots, z^{m-1}\}$. Videli sme, že

$$\mathbf{i}' = (1, 0, \dots, 0)$$

a pre všetky $w \in z^*$ musí byť vektor $\mathbf{i}'\mu'(w)$ vyjadrením vektora $\mathbf{i}''\mu''(w)$ v súradniciach bázy³⁴

$$(\mathbf{i}''\mu''(z^0), \mathbf{i}''\mu''(z^1), \dots, \mathbf{i}''\mu''(z^{m-1}))$$

priestoru $\text{Left}(\mathcal{B}'')$, z čoho špeciálne vyplýva, že

$$(\mathbf{i}'\mu'(z^0), \mathbf{i}'\mu'(z^1), \dots, \mathbf{i}'\mu'(z^{m-1}))$$

je štandardná báza $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ priestoru $\mathbb{F}^m$ a pre $j = 1, \dots, m-1$ je

$$\mathbf{e}_j\mu'(z) = \mathbf{i}'\mu'(z^{j-1})\mu'(z) = \mathbf{i}'\mu'(z^j) = \mathbf{e}_{j+1}.$$

Nutne teda

$$\mu'(z) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ c_0 & c_1 & c_2 & \dots & c_{m-1} \end{pmatrix}$$

pre nejaké $c_0, \dots, c_{m-1} \in \mathbb{F}$.³⁵ Vektor koncových váh $\mathbf{f}'$ je potom evidentne daný ako

$$\mathbf{f}' = ((r, z^0), (r, z^1), \dots, (r, z^{m-1}))^T.$$

Realizácia ľubovoľného radu $r \in \mathbb{F}\text{-Rat}(z^*) \setminus \{0\}$ konečným automatom s váhami práve opísaného špeciálneho typu vedie k charakterizácii radov racionálnych nad $\mathbb{F}$ a $\Sigma = \{z\}$ pomocou *lineárnych rekurencií*, ktorú teraz dokážeme.

Veta 3.10.1. *Nech $\mathbb{F}$ je pole a $r \in \mathbb{F}\langle\langle z^* \rangle\rangle$ – resp. $r \in \mathbb{F}[[z]]$ – je formálny mocninový rad. Potom sú nasledujúce tvrdenia ekvivalentné:*

- (i) Rad r je racionálny nad $\mathbb{F}$.
- (ii) Rad r možno vyjadriť ako podiel $r = p/q$, kde $p, q \in \mathbb{F}[z]$ sú polynómy a q má nenulový konštantný koeficient.
- (iii) Koeficienty radu r pre všetky $t \in \mathbb{N}$ vyhovujú rekurentnému vzťahu

$$(r, z^{t+m}) = c_{m-1}(r, z^{t+m-1}) + c_{m-2}(r, z^{t+m-2}) + \dots + c_0(r, z^t),$$

kde $m \in \mathbb{N} \setminus \{0\}$ a $c_0, \dots, c_{m-1} \in \mathbb{F}$.

Dôkaz. Ekvivalenciu tvrdení (i) a (ii) sme už dokázali v rámci príkladu 3.5.15. Dokážeme ekvivalenciu tvrdení (i) a (iii).

³⁴Za predpokladu, že slová ľavého bázičského jazyka $L = \{z^0, \dots, z^{m-1}\}$ uvažujeme v poradí, v akom sú uvedené – čiže $w_1 = z^0, \dots, w_m = z^{m-1}$.

³⁵Matica $\mu'(z)$ je takzvanou *sprievodnou maticou* polynómu $z^m - c_{m-1}z^{m-1} - \dots - c_0z^0$ – čo je dôležitý pojem, s ktorým sa ešte neskôr stretneme. Pre účely našich terajších úvah ale táto skutočnosť nie je zásadná.

Nech $r \in \mathbb{F}\langle\langle z^* \rangle\rangle$. Ak $r = 0$, sú evidentne pravdivé obidve tvrdenia (i) a (iii). Predpokladajme teda, že $r \neq 0$. Podľa úvah predchádzajúcich tejto vete je potom racionálnosť radu r nad $\mathbb{F}$ ekvivalentná existencii lineárnej $\mathbb{F}$ -reprezentácie $\mathcal{P} = (m, \mathbf{i}, \mu, \mathbf{f})$ nad $\Sigma = \{z\}$ takej, že $\|\mathcal{P}\| = r$, $\mathbf{i} = (1, 0, \dots, 0)$,

$$\mu(z) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ c_0 & c_1 & c_2 & \dots & c_{m-1} \end{pmatrix}$$

pre nejaké $c_0, \dots, c_{m-1} \in \mathbb{F}$ a

$$\mathbf{f} = \begin{pmatrix} (r, z^0) \\ (r, z^1) \\ \vdots \\ (r, z^{m-1}) \end{pmatrix}.$$

Takáto $\mathbb{F}$ -reprezentácia $\mathcal{P}$ musí pre všetky $t \in \mathbb{N}$ spĺňať

$$\mu(z^t)\mathbf{f} = \begin{pmatrix} (r, z^t) \\ (r, z^{t+1}) \\ \vdots \\ (r, z^{t+m-1}) \end{pmatrix};$$

pre $j = 1, \dots, m$ a j -ty vektor $\mathbf{e}_j$ štandardnej bázy priestoru $\mathbb{F}^m$ totiž

$$\mathbf{e}_j \mu(z^t)\mathbf{f} = \mathbf{i} \mu(z^{j-1}) \mu(z^t)\mathbf{f} = \mathbf{i} \mu(z^{t+j-1})\mathbf{f} = (r, z^{t+j-1}).$$

Pre všetky $t \in \mathbb{N}$ tak v dôsledku toho dostávame aj

$$\mu(z) \begin{pmatrix} (r, z^t) \\ (r, z^{t+1}) \\ \vdots \\ (r, z^{t+m-2}) \\ (r, z^{t+m-1}) \end{pmatrix} = \begin{pmatrix} (r, z^{t+1}) \\ (r, z^{t+2}) \\ \vdots \\ (r, z^{t+m-1}) \\ (r, z^{t+m}) \end{pmatrix}.$$

Reprezentácia $\mathcal{P}$ ale súčasne pre všetky $t \in \mathbb{N}$ spĺňa

$$\mu(z) \begin{pmatrix} (r, z^t) \\ (r, z^{t+1}) \\ \vdots \\ (r, z^{t+m-2}) \\ (r, z^{t+m-1}) \end{pmatrix} = \begin{pmatrix} (r, z^{t+1}) \\ (r, z^{t+2}) \\ \vdots \\ (r, z^{t+m-1}) \\ c_{m-1}(r, z^{t+m-1}) + c_{m-2}(r, z^{t+m-2}) + \dots + c_0(r, z^t) \end{pmatrix}.$$

Realizácia radu r reprezentáciou $\mathcal{P}$ je teda ekvivalentná platnosti rekurentného vzťahu

$$(r, z^{t+m}) = c_{m-1}(r, z^{t+m-1}) + c_{m-2}(r, z^{t+m-2}) + \dots + c_0(r, z^t),$$

pre všetky $t \in \mathbb{N}$ a koeficienty $(r, z^0), \dots, (r, z^{m-1})$ dané pevne ako počiatočné podmienky, čo dokazuje aj ekvivalenciu tvrdení (i) a (iii). $\square$

Formálny mocninový rad o jednej premennej z nad poľom $\mathbb{F}$ je teda racionálny práve vtedy, keď jeho koeficienty vyhovujú *lineárnej homogénnej rekurencii s konštantnými koeficientmi* – čiže rekurencii rovnakého druhu ako v tvrdení (iii) predchádzajúcej vety. Túto skutočnosť možno ľahko dokázať aj bez použitia automatov s váhami, vychádzajúc z vyjadriteľnosti racionálnych radov v podobe podielu dvoch vhodných polynómov. Takýto priamy dôkaz ekvivalencie tvrdení (ii) a (iii) z predchádzajúcej vety možno nájsť napríklad v [70].

3.11 Rozhodovacie problémy pre automaty s váhami nad poľom

Preskúmame teraz v stručnosti dva dôležité *rozhodovacie problémy* pre automaty s váhami nad poľom. Rozhodnuteľnosť *problému ekvivalencie* pre takéto automaty je jednoduchým dôsledkom existencie minimalizačného algoritmu z oddielu 3.9. Z tohto pozorovania vyplýva aj rozhodnuteľnosť ekvivalencie konečných automatov s váhami nad obormi integrity – každý obor integrity totiž možno vnoriť do jeho podielového poľa a konečný automat s váhami nad oborom integrity je tak vždy súčasne aj konečným automatom s váhami nad poľom.

Veta 3.11.1. *Problém ekvivalencie konečných automatov s váhami nad poľom – alebo oborom integrity – je rozhodnuteľný.*

Dôkaz. Nech $\mathcal{A}_1, \mathcal{A}_2$ sú konečné automaty s váhami nad poľom $\mathbb{F}$, o ktorých treba rozhodnúť, či $\|\mathcal{A}_1\| = \|\mathcal{A}_2\|$. K automatu $\mathcal{A}_2$ možno prenásobením počiatočných váh prvkom -1 poľa $\mathbb{F}$ získať automat $\mathcal{A}'_2$ taký, že $\|\mathcal{A}'_2\| = -\|\mathcal{A}_2\|$. Štandardnou konštrukciou z lemy 3.5.8 následne možno zostrojiť automat $\mathcal{A}$ taký, že

$$\|\mathcal{A}\| = \|\mathcal{A}_1\| + \|\mathcal{A}'_2\| = \|\mathcal{A}_1\| - \|\mathcal{A}_2\|.$$

Na rozhodnutie ekvivalencie automatov $\mathcal{A}_1$ a $\mathcal{A}_2$ teda stačí rozhodnúť, či $\|\mathcal{A}\| = 0$. To ale možno urobiť minimalizáciou automatu $\mathcal{A}$ pomocou algoritmu 3.9.3: rovnosť $\|\mathcal{A}\| = 0$ nastáva práve vtedy, keď minimálny automat ekvivalentný automatu $\mathcal{A}$ nemá žiaden stav. $\square$

Poznámka 3.11.2. Význam vety 3.11.1 možno lepšie oceniť vo svetle skutočnosti, že ekvivalencia konečných automatov s váhami *nie je* rozhodnuteľná nad každým polokruhom. Hoci dôkaz presahuje rámec tohto textu, spomeňme tu aspoň výsledok D. Kroba [73], ktorý dokázal nerozhodnuteľnosť ekvivalencie pre konečné automaty s váhami nad tropickým polokruhom $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$.

Zďaleka nie všetky prirodzené rozhodovacie problémy sú pre automaty s váhami nad poľami rozhodnuteľné. Dokážeme, že problém *univerzality nosiča* realizovaného radu – čiže problém neexistencie slova, ktoré má v realizovanom rade nulový koeficient – je nerozhodnuteľný už pre konečné automaty s celočíselnými váhami; tým pádom musí byť nerozhodnuteľný aj pre všetky polia obsahujúce $\mathbb{Z}$ ako podokruh, t. j. pre všetky polia nulovej charakteristiky (ako napríklad $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, atď.). Na dôkaz nerozhodnuteľnosti tohto problému sa nám zide pomocné tvrdenie o uzavretosti tried racionálnych radov na inverzné homomorfizmy, ktoré platí nad ľubovoľným polokruhom S a ktoré teraz dokážeme.

Lema 3.11.3. *Nech S je polokruh, Σ, Γ abecedy, $h: \Sigma^* \rightarrow \Gamma^*$ homomorfizmus a $r \in S\langle\Gamma^*\rangle$ je rad racionálny nad S . Potom je nad S racionálny aj rad*

$$h^{-1}(r) = \sum_{w \in \Sigma^*} (r, h(w)) w \in S\langle\Sigma^*\rangle.$$

Dôkaz. Nech $\mathcal{P} = (n, \mathbf{i}, \mu, \mathbf{f})$ je lineárna S -reprezentácia nad Γ taká, že $\|\mathcal{P}\| = r$. Uvažujme lineárnu S -reprezentáciu $\mathcal{P}' = (n, \mathbf{i}, \mu', \mathbf{f})$ nad Σ takú, že pre všetky $c \in \Sigma$ je

$$\mu'(c) = \mu(h(c)).$$

Nech $w \in \Sigma^*$, pričom $w = a_1 \dots a_t$ pre nejaké $t \in \mathbb{N}$ a $a_1, \dots, a_t \in \Sigma$. Potom

$$\begin{aligned} (\|\mathcal{P}'\|, w) &= \mathbf{i}\mu'(a_1 \dots a_t)\mathbf{f} = \mathbf{i}\mu'(a_1) \dots \mu'(a_t)\mathbf{f} = \mathbf{i}\mu(h(a_1)) \dots \mu(h(a_t))\mathbf{f} = \\ &= \mathbf{i}\mu(h(a_1) \dots h(a_t))\mathbf{f} = \mathbf{i}\mu(h(a_1 \dots a_t))\mathbf{f} = (\|\mathcal{P}\|, h(w)) = (r, h(w)), \end{aligned}$$

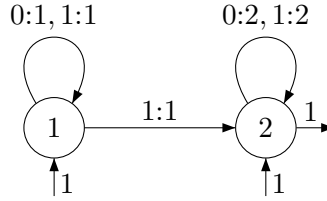
takže $\|\mathcal{P}'\| = h^{-1}(r)$ a $h^{-1}(r) \in S\text{-Rat}(\Sigma^*)$. $\square$

Veta 3.11.4. *Problém univerzality nosiča realizovaného radu je pre konečné automaty s váhami nad $\mathbb{Z}$ nerozhodnuteľný.*

Dôkaz. Uvažujme abecedu $\Sigma = \{0, 1\}$ a rad $\text{BIN} \in \mathbb{Z}\langle\langle\Sigma^*\rangle\rangle$ taký, že pre všetky $t \in \mathbb{N}$ a $a_1, \dots, a_t \in \Sigma$ je

$$(\text{BIN}, a_1 \dots a_t) = 2^t + \sum_{k=1}^t a_k 2^{t-k}.$$

Koeficientom ľubovoľného slova $w \in \Sigma^*$ v rade BIN je teda kladné celé číslo reprezentované binárnym reťazcom $1w$. Rad BIN je navyše evidentne racionálny nad $\mathbb{Z}$, keďže je realizovaný konečným automatom s váhami $\mathcal{A}$ na obrázku 3.14.



Obr. 3.14: Konečný automat $\mathcal{A}$ s váhami nad $\mathbb{Z}$ realizujúci rad BIN .

Nerozhodnuteľnosť problému univerzality nosiča radu realizovaného konečným automatom s váhami nad $\mathbb{Z}$ teraz dokážeme redukciou *Postovho korešpondenčného problému* nad binárnou abecedou na tento problém. Nech (Δ, f, g) je ľubovoľný prípad Postovho korešpondenčného problému nad abecedou $\Sigma = \{0, 1\}$ – čiže Δ je abeceda a $f, g: \Delta^* \rightarrow \Sigma^*$ sú homomorfizmy; riešením prípadu (Δ, f, g) je slovo $w \in \Delta^+$ také, že $f(w) = g(w)$. Vďaka leme 3.11.3, leme 3.5.7, leme 3.5.8, leme 3.5.9 a existencii automatu $\mathcal{A}$ realizujúceho rad BIN potom na základe (Δ, f, g) ľahko algoritmicke skonstruujeme konečný automat $\mathcal{B}$ s váhami nad $\mathbb{Z}$ a Δ taký, že

$$\|\mathcal{B}\| = f^{-1}(\text{BIN}) - g^{-1}(\text{BIN}) + 1;$$

rad $\|\mathcal{B}\|$ teda spĺňa $(\|\mathcal{B}\|, \varepsilon) = 1$ a $(\|\mathcal{B}\|, w) = (\text{BIN}, f(w)) - (\text{BIN}, g(w))$ pre všetky $w \in \Delta^+$. Keďže pre $u, v \in \Sigma^*$ je $(\text{BIN}, u) = (\text{BIN}, v)$ práve vtedy, keď $u = v$, je nosič radu $\|\mathcal{B}\|$ rovný Δ^* práve vtedy, keď neexistuje žiadne $w \in \Delta^+$ také, že $f(w) = g(w)$ – čiže práve vtedy, keď prípad Postovho korešpondenčného problému (Δ, f, g) nemá riešenie. Rozhodnuteľnosť problému univerzality nosiča radu realizovaného automatom s váhami nad $\mathbb{Z}$ by teda implikovala rozhodnuteľnosť Postovho korešpondenčného problému nad binárnou abecedou: spor. $\square$

Možno dokázať, že problém univerzality nosiča radu realizovaného konečným automatom s váhami nad poľom $\mathbb{F}$ je rozhodnuteľný práve vtedy, keď je pole $\mathbb{F}$ *lokálne konečné* – čiže každé jeho konečne generované podpole je konečné; to je ďalej ekvivalentné podmienke, že $\mathbb{F}$ je algebraickým rozšírením konečného poľa. Dôkaz tohto tvrdenia možno nájsť v [69].

Poznámka 3.11.5. Dosiaľ nie je známe, či sa problém univerzality nosiča pre automaty s váhami nad $\mathbb{Z}$ stane rozhodnuteľným v prípade, že sa obmedzíme na automaty nad unárnou abecedou. Vďaka našim pozorovaniam z oddielu 3.10 je tento rozhodovací problém ekvivalentný – vzhľadom na rozhodnuteľnosť – nasledujúcemu problému: na vstupe je daná lineárna homogénna rekurencia s celočíselnými koeficientmi a danými začiatočnými podmienkami a treba rozhodnúť, či príslušná postupnosť obsahuje aspoň jeden nulový člen.³⁶ Tento problém je v literatúre známy ako *Skolemov problém* a otázka ohľadom jeho rozhodnuteľnosti patrí k dôležitým otvoreným problémom súčasnej matematiky.

³⁶Z našich zistení z oddielu 3.10 vyplýva, že postupnosť definovaná takouto rekurenciou je súčasne aj postupnosťou koeficientov radu realizovaného automatom s váhami nad $\mathbb{Z}$ a $\Sigma = \{z\}$. Naopak tiež vieme, že postupnosť koeficientov $(a_n)_{n=0}^\infty$ každého radu $r \in \mathbb{Z}\langle\langle z^* \rangle\rangle$ racionálneho nad $\mathbb{Z}$ vyhovuje nejakej lineárnej homogénnej rekurencii s koeficientmi z $\mathbb{Q}$; pre vhodné $q \in \mathbb{N} \setminus \{0\}$ potom musí postupnosť $(q^n a_n)_{n=0}^\infty$ vyhovovať lineárnej homogénnej rekurencii s koeficientmi zo $\mathbb{Z}$, pričom pre všetky $n \in \mathbb{N}$ je $a_n = 0$ práve vtedy, keď $q^n a_n = 0$. V skutočnosti ale možno dokázať, že už samotná postupnosť $(a_n)_{n=0}^\infty$ musí vyhovovať lineárnej homogénnej rekurencii s celočíselnými koeficientmi: automat získaný minimalizáciou unárneho automatu s váhami nad $\mathbb{Z}$ je opäť automatom s váhami nad $\mathbb{Z}$ [11].

3.12 Využitie automatov s váhami na kódovanie obrazu

Konečné automaty s váhami nachádzajú – podobne ako ich klasické nekvantitatívne náprotivky – praktické uplatnenie v mnohých oblastiach aplikovanej informatiky. Obzvlášť významné sú najmä ich aplikácie pri spracovaní prirodzeného jazyka [65]. V nasledujúcom si na ukážku priblížime možnosti využitia automatov s váhami v inej oblasti: pri kódovaní obrazu. Táto aplikácia síce nepatrí k hlavným z hľadiska významu – v praxi sa obyčajne používajú iné prístupy ku kódovaniu obrazu – avšak jej prezentácia nevyžaduje takmer žiaden prípravný materiál z oblasti aplikácie a budeme ju tak môcť v rámci nasledujúcej krátkej demonštrácie preskúmať aj bez nutnosti ďalekého prekročenia hraníc oblastí, ktorými sa v tomto texte máme v úmysle zaoberať.

Obmedzíme sa tu iba na najzákladnejšie myšlienky stojace za touto metódou kódovania obrazu – podrobnejšie spracovanie problematiky možno nájsť napríklad v [3, 61].

Automaty s váhami pri tejto metóde nezohrávajú úlohu kódovacieho algoritmu, ale samotného „kódu“. Celý prístup teda bude založený na *reprezentácii* obrázkov pomocou automatov s váhami – presnejšie automatov s váhami *nad poľom*. Bitmapy stotožníme s vhodnými racionálnymi mocninovými radmi; reprezentáciou obrázku tak bude ľubovoľný automat s váhami realizujúci príslušný rad. Opíšeme algoritmus, ktorý pre daný automat vykreslí ním reprezentovanú bitmapu. Kompresia bitmapy bude naopak spočívať v tom, že sa za jej reprezentáciu vezme *minimálny* automat pre príslušný rad. Pri jeho hľadaní budeme využívať teóriu z oddielu 3.7.

Pre jednoduchosť predpokladajme, že máme danú bitmapu o $2^k \times 2^k$ pixeloch pre nejaké prirodzené číslo k . Každý pixel možno pri šedotónových obrázkoch reprezentovať prvkom poľa $\mathbb{R}$ udávajúcim jeho intenzitu – môže ísť napríklad o číslo z intervalu $[0, 1]$, kde čiernej farbe zodpovedá hodnota 0 a bielej zodpovedá hodnota 1. Pri farebných obrázkoch sú pixely prvkami množiny $\mathbb{R}^3$ – celý obrázok potom možno chápať ako „zloženie“ troch obrázkov s pixelmi z $\mathbb{R}$; v nasledujúcom sa preto bez ujmy na všeobecnosti obmedzíme na šedotónové obrázky³⁷ s pixelmi z poľa $\mathbb{R}$.

1	2
3	4

(a) Kvadranty najvyššej úrovne.

11	12	21	22
13	14	23	24
31	32	41	42
33	34	43	44

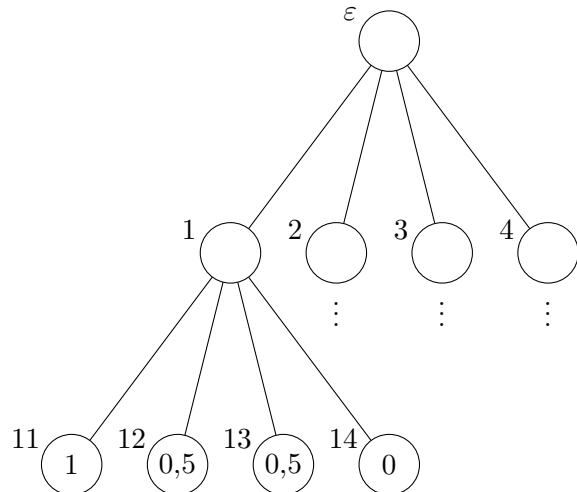
(b) Kvadranty druhej najvyššej úrovne.

Obr. 3.15: Rozdelenie bitmapy na kvadranty.

Bitmapu takéhoto typu možno reprezentovať pomocou jej *kvadrantového stromu* (angl. *quadtree*). Ide o zakorenený strom, v ktorom má každý vnútorný uzol práve štyroch synov (ktorých usporiadanie je dané). Koreň kvadrantového stromu zodpovedá celej bitmape o rozmeroch $2^k \times 2^k$. Tú možno rozdeliť na kvadranty 1, 2, 3, 4 o rozmeroch $2^{k-1} \times 2^{k-1}$ – napríklad tak, ako je to znázornené na obrázku 3.15a. Synovia koreňa zodpovedajú týmto štyrom kvadrantom. Tento proces možno následne zopakovať: každý kvadrant $c \in \{1, 2, 3, 4\}$ možno rozdeliť na štyri kvadranty $c1, c2, c3, c4$ o rozmeroch $2^{k-2} \times 2^{k-2}$ tak, ako na obrázku 3.15b; uzol kvadrantového stromu zodpovedajúci kvadrantu c pritom bude mať opäť štyroch synov zodpovedajúcich kvadrantom $c1, c2, c3, c4$. Takto možno pokračovať až ku kvadrantom o rozmeroch 1×1 pixel; výsledný kvadrantový strom pre bitmapu s rozmermi $2^k \times 2^k$ má teda hĺbku presne k a každý kvadrant o rozmeroch $2^{k-j} \times 2^{k-j}$ (kde $j \leq k$ je prirodzené číslo) je určený svojou „adresou“, ktorou je slovo z jazyka $\{1, 2, 3, 4\}^j$. Listy kvadrantového stromu zodpovedajú samotným pixelom a možno v nich uchovávať informáciu o ich intenzite alebo farbe. Príklad bitmapy o 4×4 pixeloch a časti zodpovedajúceho kvadrantového stromu je znázornený na obrázku 3.16.

³⁷Alebo ekvivalentne obrázky v škálach intenzity červenej, zelenej, alebo modrej.

1	0,5	0,5	0,5
0,5	0	0,5	0,5
1	0	0	1
0	1	1	0

(a) Bitmapa o 4×4 pixeloch.

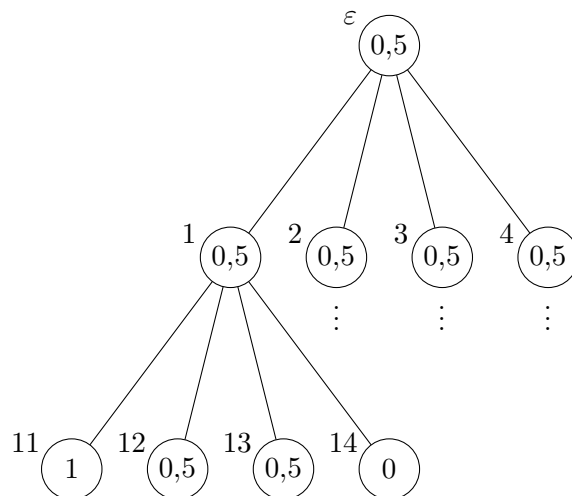
(b) Časť príslušného kvadrantového stromu.

Obr. 3.16: Bitmapa a časť jej kvadrantového stromu.

Prirodzene sa ponúka využitie vnútorných uzlov kvadrantového stromu na uchovávanie informácií o „zmenšených“ verziách pôvodnej bitmapy. Napríklad pri vykresľovaní bitmapy o $2^k \times 2^k$ pixeloch na plátno o $2^{k-1} \times 2^{k-1}$ pixeloch možno uzly kvadrantového stromu zodpovedajúce kvadrantom *druhej najnižšej* úrovne – tie majú v pôvodnej bitmape 2×2 pixely – interpretovať ako uzly prislúchajúce k jednotlivým pixelom „zmenšenej“ verzie bitmapy a pamätať si v nich intenzitu týchto pixelov. Bez ďalších obmedzení môžu byť jednotlivé „zmenšené“ verzie úplne nezávislé na pôvodnej bitmape – prirodzená sa teda javí byť požiadavka, aby intenzita C_w vnútorného uzla s „adresou“ w bola daná aritmetickým priemerom intenzít jeho synov:

$$C_w = \frac{C_{w1} + C_{w2} + C_{w3} + C_{w4}}{4};$$

budeme potom hovoriť, že kvadrantový strom je *priemer zachovávajúci*. Na obrázku 3.17 je znázornená tá istá časť kvadrantového stromu ako na obrázku 3.16b; sú doň ale doplnené intenzity jeho vnútorných uzlov tak, aby bol výsledný strom priemer zachovávajúci.

**Obr. 3.17:** Doplnenie intenzít vnútorných uzlov do časti kvadrantového stromu z obrázku 3.16b.

Aj samotné kvadrantové stromy sa dajú použiť na kompresiu obrazu: ak totiž majú všetky uzly v niektorom podstrome rovnakú intenzitu, možno celý tento podstrom reprezentovať iba jeho koreňom. Ide pritom o jeden z najbežnejších spôsobov použitia kvadrantových stromov v počítačovej grafike. My toto vylepšenie používať nebudeme, keďže kvadrantové stromy pre nás budú iba pomocným nástrojom. Ako komprimované reprezentácie bitmáp budeme používať automaty s váhami a s kvadrantovými stromami budeme pracovať len ako s nástrojom na ich konštrukciu.

V nasledujúcom budeme pracovať s „nekonečne ostrými“ bitmapami, ktoré sú dané v rozlíšení $2^k \times 2^k$ pre všetky prirodzené čísla k . Príslušnou reprezentáciou tu je *nekonečný kvadrantový strom*, pričom pixely bitmapy vo verzii o rozlíšení $2^k \times 2^k$ sú dané hodnotami v uzloch hĺbky k . Pojem *priemer zachovávajúceho* kvadrantového stromu možno prirodzeným spôsobom rozšíriť aj na nekonečné kvadrantové stromy.

Poznámka 3.12.1. Samozrejme tu ide o idealizáciu, keďže pri reálnej implementácii ťažko očakávať bitmapu danú v nekonečnom počte rozlíšení. Ak je však obrázok daný v rozlíšení 2^k pre nejaké *fixné* $k \in \mathbb{N}$, možno si zapamätať prvých $k + 1$ úrovní nekonečného kvadrantového stromu a jeho nižšie úrovne reprezentovať uzlom s označením DNC (z angl. *do not care*). Za uzol DNC možno dosadiť akúkoľvek hodnotu podľa toho, aká je akurát pre dané účely najvýhodnejšia. Algoritmy na kódovanie a dekódovanie obrazu teda môžeme opísať používajúc nekonečné kvadrantové stromy; prechod k reálnej implementácii akurát znamená nahradiť ich konečnými stromami s uzlom DNC.

Metóda reprezentácie „nekonečne ostrých“ bitmáp prostredníctvom automatov s váhami je založená na pozorovaní, že *každý nekonečný kvadrantový strom jednoznačne udáva formálny mocninový rad a naopak*. Pri tejto bijekcii sa k ľubovoľnému nekonečnému kvadrantovému stromu $\mathcal{T}$ priradí formálny mocninový rad r nad abecedou $\Sigma := \{1, 2, 3, 4\}$ a poľom $\mathbb{R}$ taký, že pre všetky $w \in \Sigma^*$ je hodnota (r, w) daná hodnotou uzla stromu $\mathcal{T}$ s „adresou“ w . Toto zobrazenie budeme mať na mysli vždy, keď budeme hovoriť o *formálnom mocninovom rade danom nekonečným kvadrantovým stromom*. Formálny mocninový rad r nazveme *priemer zachovávajúcim*, ak je priemer zachovávajúci nekonečný kvadrantový strom, ktorým je rad r daný.

Je navyše zrejmé, že pre nekonečný kvadrantový strom $\mathcal{T}$, ním daný formálny mocninový rad r a každé $c \in \{1, 2, 3, 4\}$ je rad $c^{-1}r$ daný podstromom stromu $\mathcal{T}$ zakoreneným v uzle s „adresou“ c – to jest v c -tom synovi koreňa stromu $\mathcal{T}$. Operácií ľavého kvocientu teda zodpovedá posun smerom nadol v kvadrantovom strome, čo je intuitívne to isté ako priblíženie sa do daného kvadrantu.

Na vykreslenie obrázku reprezentovaného konečným automatom $\mathcal{A}$ s váhami nad poľom $\mathbb{R}$ a abecedou $\{1, 2, 3, 4\}$ v rozlíšení $2^k \times 2^k$ stačí nájsť koeficienty všetkých slov dĺžky k v rade $\|\mathcal{A}\|$. Pokiaľ je príslušná $\mathbb{R}$ -reprezentácia daná ako $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$, je táto úloha ekvivalentná výpočtu hodnôt $\mathbf{i}\mu(w)\mathbf{f}$ pre všetky slová $w \in \{1, 2, 3, 4\}^k$. Nasledujúci algoritmus predstavuje jeden z efektívnejších spôsobov realizácie tejto úlohy. Pred násobením dvojíc matíc totiž uprednostňuje násobenie matíc s vektormi a namiesto postupného výpočtu vektorov $\mathbf{i}\mu(u)$ pre všetky slová u dĺžky najviac k postupuje z oboch strán a počíta vektory $\mathbf{i}\mu(u)$ a $\mu(v)\mathbf{f}$ pre všetky $u, v \in \{1, 2, 3, 4\}^*$ do dĺžky približne $k/2$; priamočiarým výpočtom možno overiť, že sa takto vykoná menej skalárnych operácií, než pri jednostrannom postupe zľava doprava alebo naopak.

Algoritmus 3.12.2 (Vykreslenie bitmapy reprezentovanej automatom s váhami).

Vstup: Konečný automat $\mathcal{A}$ s váhami nad poľom $\mathbb{R}$ a abecedou $\Sigma = \{1, 2, 3, 4\}$ s $\mathcal{P}_{\mathcal{A}} = (n, \mathbf{i}, \mu, \mathbf{f})$; prirodzené číslo k .

Výstup: Pre každé $w \in \Sigma^k$ číslo $x_w \in \mathbb{R}$ vyjadrujúce intenzitu pixela reprezentovanej bitmapy s „adresou“ w pri rozlíšení $2^k \times 2^k$.

1. Polož $\ell_\varepsilon := \mathbf{i}$.
2. Opakuj pre $i = 1, \dots, \lceil k/2 \rceil$:

- 2.1. Pre všetky $u \in \Sigma^{i-1}$ a $c \in \Sigma$ polož $\ell_{uc} := \ell_u \mu(c)$.
- 2.2. Uvoľni z pamäte vektory ℓ_u pre všetky $u \in \Sigma^{i-1}$.
3. Polož $\mathbf{r}_\varepsilon := \mathbf{f}$.
4. Opakuj pre $j = 1, \dots, \lfloor k/2 \rfloor$:
 - 4.1. Pre všetky $v \in \Sigma^{j-1}$ a $c \in \Sigma$ polož $\mathbf{r}_{cv} := \mu(c)\mathbf{r}_v$.
 - 4.2. Uvoľni z pamäte vektory $\mathbf{r}_v$ pre všetky $v \in \Sigma^{j-1}$.
5. Pre všetky $w = uv$, kde $u \in \Sigma^{\lfloor k/2 \rfloor}$ a $v \in \Sigma^{\lfloor k/2 \rfloor}$, polož $x_w := \ell_u \mathbf{r}_v$.

Správnosť uvedeného dekodovacieho algoritmu formálne dokazovať nebudeme; malo by však byť zrejmé, že pre ľubovoľný vstupný automat $\mathcal{A}$ a prirodzené číslo k sú jeho výstupom hodnoty x_w také, že pre všetky $w \in \Sigma^k$ je $x_w = (\|\mathcal{A}\|, w)$. Analýzu zložitosti algoritmu 3.12.2 prenechávame čitateľovi ako jedno z cvičení v závere tejto kapitoly.

Zamerajme sa teraz na opačný problém, pri ktorom treba vstupnú bitmapu danú nekonečným³⁸ kvadrantovým stromom – čiže nejakým formálnym mocninovým radom r – zakódovať do podoby automatu s váhami.

Nasledujúci algoritmus realizuje toto kódovanie, pričom na jeho výstupe je garantovaný *minimálny* automat realizujúci vstupný rad r , kedykoľvek je rad r racionálny (a teda sa preň nejaký konečný automat s váhami skonštruovať dá). V praxi je rad r racionálny vždy, keďže sú jeho koeficienty známe len pre slová do určitej dĺžky; uzol DNC kvadrantového stromu potom možno zakaždým interpretovať tak, aby automat pre r zaručene existoval. Pri opise kódovacieho algoritmu predpokladáme existenciu procedúry zisťujúcej, či je formálny mocninový rad (daný nekonečným kvadrantovým stromom) lineárnou kombináciou niekoľkých ďalších mocninových radov. Z praktického hľadiska ide o realistický predpoklad, keďže pre konečné kvadrantové stromy s uzlami DNC sa táto úloha dá vyriešiť Gaussovou elimináciou – uzlom DNC totiž možno priradiť ľubovoľnú hodnotu a zvyšok konečného kvadrantového stromu hĺbky h zodpovedá nejakému polynómu z $\mathbb{R}\langle \Sigma^0 \cup \Sigma^1 \cup \dots \cup \Sigma^h \rangle$. Vektorový priestor takýchto polynómov je pritom izomorfný s priestorom $\mathbb{R}^d$, kde $d = 4^0 + 4^1 + \dots + 4^h = (4^{h+1} - 1)/3$. Pri overovaní, či je formálny mocninový rad r lineárnou kombináciou radov $r_1, \dots, r_n$, je hĺbka konečného stromu s uzlom DNC prislúchajúceho k radu r v nasledujúcom algoritme vždy nanajvýš rovná hĺbke stromov prislúchajúcich k $r_1, \dots, r_n$ a zaujímame sa o existenciu koeficientov $x_1, \dots, x_n \in \mathbb{R}$, pre ktoré sa lineárna kombinácia $x_1 r_1 + \dots + x_n r_n$ zhoduje s radom r na všetkých slovách udávajúcich platné „adresy“ uzlov rôznych od DNC v strome prislúchajúcom k radu r .

Algoritmus 3.12.3 (Zakódovanie bitmapy do automatu s váhami).

Vstup: Racionálny formálny mocninový rad (bitmapa) r nad abecedou $\Sigma = \{1, 2, 3, 4\}$ a poľom $\mathbb{R}$ daný nekonečným kvadrantovým stromom.

Výstup: Minimálny konečný automat $\mathcal{A}$ s váhami nad $\mathbb{R}$ a abecedou Σ taký, že $\|\mathcal{A}\| = r$.

1. Ak $r = 0$, vráť na výstupe (jediný) automat $\mathcal{A}$ s 0 stavmi a skonči. Inak pokračuj krokom 2.
2. Polož $Q := \{1\}$, $a_{1,1} := 0$, $i := 1$ a $r_1 := r$.
3. Opakuj, kým je $i \leq |Q|$:
 - 3.1. Opakuj pre všetky $c \in \Sigma$:
 - 3.1.1. Nech $n = |Q|$. Zisti, či existujú $x_1, \dots, x_n \in \mathbb{R}$ také, že $c^{-1}r_i = x_1 r_1 + \dots + x_n r_n$.
 - 3.1.1.a. Ak existujú, pre $j = 1, \dots, n$ polož $a_{i,j} := a_{i,j} + x_j c$.

³⁸Pripomeňme si, že ide o idealizáciu. Prechod od nej k reálnej implementácii bol vysvetlený v poznámke 3.12.1.

3.1.1.b. Inak polož $Q := Q \cup \{n+1\}$, $a_{i,n+1} := c$, $a_{j,n+1} := 0$ pre všetky $j \in Q \setminus \{i\}$,
 $a_{n+1,j} := 0$ pre všetky $j \in Q$ a $r_{n+1} := c^{-1}r_i$.

3.2. Polož $i := i+1$.

4. Nech $n = |Q|$. Polož $\mathbf{i}[\mathcal{A}] := (1, 0, \dots, 0)$, $A[\mathcal{A}] := (a_{i,j})_{n \times n}$ a $\mathbf{f}[\mathcal{A}] := ((r_1, \varepsilon), \dots, (r_n, \varepsilon))^T$.

Zastavenie a správnosť algoritmu 3.12.3 dokazovať nebudeme – obe vlastnosti sú bezprostrednými dôsledkami teórie vybudovanej v oddiele 3.7. Nie je ťažké dokázať, že automat skonštruovaný týmto algoritmom naozaj rozoznáva rad r , pričom počet jeho stavov nemôže byť vyšší, než dimenzia priestoru kvocientov $\mathcal{Q}(r)$ radu r .

Požiadavka racionality vstupného radu r je pri algoritme 3.12.3 zásadná. Na radoch $r \notin \mathbb{R}\text{-Rat}(\Sigma^*)$ by sa tento algoritmus zo zrejmých príčin nikdy nezastavil. Pri reálnej implementácii s uzlom DNC je ale predpoklad racionality vstupného radu vždy splnený, ako sme už vysvetlili vyššie.

Keďže nám ide iba o krátku demonštráciu možnej aplikácie automatov s váhami, nebudeme sa na tomto mieste zaoberať najnáročnejším problémom súvisiacim s práve opísanou metódou kódovania obrazu, ktorým je analýza dosahovanej miery kompresie. Obmedzíme sa tu len na konštatovanie, že hoci použitie algoritmu 3.12.3 znamená určitú úsporu, je doň na dosiahnutie konkurencieschopnosti v porovnaní s najrozšírenejšími metódami kódovania obrazu potrebné zapracovať ďalšie vylepšenia. O niektorých z nich sa možno dočítať napríklad v [3, 61].

3.13 Conwayovské a čiastočné conwayovské polokruhy

V oddieloch 2.4, 2.6, 3.2, 3.3 a 3.5 sme rozvíjali dve do veľkej miery paralelné teórie formálnych mocninových radov, automatov s váhami a racionálnych výrazov. Dokázali sme, že polokruhy formálnych mocninových radov nad spočítateľne úplnými polokruhmi sú tiež spočítateľne úplné – racionálne rady nad spočítateľne úplnými polokruhmi tak možno chápať jednoducho ako racionálne prvky nejakého spočítateľne úplného polokruhu, čo umožňuje aplikovať výsledky z oddielu 1.8. Pri „bezepsilonových“ formálnych mocninových radoch nad ľubovoľným polokruhom sme naopak nenašli iné východisko, než celú teóriu vystavať nanovo a miestami konštatovať, že dôkazy jednotlivých tvrdení sú podobné ako v prípade spočítateľne úplných polokruhov. Žiadúcou sa tak zdá byť ďalšia abstrakcia, ktorá by nás od nutnosti opakovaného dokazovania podobných tvrdení odbremenila.

V stručnosti teraz naznačíme jeden zo spôsobov, ktorým možno obidve paralelné teórie – čiže teóriu *spočítateľne úplných polokruhov* so všade definovanou iteráciou a teóriu *polokruhov formálnych mocninových radov* s iteráciou definovanou na „bezepsilonových“ radoch – zjednotiť do spoločného rámca. Pôjde o prístup založený na takzvaných *čiastočných conwayovských polokruhoch* [13, 37].

Začnime s klasickým pojmom *conwayovského polokruhu*, prostredníctvom ktorého sa k čiastočným conwayovským polokruhom postupne dopracujeme. Ide o zovšeobecnenie spočítateľne úplných polokruhov, v ktorom iterácia nie je zavedená pomocou nekonečných súčtov, ale *axiomaticky*: conwayovský polokruh je ľubovoľný polokruh, na ktorom je definovaná unárna operácia iterácie spĺňajúca určité podmienky. Týmto sú vzťahy pre iteráciu súčtu a iteráciu súčinu rovnaké ako vo vetách 2.6.1 a 2.6.2 pre spočítateľne úplné polokruhy. Rozdielom je, že v tomto prípade nejde o vety, ale o súčasť definície conwayovských polokruhov – hovoríme preto o *axióme iterácie súčtu* a o *axióme iterácie súčinu*.

Definícia 3.13.1. *Conwayovský polokruh* je šesticca $(S, +, \cdot, 0, 1, *)$ taká, že $(S, +, \cdot, 0, 1)$ je polokruh a $*$: $S \rightarrow S$ je unárna operácia *iterácie* spĺňajúca nasledujúce podmienky:

(i) Pre všetky $a, b \in S$ je $(a + b)^* = (a^*b)^*a^*$.

(ii) Pre všetky $a, b \in S$ je $(a \cdot b)^* = 1 + a(ba)^*b$.

Veta 3.13.2. *Každý spočítateľne úplný polokruh je conwayovský.*

Dôkaz. Ide o bezprostredný dôsledok vety 2.6.1 a vety 2.6.2. □

Pre prvky conwayovských polokruhov možno vybudovať celú teóriu analogickú tej z oddielu 1.8.³⁹ Možno tiež ukázať, že polokruh mocninových radov nad conwayovským polokruhom je pri vhodnej definícii iterácie takisto conwayovský. V súvislosti s podobnými úvahami iba čitateľa odkážeme na [36].

My namiesto toho budeme ďalej pokračovať v abstrakcii, keďže našim cieľom je definícia triedy polokruhov zahŕňajúcej okrem spočítateľne úplných polokruhov aj *každý* polokruh formálnych mocninových radov (bez ohľadu na polokruh koeficientov). Takéto polokruhy vo všeobecnosti conwayovské nie sú, pretože iterácia je v nich definovaná len pre „bezepsilonové“ rady. Prirodzenou ideou sa teda javí byť *zoslabenie* definície 3.13.1, pri ktorom je iterácia spĺňajúca axiomy iterácie súčtu a súčinu definovaná *len pre určitú časť* prvkov polokruhu.

To očividne nebude možné doceliť pre hociktorú podmnožinu X polokruhu S . Ak požadujeme, aby axiomy iterácie súčtu a súčinu platili pre všetky $a, b \in X$, nutne musia byť v X aj prvky $a + b$ a ab – inak by nedávali zmysel *ľavé* strany rovností z týchto axióm. Aby dávala zmysel *pravá* strana rovnosti z axiomy súčtu, musí byť v X aj prvok a^*b . Prvok a^* však sám o sebe nemusí byť v X – stačí si spomenúť na to, že iteráciou „bezepsilonového“ radu nie je „bezepsilonový“ rad.⁴⁰ Od podmnožiny polokruhu s definovanými iteráciami prvkov tak budeme požadovať uzavretosť na súčet (dvoch prvkov tejto podmnožiny) a na súčin *s ľubovoľným prvkom polokruhu*. Musí teda ísť o *ideál* v polokruhu S .

Definícia 3.13.3. Nech $(S, +, \cdot, 0, 1)$ je polokruh. *Ideál* v polokruhu S je *neprázdna* množina $I \subseteq S$ taká, že:

- (i) Pre všetky $a, b \in I$ je $a + b \in I$.
- (ii) Pre všetky $a \in I$ a $b \in S$ je $ab \in I$ a $ba \in I$.

Poznámka 3.13.4. Neprázdnosť ideálu je v predchádzajúcej definícii zrejme ekvivalentná podmienke $0 \in I$. Je navyše zrejmé, že $I = S$ práve vtedy, keď $1 \in I$.

Definícia 3.13.5. *Čiastočný conwayovský polokruh* je sedmica $(S, +, \cdot, 0, 1, I, *)$, kde $(S, +, \cdot, 0, 1)$ je polokruh, I je ideál v S a $*$: $I \rightarrow S$ je unárna operácia *iterácie* spĺňajúca nasledujúce podmienky:

- (i) Pre všetky $a, b \in I$ je $(a + b)^* = (a^*b)^*a^*$.
- (ii) Pre všetky $a, b \in S$ také, že $a \in I$ alebo $b \in I$ je $(a \cdot b)^* = 1 + a(ba)^*b$.

Podobne ako pri spočítateľne úplných polokruhoch budeme – v prípade, že nehrozí nedorozumenie – namiesto o čiastočnom conwayovskom polokruhu $(S, +, \cdot, 0, 1, I, *)$ často hovoriť len o čiastočnom conwayovskom polokruhu $(S, +, \cdot, 0, 1)$ nad ideálom I alebo o čiastočnom conwayovskom polokruhu S nad ideálom I .

Poznámka 3.13.6. Každý conwayovský polokruh S je zjavne čiastočný conwayovský nad ideálom S .

Veta 3.13.7. Každý spočítateľne úplný polokruh je čiastočný conwayovský.

Dôkaz. Vyplýva z vety 3.13.2 a zo zrejmeého pozorovania z poznámky 3.13.6. □

Veta 3.13.8. Nech Σ je abeceda a S je polokruh. Potom je polokruh $S\langle\langle\Sigma^*\rangle\rangle$ čiastočný conwayovský nad ideálom $S\langle\langle\Sigma^+\rangle\rangle$.

Dôkaz. Množina $S\langle\langle\Sigma^+\rangle\rangle$ je ideálom v $S\langle\langle\Sigma^*\rangle\rangle$ vďaka tvrdeniu 2.6.3. Samotné dokazované tvrdenie je potom dôsledkom viet 2.6.7 a 2.6.8. □

³⁹To má samozrejme určité špecifiká: napríklad z lemy 1.5.5 o iterácii blokovej matice sa v prípade „ľavého horného“ bloku veľkosti 1×1 stáva definícia a pod.

⁴⁰Presnejšie táto vlastnosť platí pre všetky „bezepsilonové“ rady nad *netriviálnym* polokruhom.

Trieda čiastočných conwayovských polokruhov teda skutočne zahŕňa ako všetky spočítateľne úplné polokruhy, tak aj všetky polokruhy formálnych mocninových radov. Vo zvyšku tohto oddielu stručne naznačíme možnosť rozvinutia teórie racionálnych prvkov čiastočných conwayovských polokruhov, ktorá zjednocuje teóriu z oddielu 1.8 s tou z oddielu 3.5.

Základom takejto zjednocujúcej teórie je rozšírenie pojmu iterácie prvku z ideálu I na štvorcové matice nad I . Východiskom je tu ten prípad lemy 1.5.5 o iterácii blokovej matice nad spočítateľne úplnými polokruhmi, pri ktorom je „ľavý horný“ blok typu 1×1 .

Definícia 3.13.9. Nech S je čiastočný conwayovský polokruh nad ideálom I , $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad I . *Iteráciou matice* A nazveme maticu A^* danú indukčne: ak $n = 0$, je $A^* = A$; ak $n = 1$, je $A^* = (a_{1,1}^*)$; ak $n \geq 2$, tak pre blokový rozklad

$$A = \begin{pmatrix} a_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

uvažovanej matice A , kde $a_{1,1} \in I$, $A_{1,2} = (a_{i,1+j})_{1 \times (n-1)} \in I^{1 \times (n-1)}$, $A_{2,1} = (a_{1+i,j})_{(n-1) \times 1} \in I^{(n-1) \times 1}$ a $A_{2,2} = (a_{1+i,1+j})_{(n-1) \times (n-1)} \in I^{(n-1) \times (n-1)}$ je

$$A^* = \begin{pmatrix} a_{1,1}[*] & A_{1,2}[*] \\ A_{2,1}[*] & A_{2,2}[*] \end{pmatrix},$$

pričom jednotlivé bloky $a_{1,1}[*]$, $A_{1,2}[*]$, $A_{2,1}[*]$ a $A_{2,2}[*]$ sú dané nasledujúcimi vzťahmi:

$$\begin{aligned} a_{1,1}[*] &= (a_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^*, \\ A_{1,2}[*] &= (a_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^*A_{1,2}A_{2,2}^*, \\ A_{2,1}[*] &= (A_{2,2} + A_{2,1}A_{1,1}^*A_{1,2})^*A_{2,1}A_{1,1}^*, \\ A_{2,2}[*] &= (A_{2,2} + A_{2,1}A_{1,1}^*A_{1,2})^*. \end{aligned}$$

Dá sa dokázať, že podobné vzťahy v skutočnosti platia pre *ľubovoľný* blokový rozklad štvorcovej matice A – tieto sú sformulované v nasledujúcej leme. Relatívne technický dôkaz tejto skutočnosti vynechávame.

Lema 3.13.10. Nech S je čiastočný conwayovský polokruh nad ideálom I , $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n}$ je štvorcová matica nad I . Nech

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$$

je blokový rozklad matice A , kde $A_{1,1} = (a_{i,j})_{n_1 \times n_1}$, $A_{1,2} = (a_{i,n_1+j})_{n_1 \times n_2}$, $A_{2,1} = (a_{n_1+i,j})_{n_2 \times n_1}$ a $A_{2,2} = (a_{n_1+i,n_1+j})_{n_2 \times n_2}$ pre nejaké $n_1, n_2 \in \mathbb{N}$. Potom

$$A^* = \begin{pmatrix} A_{1,1}[*] & A_{1,2}[*] \\ A_{2,1}[*] & A_{2,2}[*] \end{pmatrix},$$

pričom jednotlivé bloky $A_{1,1}[*] = (a_{i,j}[*])_{n_1 \times n_1}$, $A_{1,2}[*] = (a_{i,n_1+j}[*])_{n_1 \times n_2}$, $A_{2,1}[*] = (a_{n_1+i,j}[*])_{n_2 \times n_1}$ a $A_{2,2}[*] = (a_{n_1+i,n_1+j}[*])_{n_2 \times n_2}$ sú dané nasledujúcimi vzťahmi:

$$\begin{aligned} A_{1,1}[*] &= (A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^*, \\ A_{1,2}[*] &= (A_{1,1} + A_{1,2}A_{2,2}^*A_{2,1})^*A_{1,2}A_{2,2}^*, \\ A_{2,1}[*] &= (A_{2,2} + A_{2,1}A_{1,1}^*A_{1,2})^*A_{2,1}A_{1,1}^*, \\ A_{2,2}[*] &= (A_{2,2} + A_{2,1}A_{1,1}^*A_{1,2})^*. \end{aligned}$$

Indukciou vzhľadom na rád štvorcových matíc $n \in \mathbb{N}$ možno dokázať aj nasledujúcu vetu; my ju opäť len uvedieme bez dôkazu.

Veta 3.13.11. *Nech S je čiastočný conwayovský polokruh nad ideálom I a $n \in \mathbb{N}$. Potom je $S^{n \times n}$ čiastočný conwayovský polokruh nad $I^{n \times n}$.*

Ak je potom S čiastočný conwayovský polokruh nad ideálom I , možno pre dané množiny $I' \subseteq I$ a $S' \subseteq S$ (kde $0 \in I'$ a $0, 1 \in S'$) definovať ako analógiu ku konečným automatom nad spočítateľne úplným polokruhom (oddiel 1.6) *konečné (S', I') -automaty nad (S, I)* . Pre takýto automat $\mathcal{A}$ daný vektorom počiatkových ohodnotení $\mathbf{i} \in (S')^{1 \times n}$, prechodovou maticou $A \in (I')^{n \times n}$ a vektorom koncových ohodnotení $\mathbf{f} \in (S')^{n \times 1}$ potom *definujeme* $\|\mathcal{A}\| = \mathbf{i}A^*\mathbf{f}$. Tento koncept zahŕňa ako konečné S' -automaty nad spočítateľne úplným polokruhom S – tam jednoducho $I = S$ a $I' = S'$ – tak aj „bezepsilonové“ konečné automaty s váhami nad polokruhom S a abecedou Σ , ktoré možno chápať ako $(S\langle \Sigma \cup \{\varepsilon\} \rangle, S\langle \Sigma \rangle)$ -automaty nad $(S\langle \Sigma^* \rangle, S\langle \Sigma^+ \rangle)$. Obdobne možno definovať aj *racionálne výrazy* nad čiastočným conwayovským polokruhom a s využitím lemy 3.13.10 dokázať ich ekvivalenciu s konečnými automatmi.

Racionálnym prvkom čiastočného conwayovského polokruhu nad ideálom I možno nazvať ľubovoľný prvok realizovaný niektorým z týchto ekvivalentných modelov. Racionálne prvky spočítateľne úplného polokruhu S tak možno chápať ako racionálne prvky čiastočného conwayovského polokruhu S nad ideálom S . Racionálne mocninové rady nad abecedou Σ a ľubovoľným polokruhom S zas možno chápať ako racionálne prvky čiastočného conwayovského polokruhu $S\langle \Sigma^* \rangle$ nad ideálom $S\langle \Sigma^+ \rangle$.

Cvičenia

1. Dokážte na plne rigoróznej úrovni vlastnosť „bezepsilonových“ automatov s váhami, chápaných v zmysle ľubovoľnej z ich alternatívnych definícií, hojne využívanú v texte: pre každý takýto automat $\mathcal{A}$ je systém radov $(\|\gamma\| \mid \gamma \in \mathcal{R}(\mathcal{A}))$ lokálne konečný a rad $\|\mathcal{A}\|$ realizovaný automatom $\mathcal{A}$ je teda dobre definovaný.
2. V oddiele 3.3 je formálne opísaná korešpondencia medzi automatmi s váhami chápanými podľa definície 3.2.1 a podľa definície 3.3.1. Formálne dokážte, že pre ľubovoľnú dvojicu automatov v tejto korešpondencii je nimi realizovaný rad rovnaký, kedykoľvek je dobre definovaný.
3. Konečné automaty s váhami nad polokruhom S a abecedou Σ sú špeciálnym prípadom konečných $S\langle \Sigma \cup \{\varepsilon\} \rangle$ -automatov nad $S\langle \Sigma^* \rangle$, v ktorom sú všetky počiatkové aj koncové ohodnotenia stavov skalárne. Spomenuli sme, že pre spočítateľne úplný polokruh S zodpovedajú konečné automaty s váhami nad S normálnemu tvaru konečných $S\langle \Sigma \cup \{\varepsilon\} \rangle$ -automatov nad $S\langle \Sigma^* \rangle$. Dokážte analogické tvrdenie pre „bezepsilonové“ automaty nad ľubovoľným polokruhom – to jest predovšetkým vhodne *definujte* prvok polokruhu $S\langle \Sigma^* \rangle$ realizovaný konečným $S\langle \Sigma \cup \{\varepsilon\} \rangle$ -automatom nad $S\langle \Sigma^* \rangle$ s „bezepsilonovou“ prechodovou maticou a následne dokážte, že každý takýto automat je ekvivalentný nejakému konečnému automatu s váhami nad S .
4. V definícii racionálnych výrazov s váhami sa narábalo s atomickými výrazmi r pre $r \in S\langle \Sigma \cup \{\varepsilon\} \rangle$. Ukážte, že je táto množina atomických výrazov zbytočne obsiahla.
5. Sformalizujte argumentáciu z dôkazu lemy 3.5.6.
6. Nech Σ je abeceda, S je polokruh a $r, s \in S\langle \Sigma^* \rangle$ sú mocninové rady. *Hadamardov súčin* radov r a s je rad $r \odot s \in S\langle \Sigma^* \rangle$ definovaný pre všetky $w \in \Sigma^*$ predpisom $(r \odot s, w) = (r, w)(s, w)$.
 - a) Ukážte, že Hadamardov súčin na $\mathbb{B}\langle \Sigma^* \rangle$ možno interpretovať ako prienik formálnych jazykov.
 - b) Dokážte, že pre všeobecný polokruh S a abecedu Σ množina radov $S\text{-Rat}(\Sigma^*)$ nemusí byť uzavretá na Hadamardov súčin.
 - c) Dokážte, že množina radov $S\text{-Rat}(\Sigma^*)$ je uzavretá na Hadamardov súčin pre ľubovoľný komutatívny polokruh S a ľubovoľnú abecedu Σ .

7. Nech Σ je abeceda a $r \in \mathbb{N}\langle\langle\Sigma^*\rangle\rangle$ je rad racionálny nad $\mathbb{N}$. Zistite, či musí byť v takom prípade racionálny jazyk $\text{supp}(r)$.
8. Nech Σ je abeceda a $r \in \mathbb{Z}\langle\langle\Sigma^*\rangle\rangle$ je rad racionálny nad $\mathbb{Z}$. Zistite, či musí byť v takom prípade racionálny jazyk $\text{supp}(r)$.

Hovoríme, že polokruh T je *Fatouovým rozšírením* polokruhu S , ak je S podpolokruhom T a súčasne pre ľubovoľnú abecedu Σ a rad $r \in S\langle\langle\Sigma^*\rangle\rangle$ racionálny nad T je rad r racionálny aj nad S .

9. Dokážte, že okruh $\mathbb{Z}$ nie je Fatouovým rozšírením polokruhu $\mathbb{N}$.
10. Dokážte, že každé rozšírenie $\mathbb{K}$ poľa $\mathbb{F}$ je aj jeho Fatouovým rozšírením.
11. Dokážte tvrdenie 3.7.6.
12. Nech $\Sigma = \{a, b\}$. Zistite, či existuje formálny mocninový rad $r \in \mathbb{Z}\langle\langle\Sigma^*\rangle\rangle$ racionálny nad $\mathbb{Z}$ taký, že $\text{supp}(r) = \{w \in \Sigma^* \mid |w|_a = |w|_b\}$.
13. V leme 3.7.10 sme dokázali, že pre každý „bezepsilonový“ konečný automat $\mathcal{A} = (Q, \varphi, \iota, \tau)$ nad abecedou Σ s váhami nad poľom $\mathbb{F}$ je $\dim(\mathcal{Q}(\|\mathcal{A}\|)) \leq \dim(\mathcal{V}(\mathcal{A})) \leq |Q|$. Nájdite príklad automatu $\mathcal{A}$, pre ktorý sú obidve tieto rovnosti ostré (alebo dokážte, že taký automat neexistuje).
14. Analogicky k ľavým kvocientom definujte *pravý kvocient* rx^{-1} formálneho mocninového radu $r \in \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ podľa slova x . Dajte do súvisu dimenziu vektorového priestoru generovaného pravými kvocientmi racionálneho mocninového radu r a počet stavov minimálneho automatu realizujúceho rad r . Vysvetlite, ako by sa pomocou pravých kvocientov dala teória z oddielu 3.7 „urobiť symetrickou“.
15. a) Dokážte, že pre ľubovoľnú maticu A typu $m \times n$ nad poľom $\mathbb{F}$ s riadkovou hodnotou m je jej pravou inverznou maticou (okrem iného aj) matica $Z := A^T(AA^T)^{-1}$. *Pozor*: treba overiť aj korektnosť definície matice Z .
b) Sformulujte a dokážte analogické tvrdenie pre ľavé inverzné matice.
16. Nech Σ je abeceda a $\mathbb{F}$ je pole. *Reverzom* formálneho mocninového radu $r \in \mathbb{F}\langle\langle\Sigma^*\rangle\rangle$ môžeme nazvať, podobne ako pri jazykoch, rad

$$r^R = \sum_{w \in \Sigma^*} (r, w^R)w,$$

kde w^R označuje bežný reverz – alebo zrkadlový obraz – slova w .

- a) Dokážte, že ak $\|\mathcal{P}\| = r$ pre $\mathbb{F}$ -reprezentáciu $\mathcal{P} = (\mathbf{i}, \mu, \mathbf{f})$ rádu n , tak pre *transponovanú reprezentáciu* $\mathcal{P}^T = (\mathbf{f}^T, \mu^T, \mathbf{i}^T)$, kde homomorfizmus $\mu^T: \Sigma^* \rightarrow \mathbb{F}^{n \times n}$ je pre všetky $c \in \Sigma$ daný ako $\mu^T(c) = \mu(c)^T$, je $\|\mathcal{P}^T\| = r^R$.
- b) Dokážte, že počet stavov minimálneho automatu s váhami nad poľom $\mathbb{F}$ realizujúceho rad r je rovnaký ako počet stavov minimálneho automatu realizujúceho rad r^R .
- c) Vyvoďte z predchádzajúcich pozorovaní záver, že minimalizáciu konečných automatov s váhami nad poľom $\mathbb{F}$ možno realizovať aj „opačne“ ako v algoritme 3.9.3 – to jest najprv možno aplikovať konštrukciu z lemy 3.9.1 a až následne konštrukciu z lemy 3.9.2.
17. Dokážte tvrdenie 3.9.7.
18. Dokážte rozhodnuteľnosť problému ekvivalencie „bezepsilonových“ konečných automatov s váhami nad *konečnými* polokruhmi.

19. Preskúmajte rozhodnuteľnosť problému univerzality nosiča realizovaného formálneho mocninného radu pre „bezepsilonové“ konečné automaty s váhami nad polokruhom prirodzených čísel $(\mathbb{N}, +, \cdot, 0, 1)$ a nad tropickým polokruhom $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$.
20. Uvažujme nasledujúci rozhodovací problém: na vstupe je daný „bezepsilonový“ konečný automat $\mathcal{A}$ s váhami nad poľom $\mathbb{Q}$ a abecedou Σ a treba rozhodnúť, či pre všetky $w \in \Sigma^*$ je $(\|\mathcal{A}\|, w) \in \{0, 1\}$. Je tento problém rozhodnuteľný? (*Návod:* použite uzavretosť množiny radov $\mathbb{Q}\text{-Rat}(\Sigma^*)$ na Hadamardov súčin dokázanú v rámci úlohy 6.)
21. Uvažujme nasledujúci rozhodovací problém: na vstupe je daný „bezepsilonový“ konečný automat $\mathcal{A}$ s váhami nad poľom $\mathbb{Q}$ a abecedou Σ a treba rozhodnúť, či pre všetky $w \in \Sigma^*$ je $(\|\mathcal{A}\|, w) > 0$. Je tento problém rozhodnuteľný? (*Návod:* použite uzavretosť množiny radov $\mathbb{Q}\text{-Rat}(\Sigma^*)$ na Hadamardov súčin dokázanú v rámci úlohy 6.)
22. „Bezepsilonový“ automat s váhami nad polokruhom S a abecedou Σ nazveme *deterministickým*, ak najviac jeden spomedzi jeho stavov má nenulovú počiatočnú váhu a ak z každého jeho stavu vedie na každé písmeno abecedy Σ najviac jeden prechod. Dokážte, že automaty s váhami vo všeobecnosti nemusia byť determinizovateľné: nájdite príklad radu $r \in \mathbb{N}\langle\langle\Sigma^*\rangle\rangle$ racionálneho nad $\mathbb{N}$, ktorý nie je realizovaný žiadnym deterministickým konečným automatom s váhami nad $\mathbb{N}$ a Σ .
23. Dokážte, že „bezepsilonové“ konečné automaty s váhami nad *konečnými* polokruhmi sú vždy determinizovateľné.
24. Asymptoticky odhadnite počet vykonaných skalárnych operácií – najmä násobení reálnych čísel – v algoritme 3.12.2. Porovnajme takto chápanú časovú zložitosť algoritmu 3.12.2 so zložitou alternatívneho algoritmu, v ktorom sa namiesto obojstranného prístupu aplikuje priamočiarejší jednostranný prístup spočívajúci v postupnom výpočte vektorov $\mathbf{i}\mu(u)$ pre všetky $u \in \{1, 2, 3, 4\}^*$ dĺžky najviac k , pričom výsledná hodnota x_w je pre každé $w \in \{1, 2, 3, 4\}^k$ získaná pre násobením riadkového vektora $\mathbf{i}\mu(w)$ stĺpcovým vektorom $\mathbf{f}$.

Poznámky a odkazy na samoštúdium

Teória automatov s váhami vychádza z práce M.-P. Schützenbergera [98], ktorý študoval automaty s váhami nad poľom a opísal aj minimalizačný algoritmus pre takéto automaty. Najobsiahlejší zdroj informácií o automatoch s váhami je v súčasnosti monografia [30] – z hľadiska teórie sú pre naše účely relevantné predovšetkým jej kapitoly [29, 36, 95] a aplikácii automatov s váhami pri kódovaní obrazu sa venuje kapitola [3]; o aplikáciách automatov s váhami pri spracovaní prirodzeného jazyka sa možno dočítať v kapitole [65]. Do širšieho kontextu modernej teórie automatov sú automaty s váhami zasadené v učebnici [94]. Z ďalších zdrojov spomeňme knihy [11, 75, 97] a prehľadovú kapitolu [31].

Podobne ako pri iterácii formálnych mocninných radov, dá sa aj pri definícii radov realizovaných automatmi s váhami predpoklad „bezepsilonovosti“ automatu zovšeobecniť na predpoklad „bezepsilonovosti“ niektorej mocniny jeho prechodovej matice (rad potom bude okrem iného dobre definovaný pre všetky automaty neobsahujúce orientovanú kružnicu tvorenú prechodmi na ε , a to nad *ľubovoľným* polokruhom). Podrobnosti možno nájsť v [36].

Prezentácia teórie okolo minimálnych automatov s váhami nad poľom je v tomto texte do veľkej miery založená na [95], miestami je ale mierne upravená s cieľom využívať len čo možno najelementárnejšie techniky. Minimalizačný algoritmus pre automaty s váhami nad poľom je opísaný aj v [11, 94]. V [94, 95] sa dokonca uvažujú o niečo všeobecnejšie automaty nad *telesom* (angl. *division ring* alebo *skew field*, pričom prívlastok „*skew*“ sa pri druhom pomenovaní v závislosti od kontextu občas vynecháva). Ako prvý skúmal minimalizáciu automatov nad poľom M.-P. Schützenberger [98]. Verziu minimalizačného algoritmu opísanú v tomto texte po prvý raz publikovali A. Cardon a M. Crochemore [19].

Nerozhodnuteľnosť ekvivalencie konečných automatov s váhami nad tropickým polokruhom $(\mathbb{N} \cup \{\infty\}, \min, +, \infty, 0)$ ako prvý dokázal D. Krob [73]; jeho dôkazová technika využíva redukciu Hilbertovho desiateho problému. S jednoduchším dôkazom, založeným na redukcii problému zastavenia pre Minského stroje s dvoma počítadlami, neskôr prišli S. Almagor, U. Boker a O. Kupfermanová [4, 5] a nezávisle na nich T. Colcombet, ktorého dôkaz možno nájsť v [31].

S myšlienkou kódovania obrazu pomocou automatov s váhami po prvý raz prišli K. Čulík II a J. Kari [24]; ide pritom o jednu z najstarších, hoci pravdepodobne nie najvýznamnejších, aplikácií automatov s váhami. Prístupnými úvodmi do problematiky sú [3, 61], kde sú spracované aj jej praktickejšie aspekty.

Dnes už pomerne dobre preskúmané conwayovské polokruhy [29, 36] sú pomenované po J. H. Conwayovi, ktorý ako prvý identifikoval kľúčový význam axióm iterácie súčtu a iterácie súčinu [22]. S konceptom čiastočných conwayovských polokruhov nad ideálom prišli v roku 2008 S. L. Bloom, Z. Ésik a W. Kuich [13, 37]. Motiváciou za týmto pojmom bolo predovšetkým zjednotenie teórie konečných automatov nad conwayovskými (a teda aj spočítateľne úplnými a úplnými) polokruhmi s teóriou „bezepsilonových“ automatov s váhami nad ľubovoľným polokruhom. Popri čiastočných conwayovských polokruhoch sú spoločnou abstrakciou spočítateľne úplných polokruhov s iteráciou definovanou na všetkých prvkoch a polokruhov formálnych mocninových radov s iteráciou „bezepsilonových“ radov aj takzvané *sumačné polokruhy*, známe aj ako Σ -*polokruhy* [52, 67, 68].

Popri teórii racionálnych jazykov možno na formálne mocninové rady s nekomutujúcimi premennými zovšeobecniť napríklad aj teóriu bezkontextových jazykov. Podobne ako sú racionálne rady obdobou racionálnych funkcií, sú rady realizované bezkontextovými gramatikami (prípadne zásobníkovými automatmi) s váhami obdobou *algebraických* funkcií – hovoríme teda o *algebraických formálnych mocninových radoch* a bezkontextové gramatiky s váhami sa často študujú v podobe „izomorfného“ modelu *algebraických systémov*. Tu čitateľa odkazujeme napríklad na [75, 97, 91, 74, 38].

Kapitola 4

Trocha lineárnej algebry

Teória okolo *vlastných čísel a vlastných vektorov* nepochybne patrí k najužitočnejším partiám lineárnej algebry a základným prostriedkom bude v nasledujúcich kapitolách aj pre nás. V ideálnom prípade je čitateľovi už známa – túto kapitolu potom možno chápať ako jej stručné zopakovanie. V menej ideálnom komplementárnom prípade, ktorý u predpokladaného čitateľa takisto nie je vylúčený, môže táto kapitola slúžiť aj ako samostatný úvod do problematiky postačujúci pre účely tohto textu.

Väčšina tejto kapitoly tak bude venovaná štandardnému materiálu, ktorý je podrobnejšie a v širších súvislostiach rozpracovaný v prakticky ľubovoľnej dobrej učebnici lineárnej algebry alebo maticového počtu – za všetky spomeňme napríklad [105, 7, 77, 57]. Výnimkou bude len oddiel 4.3 tvoriaci akýsi predvoj kapitol 5 a 6. V nich svojím spôsobom nadviažeme na kapitolu 1, z ktorej sú nám známe viaceré aplikácie umocňovania štvorcovej matice. Uvidíme, že pre ľubovoľnú maticu A nad poľom $\mathbb{C}$ – a teda aj pre matice nad $\mathbb{N}$ alebo $\mathbb{R}$ – možno metódami lineárnej algebry vyjadriť jednotlivé prvky matice A^t v uzavretom tvare. V kapitole 5 toto pozorovanie aplikujeme na úlohy súvisiace s grafmi a formálnymi jazykmi a v kapitole 6 na riešenie určitých typov rekurencií – alebo *diferenčných rovníc* – a ich systémov. V oddiele 4.3 si tieto aplikácie ukážeme na špeciálnom prípade tzv. *diagonalizovateľných* matíc: ten síce nebude univerzálne použiteľný, zato sa však ukáže ako mimoriadne jednoduchý a kľúčové myšlienky neskorších kapitol tak budeme môcť demonštrovať ešte predtým, než prejdeme k technickejším pasážam okolo zovšeobecnených vlastných vektorov a Jordanovho kanonického tvaru.

V tejto kapitole sa obmedzíme na matice s prvkami z poľa $\mathbb{C}$. Analogickú teóriu ale možno rozvinúť pre matice nad ľubovoľným algebraicky uzavretým poľom [105] – čiže poľom $\mathbb{F}$ takým, že každý nekonštantný polynóm $p(x) \in \mathbb{F}[x]$ má koreň v $\mathbb{F}$.

Namiesto samotných matíc nad $\mathbb{C}$ – čiže lineárnych zobrazení na $\mathbb{C}^n$ – však budeme často pracovať vo všeobecnejšom kontexte *lineárnych zobrazení na konečnorozmernom vektorovom priestore* $\mathcal{V}$ nad $\mathbb{C}$.¹ Ide o zvyčajný prístup výhodný najmä vďaka tomu, že umožňuje pre dané zobrazenie φ na $\mathbb{C}^n$ skúmať jeho zúženia na podpriestory $\mathbb{C}^n$ invariantné vzhľadom na φ aj bez nutnosti zmeny súradníc.

Poznámka 4.0.1. V tejto a aj v nasledujúcich kapitolách budeme množinu $\mathbb{C}^n$ chápať ako $\mathbb{C}^{n \times 1}$ – vektory komplexných čísel tak budú prevažne považované za stĺpcové a *lineárne zobrazenie* $\varphi_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ prislúchajúce k matici $A \in \mathbb{C}^{n \times n}$ bude pre všetky $\mathbf{x} \in \mathbb{C}^n$ dané ako $\varphi_A: \mathbf{x} \mapsto A\mathbf{x}$.

Nech je naopak $\mathcal{V}$ konečnorozmerný vektorový priestor nad poľom $\mathbb{C}$ s bázou $\mathcal{B} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. *Súradnicami vektora* $\mathbf{x} \in \mathcal{V}$ *vzhľadom na bázu* $\mathcal{B}$ nazveme vektor $\mathbf{x}[\mathcal{B}] = (a_1, \dots, a_n)^T \in \mathbb{C}^n$, kde $a_1, \dots, a_n \in \mathbb{C}$ sú konštanty také, že $\mathbf{x} = a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n$. *Maticou zobrazenia* φ *vzhľadom na bázu* $\mathcal{B}$ nazveme maticu

$$A_\varphi[\mathcal{B}] = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \varphi(\mathbf{u}_1)[\mathcal{B}] & \varphi(\mathbf{u}_2)[\mathcal{B}] & \dots & \varphi(\mathbf{u}_n)[\mathcal{B}] \\ \downarrow & \downarrow & & \downarrow \end{pmatrix};$$

¹Presnejšie pôjde o lineárne zobrazenia typu $\varphi: \mathcal{V} \rightarrow \mathcal{V}$. Takéto lineárne zobrazenia s rovnakým oborom a kooborom sa niekedy nazývajú aj *lineárne operátory* [105]; my túto terminológiu používať nebudeme a pojem lineárneho operátora si rezervujeme pre lineárne zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ na *nekonečnorozmernom* vektorovom priestore $\mathcal{V}$.

pripomeňme si, že šípky v uvedenom zápise slúžia len na vizuálne zdôraznenie stĺpcovosti vektorov. Pre všetky $\mathbf{x} \in \mathcal{V}$ potom evidentne $\varphi(\mathbf{x})[\mathcal{B}] = A_\varphi[\mathcal{B}]\mathbf{x}[\mathcal{B}]$. Ak $\mathcal{V} = \mathbb{C}^n$ pre nejaké $n \in \mathbb{N}$ a $\mathcal{B}$ je štandardná báza $\mathbb{C}^n$, píšeme namiesto $A_\varphi[\mathcal{B}]$ len A_φ . Pre všetky $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$ potom evidentne $A_{\varphi_A} = A$, kým pre všetky lineárne zobrazenia $\varphi: \mathbb{C}^n \rightarrow \mathbb{C}^n$ je $\varphi_{A_\varphi} = \varphi$.

Ak je číslo n zrejmé z kontextu, budeme pre jednotkovú maticu $\mathbf{I}_n$ písať len $\mathbf{I}$. *Identické zobrazenie* $\text{Id}_\mathcal{V}: \mathcal{V} \rightarrow \mathcal{V}$ na vektorovom priestore $\mathcal{V}$ budeme často označovať len ako Id .

4.1 Vlastné čísla, vlastné vektory a diagonalizácia matice

Štvorcovú maticu A typu $n \times n$ nad poľom $\mathbb{C}$ možno v zmysle poznámky 4.0.1 chápať ako maticu lineárneho zobrazenia φ_A vzhľadom na štandardnú bázu priestoru $\mathbb{C}^n$. Teória z tejto kapitoly je do veľkej miery hnaná snahou opísať zobrazenie φ_A čo možno najjednoduchším spôsobom. Základnou myšlienkou je pritom nájsť určité špeciálne vektory z $\mathbb{C}^n$ tak, aby sa „pôsobenie“ zobrazenia φ_A na týchto vektoroch dalo vyjadriť pokiaľ možno jednoduchšie, než pomocou matice A ; ak tieto vektory tvoria bázu $\mathbb{C}^n$, možno zobrazenie φ_A charakterizovať jeho maticou vzhľadom na túto bázu. Najjednoduchším typom takýchto vektorov sú *vlastné vektory*, ktoré teraz definujeme. Z nich však ešte nebude zakaždým možné vybrať bázu – túto vlastnosť budú mať až zovšeobecnené vlastné vektory (oddiel 4.5).

Definícia 4.1.1. Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Ak pre nejaký nenulový vektor $\mathbf{v} \in \mathcal{V}$ a nejaké číslo $\lambda \in \mathbb{C}$ je

$$\varphi(\mathbf{v}) = \lambda \mathbf{v},$$

hovoríme, že $\mathbf{v}$ je *vlastný vektor* zobrazenia φ a λ je *vlastné číslo* zobrazenia φ prislúchajúce k $\mathbf{v}$.

Definícia 4.1.2. Nech $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$ je štvorcová matica. *Vlastným vektorom* matice A nazveme ľubovoľný vlastný vektor $\mathbf{v} \in \mathbb{C}^n$ zobrazenia φ_A a *vlastným číslom* matice A prislúchajúcim vektoru $\mathbf{v}$ nazveme vlastné číslo λ zobrazenia φ_A prislúchajúce vektoru $\mathbf{v}$.

Vektor $\mathbf{v} \in \mathbb{C}^n$ je teda *vlastným vektorom* matice $A \in \mathbb{C}^{n \times n}$ s prislúchajúcim vlastným číslom $\lambda \in \mathbb{C}$ práve vtedy, keď $\mathbf{v} \neq \mathbf{0}$ a

$$A\mathbf{v} = \lambda \mathbf{v}.$$

Je zrejmé, že podpriestor $[\mathbf{v}]$ vektorového priestoru $\mathcal{V}$ generovaný vlastným vektorom $\mathbf{v}$ zobrazenia φ prislúchajúcim k nejakému vlastnému číslu λ je *invariantný* vzhľadom na φ : pre všetky $\mathbf{x} \in [\mathbf{v}]$ patrí do $[\mathbf{v}]$ aj vektor $\varphi(\mathbf{x})$; navyše nutne $\varphi(\mathbf{x}) = \lambda \mathbf{x}$ a vektor $\mathbf{x}$ teda je, pokiaľ je nenulový, *tiež vlastným vektorom* prislúchajúcim k λ . Na podpriestore $[\mathbf{v}]$ sa tak lineárne zobrazenie φ dá namiesto matice opísať iba konštantou λ . Za účelom dôkazu týchto tvrdení si stačí uvedomiť, že každý vektor $\mathbf{x} \in [\mathbf{v}]$ sa dá pre nejaké $c \in \mathbb{C}$ vyjadriť ako $\mathbf{x} = c\mathbf{v}$, z čoho $\varphi(\mathbf{x}) = \varphi(c\mathbf{v}) = c\varphi(\mathbf{v}) = c\lambda \mathbf{v} = \lambda c\mathbf{v} = \lambda \mathbf{x}$.

Ku každému vlastnému vektoru očividne prislúcha práve jedno vlastné číslo. Z argumentácie vyššie vyplýva, že naopak táto vlastnosť neplatí: ak vlastné číslo λ prislúcha k vlastnému vektoru $\mathbf{v}$, sú vlastnými vektormi prislúchajúcimi k λ prinajmenšom aj všetky ostatné nenulové vektory z jednorozmerného priestoru generovaného vektorom $\mathbf{v}$. Ako uvidíme neskôr, priestor generovaný vlastnými vektormi prislúchajúcimi k danému vlastnému číslu nemusí byť jednorozmerný: k vlastnému číslu λ môže prislúchať viacero lineárne nezávislých vlastných vektorov. To motivuje nasledujúce definície.

Definícia 4.1.3. Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie a $\lambda \in \mathbb{C}$ je jeho vlastné číslo. *Vlastným priestorom* zobrazenia φ prislúchajúcim k vlastnému číslu λ nazveme podpriestor $E(\varphi, \lambda)$ priestoru $\mathcal{V}$ generovaný množinou všetkých vlastných vektorov zobrazenia φ prislúchajúcich k λ .

Definícia 4.1.4. Nech $n \in \mathbb{N}$, $A \in \mathbb{C}^{n \times n}$ je štvorcová matica a $\lambda \in \mathbb{C}$ je jej vlastné číslo. *Vlastným priestorom* matice A prislúchajúcim k vlastnému číslu λ nazveme priestor $E(A, \lambda) := E(\varphi_A, \lambda)$.

Je zrejmé, že $E(A, \lambda)$ je podpriestor priestoru $\mathbb{C}^n$ generovaný množinou všetkých vlastných vektorov matice A prislúchajúcich k λ .

Nech je teraz $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ lineárne zobrazenie, λ jeho vlastné číslo a $\mathbf{x} \in E(\varphi, \lambda)$. Potom existuje prirodzené číslo m , vlastné vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ zobrazenia φ prislúchajúce k λ a čísla $a_1, \dots, a_m \in \mathbb{C}$ také, že $\mathbf{x} = a_1\mathbf{v}_1 + \dots + a_m\mathbf{v}_m$. Z toho

$$\varphi(\mathbf{x}) = a_1\varphi(\mathbf{v}_1) + \dots + a_m\varphi(\mathbf{v}_m) = a_1\lambda\mathbf{v}_1 + \dots + a_m\lambda\mathbf{v}_m = \lambda(a_1\mathbf{v}_1 + \dots + a_m\mathbf{v}_m) = \lambda\mathbf{x}.$$

Môžeme teda uzavrieť, že pre vlastný priestor $E(\varphi, \lambda)$ platí

$$E(\varphi, \lambda) = \{\mathbf{x} \in \mathcal{V} \mid \varphi(\mathbf{x}) = \lambda\mathbf{x}\}; \quad (4.1)$$

pozostáva teda z práve všetkých vlastných vektorov zobrazenia φ prislúchajúcich k λ a z nulového vektora. Nasledujúce tvrdenie je len iným vyjadrením tejto vlastnosti.

Tvrdenie 4.1.5. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie a $\lambda \in \mathbb{C}$ je vlastné číslo zobrazenia φ . Potom*

$$E(\varphi, \lambda) = \text{Ker}(\lambda\text{Id} - \varphi) = \text{Ker}(\varphi - \lambda\text{Id}).$$

Dôkaz. Z rovnosti (4.1) vyplýva, že vektor $\mathbf{x}$ patrí do $E(\varphi, \lambda)$ práve vtedy, keď $\varphi(\mathbf{x}) = \lambda\mathbf{x} = \lambda\text{Id}(\mathbf{x})$. To je práve vtedy, keď $\lambda\text{Id}(\mathbf{x}) - \varphi(\mathbf{x}) = \varphi(\mathbf{x}) - \lambda\text{Id}(\mathbf{x}) = \mathbf{0}$, čiže keď $(\lambda\text{Id} - \varphi)(\mathbf{x}) = (\varphi - \lambda\text{Id})(\mathbf{x}) = \mathbf{0}$. Táto rovnosť je ale z definície jadra splnená práve vtedy, keď $\mathbf{x} \in \text{Ker}(\lambda\text{Id} - \varphi) = \text{Ker}(\varphi - \lambda\text{Id})$. $\square$

Dôsledok 4.1.6. *Nech $n \in \mathbb{N}$, $A \in \mathbb{C}^{n \times n}$ je štvorcová matica a $\lambda \in \mathbb{C}$ je vlastné číslo matice A . Potom*

$$E(A, \lambda) = \text{Ker}(\lambda\mathbf{I} - A) = \text{Ker}(A - \lambda\mathbf{I}).$$

Budeme sa teraz zaoberať vyjadreniami lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$, kde $\mathcal{V}$ je n -rozmerný vektorový priestor nad $\mathbb{C}$, vzhľadom na rôzne bázy $\mathcal{V}$. Pre dvojicu (usporiadaných) báz $\mathcal{B}_1 = (\mathbf{x}_1, \dots, \mathbf{x}_n)$, $\mathcal{B}_2 = (\mathbf{y}_1, \dots, \mathbf{y}_n)$ priestoru $\mathcal{V}$ nazveme *maticou prechodu od bázy $\mathcal{B}_1$ k báze $\mathcal{B}_2$* štvorcovú maticu

$$P[\mathcal{B}_2 \leftarrow \mathcal{B}_1] = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{x}_1[\mathcal{B}_2] & \mathbf{x}_2[\mathcal{B}_2] & \dots & \mathbf{x}_n[\mathcal{B}_2] \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

typu $n \times n$ nad poľom $\mathbb{C}$. Ľahko vidieť, že pre všetky $\mathbf{x} \in \mathcal{V}$ je

$$P[\mathcal{B}_2 \leftarrow \mathcal{B}_1]\mathbf{x}[\mathcal{B}_1] = \mathbf{x}[\mathcal{B}_2].$$

Z toho ďalej vyplýva, že matica $P[\mathcal{B}_2 \leftarrow \mathcal{B}_1]$ je pre každú dvojicu báz $\mathcal{B}_1, \mathcal{B}_2$ nesingulárna, pričom

$$P[\mathcal{B}_2 \leftarrow \mathcal{B}_1]^{-1} = P[\mathcal{B}_1 \leftarrow \mathcal{B}_2].$$

Poznámka 4.1.7. Uvažujme teraz špeciálny prípad, keď $\mathcal{V} = \mathbb{C}^n$ pre nejaké $n \in \mathbb{N}$. Označme $\mathcal{B}_{st}(\mathbb{C}^n)$ *standardnú bázu* $\mathcal{B}_{st}(\mathbb{C}^n) = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ priestoru $\mathbb{C}^n$ tvorenú vektormi

$$\mathbf{e}_1 = (1, 0, \dots, 0)^T, \mathbf{e}_2 = (0, 1, \dots, 0)^T, \dots, \mathbf{e}_n = (0, 0, \dots, 1)^T.$$

Ľahko potom vidieť, že pre každú nesingulárnu maticu

$$A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \in \mathbb{C}^{n \times n}$$

je $A = P[\mathcal{B}_{st}(\mathbb{C}^n) \leftarrow \mathcal{B}]$, kde $\mathcal{B}$ označuje bázu $\mathcal{B} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$. Každú nesingulárnu štvorcovú maticu typu $n \times n$ nad poľom $\mathbb{C}$ tak možno interpretovať ako maticu prechodu od bázy priestoru $\mathbb{C}^n$ tvorenej jej stĺpcami k standardnej báze.

Množina všetkých nesingulárnych matíc v $\mathbb{C}^{n \times n}$ je teda rovná množine všetkých matíc prechodu medzi bázami vektorového priestoru $\mathbb{C}^n$. Nie je ťažké ukázať, že táto množina je ďalej rovná množine všetkých matíc prechodu medzi bázami ľubovoľného n -rozmerného vektorového priestoru $\mathcal{V}$; nám však bude stačiť už učené pozorovanie, podľa ktorého je každá takáto matica prechodu nesingulárna.

Tvrdenie 4.1.8. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie a $\mathcal{B}_1, \mathcal{B}_2$ sú bázy priestoru $\mathcal{V}$. Potom*

$$A_\varphi[\mathcal{B}_2] = P[\mathcal{B}_2 \leftarrow \mathcal{B}_1] A_\varphi[\mathcal{B}_1] P[\mathcal{B}_1 \leftarrow \mathcal{B}_2].$$

Dôkaz. Jednoduché cvičenie prenechané čitateľovi. □

Matice $A, B \in \mathbb{C}^{n \times n}$ nazveme *podobnými*, ak ide o matice rovnakého lineárneho zobrazenia na $\mathbb{C}^n$, prípadne nad rôznymi bázami. Existujú teda bázy $\mathcal{B}_1, \mathcal{B}_2$ také, že $B = P[\mathcal{B}_2 \leftarrow \mathcal{B}_1] A P[\mathcal{B}_1 \leftarrow \mathcal{B}_2]$. Z pozorovaní učiných vyššie vyplýva, že táto situácia nastane práve vtedy, keď existuje nesingulárna matica P taká, že $B = P^{-1}AP$. Práve túto alternatívnu formuláciu využijeme v nasledujúcej definícii.

Definícia 4.1.9. Nech $n \in \mathbb{N}$ a $A, B \in \mathbb{C}^{n \times n}$ sú štvorcové matice. Hovoríme, že matice A, B sú *podobné*, ak existuje nesingulárna matica $P \in \mathbb{C}^{n \times n}$ taká, že

$$B = P^{-1}AP.$$

Z tvrdenia 4.1.8 vyplýva, že pre všetky lineárne zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ na konečnorozmernom vektorovom priestore $\mathcal{V}$ a všetky bázy $\mathcal{B}_1, \mathcal{B}_2$ priestoru $\mathcal{V}$ sú matice $A_\varphi[\mathcal{B}_1]$ a $A_\varphi[\mathcal{B}_2]$ navzájom podobné – vieme totiž, že $P[\mathcal{B}_1 \leftarrow \mathcal{B}_2]^{-1} = P[\mathcal{B}_2 \leftarrow \mathcal{B}_1]$.

Relácia „byť podobný“ je reláciou ekvivalencie na $\mathbb{C}^{n \times n}$. Dôkaz tejto skutočnosti prenechávame čitateľovi ako jednoduché – avšak dôležité – cvičenie.

Ako už bolo spomenuté, v tejto kapitole sa úsilie sústreďuje na nájdenie metód umožňujúcich vyjadriť danú maticu A – resp. príslušné lineárne zobrazenie φ_A – „čo možno najjednoduchším“ spôsobom. Presnejšie pôjde o nájdenie „čo možno najjednoduchšej“ matice B podobnej matici A . Jedna z najjednoduchších netriviálnych tried matíc je pritom tvorená diagonálnymi maticami. Úlohu teda bude možné považovať za splnenú, keď zistíme, že je matica A podobná nejakej diagonálnej matici – takéto matice nazveme *diagonalizovateľnými*. Neskôr uvidíme, že hoci nie všetky matice sú diagonalizovateľné, sú všetky matice podobné matici, ktorá je „veľmi blízka“ diagonálnej. Toto pozorovanie však učiníme až v oddiele 4.6; zatiaľ sa obmedzíme na dôkaz charakterizácie triedy diagonalizovateľných matíc, podľa ktorej je matica diagonalizovateľná práve vtedy, keď existuje báza priestoru $\mathbb{C}^n$ zložená z jej vlastných vektorov.

Definícia 4.1.10. Nech $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$ je štvorcová matica. Matica A je *diagonalizovateľná*, ak je podobná diagonálnej matici.

Veta 4.1.11. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$. Matica A je diagonalizovateľná práve vtedy, keď existuje n lineárne nezávislých vlastných vektorov $\mathbf{v}_1, \dots, \mathbf{v}_n$ matice A . V takom prípade navyše*

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

kde λ_k je pre $k = 1, \dots, n$ vlastným číslom prislúchajúcim k vlastnému vektoru $\mathbf{v}_k$ a P je štvorcová matica

$$P = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.$$

Dôkaz. Predpokladajme najprv, že je matica A diagonalizovateľná. Potom existuje nesingulárna matica

$$R = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

a čísla $\kappa_1, \dots, \kappa_n \in \mathbb{C}$ také, že

$$R^{-1}AR = \begin{pmatrix} \kappa_1 & 0 & \dots & 0 \\ 0 & \kappa_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \kappa_n \end{pmatrix}.$$

Maticu R možno v duchu poznámky 4.1.7 interpretovať ako maticu prechodu $P[\mathcal{B}_{st}(\mathbb{C}^n) \leftarrow \mathcal{B}]$ od bázy $\mathcal{B} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ k štandardnej báze priestoru $\mathbb{C}^n$. Pre lineárne zobrazenie φ_A prislúchajúce k matici A tak z tvrdenia 4.1.8 dostávame

$$A_{\varphi_A}[\mathcal{B}] = \begin{pmatrix} \kappa_1 & 0 & \dots & 0 \\ 0 & \kappa_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \kappa_n \end{pmatrix},$$

z čoho vyplýva, že pre $k = 1, \dots, n$ je $\varphi_A(\mathbf{x}_k) = \kappa_k \mathbf{x}_k$, a teda aj $A\mathbf{x}_k = \kappa_k \mathbf{x}_k$. Vektory $\mathbf{x}_1, \dots, \mathbf{x}_n$ sú teda vlastnými vektormi matice A s prislúchajúcimi vlastnými číslami $\kappa_1, \dots, \kappa_n$; na zavŕšenie dôkazu implikácie „zľava doprava“ teda stačí pre $k = 1, \dots, n$ položiť $\mathbf{v}_k := \mathbf{x}_k$.

Za účelom dôkazu opačnej implikácie predpokladajme, že $\mathbf{v}_1, \dots, \mathbf{v}_n$ sú lineárne nezávislé vlastné vektory matice A prislúchajúce k vlastným číslam $\lambda_1, \dots, \lambda_n$ a opäť uvažujme lineárne zobrazenie φ_A prislúchajúce k matici A . Pre $k = 1, \dots, n$ potom $A\mathbf{v}_k = \lambda_k \mathbf{v}_k$ a v dôsledku toho aj $\varphi_A(\mathbf{v}_k) = \lambda_k \mathbf{v}_k$. Ak teraz označíme ako $\mathcal{B}$ bázu $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, nutne

$$A_{\varphi_A}[\mathcal{B}] = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

V duchu poznámky 4.1.7 možno maticu P interpretovať ako maticu $P[\mathcal{B}_{st}(\mathbb{C}^n) \leftarrow \mathcal{B}]$ prechodu od bázy $\mathcal{B}$ k štandardnej báze, pričom triviálne $A_{\varphi_A}[\mathcal{B}_{st}(\mathbb{C}^n)] = A$. Z tvrdenia 4.1.8 teda dostávame

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = A_{\varphi_A}[\mathcal{B}] = P[\mathcal{B} \leftarrow \mathcal{B}_{st}(\mathbb{C}^n)] A P[\mathcal{B}_{st}(\mathbb{C}^n) \leftarrow \mathcal{B}] = P^{-1}AP.$$

Tým je veta dokázaná. □

Definovali sme vlastné čísla a vlastné vektory štvorcovej matice, no zatiaľ nemáme k dispozícii žiaden spôsob, ako pre danú maticu A jej vlastné čísla a vektory aj skutočne nájsť. Tento nedostatok teraz napravíme – ukážeme, že vlastné čísla matice A zodpovedajú koreňom polynómu, ktorý nazveme *charakteristickým polynómom* matice A . Vlastné vektory prislúchajúce k danému vlastnému číslu už potom bude možné vypočítať riešením homogénnej sústavy lineárnych rovníc.

Definícia 4.1.12. Nech $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ je štvorcová matica. *Charakteristickým polynómom* matice A nazveme polynóm

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z - a_{1,1} & -a_{1,2} & \dots & -a_{1,n} \\ -a_{2,1} & z - a_{2,2} & \dots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \dots & z - a_{n,n} \end{pmatrix}.$$

Charakteristickou rovnicou matice A nazývame rovnicu $\text{ch}_A(z) = 0$ o neznámej z . Nasledujúca veta hovorí, že λ je vlastné číslo matice A práve vtedy, keď vyhovuje jej charakteristickej rovnici.

Veta 4.1.13. *Nech $n \in \mathbb{N}$, $A \in \mathbb{C}^{n \times n}$ je štvorcová matica a $\lambda \in \mathbb{C}$. Potom je λ vlastným číslom matice A práve vtedy, keď*

$$\text{ch}_A(\lambda) = \det(\lambda \mathbf{I} - A) = 0.$$

Dôkaz. Číslo λ je vlastným číslom matice A práve vtedy, keď existuje *nenuťový* vektor $\mathbf{v} \in \mathbb{C}^n$ taký, že $A\mathbf{v} = \lambda\mathbf{v}$, alebo ekvivalentne $(\lambda \mathbf{I} - A)\mathbf{v} = \mathbf{0}$. Zisťujeme teda, že λ je vlastným číslom matice A práve vtedy, keď má homogénna sústava lineárnych rovníc $(\lambda \mathbf{I} - A)\mathbf{x} = \mathbf{0}$ aspoň jedno nenuťové riešenie. To nastane práve vtedy, keď je matica $\lambda \mathbf{I} - A$ singulárna; čiže práve vtedy, keď $\det(\lambda \mathbf{I} - A) = 0$. $\square$

Na nájdenie vlastných čísel matice A je teda potrebné nájsť korene charakteristického polynómu $\text{ch}_A(z)$. Nájdenie vlastných vektorov prislúchajúcich k vlastnému číslu λ je už potom jednoduchou záležitosťou: stačí vyriešiť homogénnu sústavu lineárnych rovníc

$$(\lambda \mathbf{I} - A)\mathbf{x} = \mathbf{0} \quad \text{alebo ekvivalentne} \quad (A - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}.$$

V nasledujúcom tvrdení dokážeme, že podobné matice majú rovnaké charakteristické polynómy. V dôsledku toho majú aj rovnaké vlastné čísla.

Tvrdenie 4.1.14. *Nech $n \in \mathbb{N}$ a $A, B \in \mathbb{C}^{n \times n}$ sú podobné štvorcové matice. Potom $\text{ch}_A(z) = \text{ch}_B(z)$.*

Dôkaz. Nech $P \in \mathbb{C}^{n \times n}$ je nesingulárna matica taká, že $B = P^{-1}AP$. S využitím známej skutočnosti, podľa ktorej je determinant súčinu matíc nad ľubovoľným poľom – a teda aj nad podielovým poľom $\mathbb{C}(z)$ oboru integrity $\mathbb{C}[z]$ – rovný súčinu determinantov, dostávame

$$\begin{aligned} \text{ch}_B(z) &= \det(z\mathbf{I} - B) = \det(zP^{-1}P - P^{-1}AP) = \det(P^{-1}zP - P^{-1}AP) = \\ &= \det(P^{-1}(z\mathbf{I})P - P^{-1}AP) = \det(P^{-1}(z\mathbf{I} - A)P) = \det(P^{-1}) \det(z\mathbf{I} - A) \det(P) = \\ &= \det(P^{-1}) \det(P) \text{ch}_A(z) = \det(P^{-1}P) \text{ch}_A(z) = \det(\mathbf{I}) \text{ch}_A(z) = \text{ch}_A(z), \end{aligned}$$

čo bolo treba dokázať. $\square$

Vďaka práve dokázanému tvrdeniu môžeme rozšíriť pojem charakteristického polynómu, doposiaľ definovaný iba pre matice, aj na lineárne zobrazenia na konečnorozmerných vektorových priestoroch. Pre dané zobrazenie φ sú všetky jeho matice vzhľadom na rôzne bázy podobné – vďaka tvrdeniu 4.1.14 preto majú aj rovnaký charakteristický polynóm. Charakteristický polynóm zobrazenia φ tak možno definovať ako charakteristický polynóm ľubovoľnej z týchto matíc.

Definícia 4.1.15. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Charakteristický polynóm zobrazenia φ je polynóm*

$$\text{ch}_\varphi(z) = \text{ch}_{A_\varphi[\mathcal{B}]}(z),$$

kde $\mathcal{B}$ je ľubovoľná báza priestoru $\mathcal{V}$.

V tejto súvislosti ešte zostáva dokázať relatívne očakávateľné tvrdenie, podľa ktorého sú korene charakteristického polynómu $\text{ch}_\varphi(z)$ nielen vlastnými číslami všetkých matíc lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ vzhľadom na rôzne bázy priestoru $\mathcal{V}$, ale aj vlastnými číslami zobrazenia φ samotného. Z toho okrem iného vyplýva, že zobrazenie φ má rovnaké vlastné čísla ako všetky (navzájom podobné) matice zobrazenia φ .

Dôkaz tohto tvrdenia bude formalizáciou veľmi jednoduchej myšlienky: prenášobenie vektora $\mathbf{v}$ z priestoru $\mathcal{V}$ konštantou λ je pre ľubovoľnú n -prvkovú bázu $\mathcal{B}$ priestoru $\mathcal{V}$ „ekvivalentné“ prenášobeniu vektora jeho súradníc $\mathbf{v}[\mathcal{B}] \in \mathbb{C}^n$ taktiež konštantou λ .

Tvrdenie 4.1.16. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie a $\lambda \in \mathbb{C}$. Potom je λ vlastné číslo zobrazenia φ práve vtedy, keď $\text{ch}_\varphi(\lambda) = 0$.*

Dôkaz. Nech $\mathcal{B} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ je ľubovoľná báza priestoru $\mathcal{V}$. Zjavne stačí ukázať, že λ je vlastným číslom zobrazenia φ práve vtedy, keď je vlastným číslom matice

$$A_\varphi[\mathcal{B}] = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \varphi(\mathbf{u}_1)[\mathcal{B}] & \varphi(\mathbf{u}_2)[\mathcal{B}] & \dots & \varphi(\mathbf{u}_n)[\mathcal{B}] \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.$$

Číslo λ je vlastným číslom matice $A_\varphi[\mathcal{B}]$ práve vtedy, keď existuje aspoň jeden nenulový vektor $\mathbf{v} = (a_1, \dots, a_n)^T \in \mathbb{C}^n$ taký, že

$$A_\varphi[\mathcal{B}]\mathbf{v} = \lambda\mathbf{v}.$$

Z tvaru matice $A_\varphi[\mathcal{B}]$ vidieť, že táto situácia nastáva práve vtedy, keď

$$a_1\varphi(\mathbf{u}_1)[\mathcal{B}] + \dots + a_n\varphi(\mathbf{u}_n)[\mathcal{B}] = \lambda\mathbf{v}. \quad (4.2)$$

Zrejme ale

$$a_1\varphi(\mathbf{u}_1)[\mathcal{B}] + \dots + a_n\varphi(\mathbf{u}_n)[\mathcal{B}] = (a_1\varphi(\mathbf{u}_1) + \dots + a_n\varphi(\mathbf{u}_n))[\mathcal{B}] = \varphi(a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n)[\mathcal{B}],$$

čo znamená, že rovnosť (4.2) je ekvivalentná rovnosti

$$\varphi(a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n)[\mathcal{B}] = \lambda(a_1, \dots, a_n)^T,$$

ktorá je splnená práve vtedy, keď

$$\varphi(a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n) = \lambda(a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n),$$

čiže keď λ je vlastným číslom zobrazenia φ prislúchajúcim k vektoru $a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n$. $\square$

Pripomeňme si, že *lineárny súčet* podpriestorov $\mathcal{S}_1, \dots, \mathcal{S}_n$ vektorového priestoru $\mathcal{V}$ je vektorový priestor

$$\mathcal{S}_1 + \dots + \mathcal{S}_n = \{\mathbf{x}_1 + \dots + \mathbf{x}_n \mid \mathbf{x}_1 \in \mathcal{S}_1, \dots, \mathbf{x}_n \in \mathcal{S}_n\},$$

ktorý je očividne najmenším podpriestorom priestoru $\mathcal{V}$ obsahujúcim všetky podpriestory $\mathcal{S}_1, \dots, \mathcal{S}_n$. Takéto súčty podpriestorov sú zaujímavé najmä vtedy, keď každý vektor $\mathbf{x} \in \mathcal{S}_1 + \dots + \mathcal{S}_n$ možno *jednoznačne* vyjadriť ako súčet $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_n$, kde pre $k = 1, \dots, n$ je $\mathbf{x}_k \in \mathcal{S}_k$. V takom prípade hovoríme o *priamom súčte* podpriestorov a namiesto $\mathcal{S}_1 + \dots + \mathcal{S}_n$ píšeme $\bigoplus_{k=1}^n \mathcal{S}_k = \mathcal{S}_1 \oplus \dots \oplus \mathcal{S}_n$.

Tvrdenie 4.1.17. *Nech $\mathcal{V}$ je vektorový priestor nad $\mathbb{C}$, nech $n \geq 2$ je prirodzené číslo a $\mathcal{S}_1, \dots, \mathcal{S}_n$ sú podpriestory priestoru $\mathcal{V}$. Potom je súčet $\mathcal{S}_1 + \dots + \mathcal{S}_n$ priamy práve vtedy, keď pre $k = 1, \dots, n$ je*

$$\mathcal{S}_k \cap (\mathcal{S}_1 + \dots + \mathcal{S}_{k-1} + \mathcal{S}_{k+1} + \dots + \mathcal{S}_n) = \{\mathbf{0}\}.$$

Špeciálne súčet $\mathcal{S} + \mathcal{T}$ dvojice podpriestorov $\mathcal{S}, \mathcal{T}$ priestoru $\mathcal{V}$ je priamy práve vtedy, keď $\mathcal{S} \cap \mathcal{T} = \{\mathbf{0}\}$.

Dôkaz. Ak $\mathbf{v} \in \mathcal{S}_k \cap (\mathcal{S}_1 + \dots + \mathcal{S}_{k-1} + \mathcal{S}_{k+1} + \dots + \mathcal{S}_n)$ pre nejaké $k \in [n]$, možno tento vektor vyjadriť ako $\mathbf{v} = \mathbf{x}_1 + \dots + \mathbf{x}_n$, kde pre $j = 1, \dots, n$ je $\mathbf{x}_j \in \mathcal{S}_j$ a $\mathbf{x}_k = \mathbf{v}$, aj ako $\mathbf{v} = \mathbf{y}_1 + \dots + \mathbf{y}_n$, kde pre $j = 1, \dots, n$ je $\mathbf{y}_j \in \mathcal{S}_j$ a $\mathbf{y}_k = \mathbf{0}$. Aby bol súčet $\mathcal{S}_1 + \dots + \mathcal{S}_n$ priamy, musí byť $\mathbf{x}_j = \mathbf{y}_j$ pre $j = 1, \dots, n$, a teda aj $\mathbf{v} = \mathbf{x}_k = \mathbf{y}_k = \mathbf{0}$.

Predpokladajme teraz, že pre $k = 1, \dots, n$ je $\mathcal{S}_k \cap (\mathcal{S}_1 + \dots + \mathcal{S}_{k-1} + \mathcal{S}_{k+1} + \dots + \mathcal{S}_n) = \{\mathbf{0}\}$ a $\mathbf{v} \in \mathcal{S}_1 + \dots + \mathcal{S}_n$ je vyjadriteľný ako $\mathbf{v} = \mathbf{x}_1 + \dots + \mathbf{x}_n = \mathbf{y}_1 + \dots + \mathbf{y}_n$, kde pre $j = 1, \dots, n$ je $\mathbf{x}_j, \mathbf{y}_j \in \mathcal{S}_j$. Keby existovalo $k \in [n]$ také, že $\mathbf{x}_k \neq \mathbf{y}_k$, bol by vektor

$$\mathbf{x}_k - \mathbf{y}_k = (\mathbf{y}_1 - \mathbf{x}_1) + \dots + (\mathbf{y}_{k-1} - \mathbf{x}_{k-1}) + (\mathbf{y}_{k+1} - \mathbf{x}_{k+1}) + \dots + (\mathbf{y}_n - \mathbf{x}_n)$$

nenulový a súčasne by patrilo do priestoru $\mathcal{S}_k \cap (\mathcal{S}_1 + \dots + \mathcal{S}_{k-1} + \mathcal{S}_{k+1} + \dots + \mathcal{S}_n)$. Taký vektor ale neexistuje – nutne teda $\mathbf{x}_k = \mathbf{y}_k$ pre $k = 1, \dots, n$ a súčet $\mathcal{S}_1 + \dots + \mathcal{S}_n$ je priamy. $\square$

Ak je teraz $\mathcal{S}_1 \oplus \dots \oplus \mathcal{S}_n$ priamy súčet podpriestorov $\mathcal{S}_1, \dots, \mathcal{S}_n$ priestoru $\mathcal{V}$ a pre $k = 1, \dots, n$ je $(\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,s_k})$ bázou priestoru $\mathcal{S}_k$, môže byť jedna z báz vektorového priestoru $\mathcal{S}_1 \oplus \dots \oplus \mathcal{S}_n$ evidentne daná ako $(\mathbf{x}_{1,1}, \dots, \mathbf{x}_{1,s_1}, \dots, \mathbf{x}_{n,1}, \dots, \mathbf{x}_{n,s_n})$. Preto tiež $\dim(\mathcal{S}_1 \oplus \dots \oplus \mathcal{S}_n) = \dim(\mathcal{S}_1) + \dots + \dim(\mathcal{S}_n)$.

Podpriestor $\mathcal{S}$ vektorového priestoru $\mathcal{V}$ je *invariantný* vzhľadom na lineárne zobrazenie $\varphi: \mathcal{V} \rightarrow \mathcal{V}$, ak pre všetky $\mathbf{x} \in \mathcal{S}$ je aj $\varphi(\mathbf{x}) \in \mathcal{S}$. Nech je teraz $\mathcal{V}$ konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Ak $\mathcal{S}$ je podpriestor priestoru $\mathcal{V}$ invariantný vzhľadom na φ , označíme $\varphi|_{\mathcal{S}}$ *zúženie* zobrazenia φ nad podpriestor $\mathcal{S}$ – to je lineárne zobrazenie $\varphi|_{\mathcal{S}}: \mathcal{S} \rightarrow \mathcal{S}$ definované pre všetky $\mathbf{x} \in \mathcal{S}$ ako $\varphi|_{\mathcal{S}}(\mathbf{x}) = \varphi(\mathbf{x})$. Uvažujme teraz ľubovoľnú bázu $\mathcal{B} = (\mathbf{x}_1, \dots, \mathbf{x}_s, \mathbf{y}_1, \dots, \mathbf{y}_t)$ priestoru $\mathcal{V}$ takú, že $\mathcal{B}' = (\mathbf{x}_1, \dots, \mathbf{x}_s)$ je bázou $\mathcal{S}$. Maticu $A_{\varphi}[\mathcal{B}]$ potom možno vyjadriť v blokovom tvare

$$A_{\varphi}[\mathcal{B}] = \begin{pmatrix} A_{\varphi|_{\mathcal{S}}}[\mathcal{B}'] & A_{1,2} \\ \mathbf{0} & A_{2,2} \end{pmatrix} \quad (4.3)$$

pre nejaké matice $A_{1,2} \in \mathbb{C}^{s \times t}$ a $A_{2,2} \in \mathbb{C}^{t \times t}$. Ak navyše $\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$, kde $\mathcal{T}$ je tiež podpriestor priestoru $\mathcal{V}$ invariantný vzhľadom na φ s bázou $\mathcal{B}'' = (\mathbf{y}_1, \dots, \mathbf{y}_t)$, tak

$$A_{\varphi}[\mathcal{B}] = \begin{pmatrix} A_{\varphi|_{\mathcal{S}}}[\mathcal{B}'] & \mathbf{0} \\ \mathbf{0} & A_{\varphi|_{\mathcal{T}}}[\mathcal{B}''] \end{pmatrix}. \quad (4.4)$$

Nasledujúce tvrdenie je jednoduchým pozorovaním o charakteristických polynómoch týchto matíc.

Tvrdenie 4.1.18. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie a $\mathcal{S}$ je podpriestor priestoru $\mathcal{V}$ invariantný vzhľadom na φ . Potom existuje polynóm $p(z)$ taký, že*

$$\text{ch}_{\varphi}(z) = \text{ch}_{\varphi|_{\mathcal{S}}}(z) \cdot p(z).$$

Ak navyše $\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$, kde $\mathcal{S}, \mathcal{T}$ sú podpriestory $\mathcal{V}$ invariantné vzhľadom na φ , tak

$$\text{ch}_{\varphi}(z) = \text{ch}_{\varphi|_{\mathcal{S}}}(z) \cdot \text{ch}_{\varphi|_{\mathcal{T}}}(z).$$

Dôkaz. Nech $\mathcal{B} = (\mathbf{x}_1, \dots, \mathbf{x}_s, \mathbf{y}_1, \dots, \mathbf{y}_t)$ je báza priestoru $\mathcal{V}$ taká, že $\mathcal{B}' = (\mathbf{x}_1, \dots, \mathbf{x}_s)$ je bázou $\mathcal{S}$. Matica $A_{\varphi}[\mathcal{B}]$ potom nadobúda blokový rozklad (4.3). Z toho dostávame

$$\begin{aligned} \text{ch}_{\varphi}(z) &= \text{ch}_{A_{\varphi}[\mathcal{B}]}(z) = \det \begin{pmatrix} z\mathbf{I} - A_{\varphi|_{\mathcal{S}}}[\mathcal{B}'] & -A_{1,2} \\ \mathbf{0} & z\mathbf{I} - A_{2,2} \end{pmatrix} = \det(z\mathbf{I} - A_{\varphi|_{\mathcal{S}}}[\mathcal{B}']) \cdot \det(z\mathbf{I} - A_{2,2}) = \\ &= \text{ch}_{A_{\varphi|_{\mathcal{S}}}[\mathcal{B}']}(z) \cdot \det(z\mathbf{I} - A_{2,2}) = \text{ch}_{\varphi|_{\mathcal{S}}}(z) \cdot \det(z\mathbf{I} - A_{2,2}). \end{aligned}$$

Na dôkaz prvej časti tvrdenia teda stačí položiť $p(z) := \det(z\mathbf{I} - A_{2,2})$. Na dôkaz druhej časti tvrdenia si stačí uvedomiť, že ak $\mathcal{V} = \mathcal{S} \oplus \mathcal{T}$, kde $\mathcal{S}, \mathcal{T}$ sú podpriestory $\mathcal{V}$ invariantné vzhľadom na φ , nadobúda $A_{\varphi}[\mathcal{B}]$ blokový rozklad (4.4), z čoho pre bázu $\mathcal{B}'' = (\mathbf{y}_1, \dots, \mathbf{y}_t)$ priestoru $\mathcal{T}$ dostávame

$$\begin{aligned} \text{ch}_{\varphi}(z) &= \text{ch}_{A_{\varphi}[\mathcal{B}]}(z) = \det \begin{pmatrix} z\mathbf{I} - A_{\varphi|_{\mathcal{S}}}[\mathcal{B}'] & \mathbf{0} \\ \mathbf{0} & z\mathbf{I} - A_{\varphi|_{\mathcal{T}}}[\mathcal{B}''] \end{pmatrix} = \\ &= \det(z\mathbf{I} - A_{\varphi|_{\mathcal{S}}}[\mathcal{B}']) \cdot \det(z\mathbf{I} - A_{\varphi|_{\mathcal{T}}}[\mathcal{B}'']) = \\ &= \text{ch}_{A_{\varphi|_{\mathcal{S}}}[\mathcal{B}']}(z) \cdot \text{ch}_{A_{\varphi|_{\mathcal{T}}}[\mathcal{B}'']}(z) = \text{ch}_{\varphi|_{\mathcal{S}}}(z) \cdot \text{ch}_{\varphi|_{\mathcal{T}}}(z). \end{aligned}$$

Tým je tvrdenie dokázané. □

Tento oddiel zakončíme niekoľkými príkladmi, v ktorých zakaždým nájdeme všetky vlastné čísla danej štvorcovej matice a ku každému vlastnému číslu nájdeme príslušný vlastný priestor (čiže v podstate všetky vlastné vektory). V rámci príprav na naše neskoršie úvahy o súvisi vlastných čísel s grafmi môže byť užitočným cvičením interpretovať matice z nasledujúcich príkladov ako matice ohodnotených grafov nad $\mathbb{N} \subseteq \mathbb{C}$.

Príklad 4.1.19. Uvažujme maticu

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

Charakteristickým polynómom matice A je

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z-1 & -1 \\ -1 & z \end{pmatrix} = z^2 - z - 1.$$

Vlastnými číslami matice A sú korene polynómu $\text{ch}_A(z)$ dané ako

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} \quad \text{a} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

Riešením homogénnych sústav lineárnych rovníc sa následne možno presvedčiť, že vlastnými priestormi prislúchajúcimi týmto vlastným číslam sú

$$E(A, \lambda_1) = \text{Ker}(\lambda_1\mathbf{I} - A) = \{c(\lambda_1, 1)^T \mid c \in \mathbb{C}\}$$

a

$$E(A, \lambda_2) = \text{Ker}(\lambda_2\mathbf{I} - A) = \{c(\lambda_2, 1)^T \mid c \in \mathbb{C}\}.$$

Matica A má teda dve vlastné čísla, ktoré sú jednoduchými koreňmi jej charakteristického polynómu. Ku každému vlastnému číslu prislúcha jednorozmerný vlastný priestor.

Príklad 4.1.20. Uvažujme teraz maticu

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

Keďže je matica A diagonálna, ľahko vidieť, že

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z-2 & 0 & 0 \\ 0 & z-2 & 0 \\ 0 & 0 & z-2 \end{pmatrix} = (z-2)^3.$$

Jediným vlastným číslom matice A je teda trojnásobný koreň polynómu $\text{ch}_A(z)$:

$$\lambda = \lambda_1 = \lambda_2 = \lambda_3 = 2.$$

Keďže je matica $\lambda\mathbf{I} - A$ nulová, ľahko tiež vidieť, že vlastný priestor prislúchajúci tomuto jedinému vlastnému číslu je daný ako

$$E(A, \lambda) = \text{Ker}(\lambda\mathbf{I} - A) = \mathbb{C}^3.$$

Matica A má teda jediné vlastné číslo, ktoré je trojnásobným koreňom jej charakteristického polynómu. Zodpovedá mu vlastný priestor dimenzie 3.

Príklad 4.1.21. Nech teraz

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

Matica A je horná trojuholníková; bez problémov teda opäť vidieť, že

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z-2 & -1 & 0 \\ 0 & z-2 & 0 \\ 0 & 0 & z-2 \end{pmatrix} = (z-2)^3.$$

Jediným vlastným číslom matice A je tak znova

$$\lambda = \lambda_1 = \lambda_2 = \lambda_3 = 2.$$

Riešením systému lineárnych rovníc $(\lambda \mathbf{I} - A)\mathbf{x} = \mathbf{0}$ sa však ľahko možno presvedčiť o tom, že

$$E(A, \lambda) = \text{Ker}(\lambda \mathbf{I} - A) = \{a(1, 0, 0)^T + b(0, 0, 1)^T \mid a, b \in \mathbb{C}\}.$$

Matica A má teda opäť jediné vlastné číslo, ktoré je trojnásobným koreňom charakteristického polynómu; tentokrát mu ale prislúcha vlastný priestor dimenzie 2.

Príklad 4.1.22. Uvažujme nakoniec maticu

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

kde rovnako ako v predchádzajúcom príklade zisťujeme, že jej jediným vlastným číslom je

$$\lambda = \lambda_1 = \lambda_2 = \lambda_3 = 2.$$

Riešením systému rovníc $(\lambda \mathbf{I} - A)\mathbf{x} = \mathbf{0}$ ďalej možno nájsť príslušný vlastný priestor:

$$E(A, \lambda) = \text{Ker}(\lambda \mathbf{I} - A) = \{c(1, 0, 0)^T \mid c \in \mathbb{C}\}.$$

Matica A teda má jediné vlastné číslo; to je opäť trojnásobným koreňom charakteristického polynómu, no tentokrát mu zodpovedá iba jednorozmerný vlastný priestor.

4.2 Spektrum, algebraická násobnosť a geometrická násobnosť

Pokračujúc v skúmaní načatom v rámci predchádzajúceho oddielu sa teraz bližšie zameriame na vlastnosti matíc, na ktoré sme sa sústredili už v príkladoch 4.1.19 až 4.1.22 – teda na množinu ich vlastných čísel, na násobnosti jednotlivých vlastných čísel ako koreňov charakteristického polynómu a napokon na dimenzie príslušných vlastných priestorov. Ukazuje sa, že tieto „veličiny“ sú natoľko významné, že si zaslúžia vlastné pomenovania: budeme tak hovoriť o *spektre* matice, o *algebraickej násobnosti* vlastného čísla a o *geometrickej násobnosti* vlastného čísla. Podobne ako v prípade pojmov zavedených v predchádzajúcom oddiele, možno aj definície týchto sformulovať v o niečo všeobecnejšom kontexte lineárnych zobrazení na konečnorozmerných vektorových priestoroch nad poľom komplexných čísel.

Definícia 4.2.1. Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Potom:

- Spektrum* zobrazenia φ nazveme množinu $\sigma(\varphi)$ všetkých vlastných čísel zobrazenia φ .
- Algebraickou násobnosťou* $\alpha(\varphi, \lambda)$ vlastného čísla $\lambda \in \sigma(\varphi)$ zobrazenia φ nazveme najväčšie prirodzené číslo m také, že polynóm $(z - \lambda)^m$ delí $\text{ch}_\varphi(z)$.²
- Geometrickou násobnosťou* $\gamma(\varphi, \lambda)$ vlastného čísla $\lambda \in \sigma(\varphi)$ zobrazenia φ nazveme dimenziu vlastného priestoru $E(\varphi, \lambda)$.

Nasledujúca definícia zavádza pojmy, ktoré sme práve definovali pre lineárne zobrazenia na konečnorozmerných vektorových priestoroch, v kontexte štvorcových matíc.

²Ide teda o násobnosť koreňa λ charakteristického polynómu $\text{ch}_\varphi(z)$.

Definícia 4.2.2. Nech $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$ je štvorcová matica. Potom:

- a) *Spektrom* matice A nazveme množinu $\sigma(A)$ všetkých vlastných čísel matice A .
- b) *Algebraickou násobnosťou* $\alpha(A, \lambda)$ vlastného čísla $\lambda \in \sigma(A)$ matice A nazveme najväčšie prirodzené číslo m také, že polynóm $(z - \lambda)^m$ delí $\text{ch}_A(z)$.
- c) *Geometrickou násobnosťou* $\gamma(A, \lambda)$ vlastného čísla $\lambda \in \sigma(A)$ matice A nazveme dimenziu vlastného priestoru $E(A, \lambda)$.

Nasledujúce tvrdenie dáva predchádzajúce definície do vzájomného súvisu a čitateľ by ho mal v ideálnom prípade považovať za zrejmé.

Tvrdenie 4.2.3. Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie a $\mathcal{B} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ je ľubovoľná báza priestoru $\mathcal{V}$. Potom $\sigma(\varphi) = \sigma(A_\varphi[\mathcal{B}])$ a pre všetky $\lambda \in \sigma(\varphi) = \sigma(A_\varphi[\mathcal{B}])$ je $\alpha(\varphi, \lambda) = \alpha(A_\varphi[\mathcal{B}], \lambda)$ a $\gamma(\varphi, \lambda) = \gamma(A_\varphi[\mathcal{B}], \lambda)$.

Dôkaz. Rovnosť spektier a algebraických násobností jednotlivých vlastných čísel je priamym dôsledkom vety 4.1.13, definície 4.1.15 a tvrdenia 4.1.16. Zostáva pre každé $\lambda \in \sigma(\varphi) = \sigma(A_\varphi[\mathcal{B}])$ dokázať rovnosť geometrických násobností $\gamma(\varphi, \lambda)$ a $\gamma(A_\varphi[\mathcal{B}], \lambda)$, čiže dimenzií priestorov $E(\varphi, \lambda)$ a $E(A_\varphi[\mathcal{B}], \lambda)$.

Nech $\lambda \in \sigma(\varphi) = \sigma(A_\varphi[\mathcal{B}])$ je vlastné číslo. Rovnako ako v dôkaze tvrdenia 4.1.16 zisťujeme, že $\mathbf{v} = (a_1, \dots, a_n)^T \in \mathbb{C}^n$ je vlastným vektorom matice $A_\varphi[\mathcal{B}]$ prislúchajúcim k vlastnému číslu λ práve vtedy, keď

$$a_1 \mathbf{u}_1 + \dots + a_n \mathbf{u}_n \in \mathcal{V}$$

je vlastným vektorom zobrazenia φ prislúchajúcim k vlastnému číslu λ . Teda

$$E(\varphi, \lambda) = \{a_1 \mathbf{u}_1 + \dots + a_n \mathbf{u}_n \mid (a_1, \dots, a_n)^T \in E(A_\varphi[\mathcal{B}], \lambda)\}.$$

Pretože sú ale vektory $\mathbf{u}_1, \dots, \mathbf{u}_n$ lineárne nezávislé, musí tiež byť $\dim(E(\varphi, \lambda)) = \dim(E(A_\varphi[\mathcal{B}], \lambda))$ a v dôsledku toho aj $\gamma(\varphi, \lambda) = \gamma(A_\varphi[\mathcal{B}], \lambda)$. $\square$

Dôsledok 4.2.4. Nech $n \in \mathbb{N}$ a $A, B \in \mathbb{C}^{n \times n}$ sú podobné štvorcové matice. Potom $\sigma(A) = \sigma(B)$ a pre všetky $\lambda \in \sigma(A) = \sigma(B)$ je $\alpha(A, \lambda) = \alpha(B, \lambda)$ a $\gamma(A, \lambda) = \gamma(B, \lambda)$.

Dôkaz. Vyplýva bezprostredne z tvrdenia 4.2.3 a zo skutočnosti, že matice A a B sú, keďže sú podobné, maticami toho istého lineárneho zobrazenia na $\mathbb{C}^n$, prípadne však nad rôznymi bázami. $\square$

Niekoľko tvrdení o spektrách, algebraických násobnostiach a geometrických násobnostiach teraz vyslovíme iba v reči lineárnych zobrazení; lineárne zobrazenie $\varphi_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ prislúchajúce k matici $A \in \mathbb{C}^{n \times n}$ skrz predpis $\varphi_A: \mathbf{x} \mapsto A\mathbf{x}$ bude spolu s tvrdením 4.2.3 zárukou platnosti analogických tvrdení aj pre štvorcové matice.

V prvom rade si všimnime, že súčet *algebraických* násobností vlastných čísel lineárneho zobrazenia na konečnorozmernom priestore $\mathcal{V}$ je vždy rovný dimenzii priestoru $\mathcal{V}$. V reči matíc toto tvrdenie hovorí, že súčet algebraických násobností vlastných čísel matice $A \in \mathbb{C}^{n \times n}$ je vždy n .

Tvrdenie 4.2.5. Nech n je prirodzené číslo, $\mathcal{V}$ je n -rozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Potom

$$\sum_{\lambda \in \sigma(\varphi)} \alpha(\varphi, \lambda) = n.$$

Dôkaz. Vďaka tvrdeniu 4.1.16 je $\lambda \in \sigma(\varphi)$ práve vtedy, keď je λ koreňom $\text{ch}_\varphi(z)$; násobnosťou tohto koreňa je číslo $\alpha(\varphi, \lambda)$. Charakteristický polynóm $\text{ch}_\varphi(z)$ je ale polynómom stupňa n nad algebraicky uzavretým poľom komplexných čísel, ktorý tak možno jednoznačne rozložiť na koreňové činitele – súčet násobností jeho koreňov teda musí byť n . $\square$

V príkladoch 4.1.21 a 4.1.22 sme videli, že pre *geometrické* násobnosti vlastných čísel obdoba tvrdenia 4.2.5 platiť nemusí. Dokážeme však, že geometrická násobnosť každého vlastného čísla je *zhora* ohraničená jeho algebraickou násobnosťou; súčet geometrických násobností vlastných čísel lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ na n -rozmernom vektorovom priestore $\mathcal{V}$ – prípadne štvorcovej matice $A \in \mathbb{C}^{n \times n}$ – je tak *nanajvýš* rovný n .

Tvrdenie 4.2.6. *Nech n je prirodzené číslo, $\mathcal{V}$ je n -rozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Pre každé $\lambda \in \sigma(\varphi)$ potom $\gamma(\varphi, \lambda) \leq \alpha(\varphi, \lambda)$ a v dôsledku toho*

$$\sum_{\lambda \in \sigma(\varphi)} \gamma(\varphi, \lambda) \leq n.$$

Dôkaz. Nech $\lambda \in \sigma(\varphi)$. Geometrická násobnosť $\gamma(\varphi, \lambda)$ je daná dimenziou vlastného priestoru $E(\varphi, \lambda)$. Vlastný priestor $E(\varphi, \lambda)$ je invariantný vzhľadom na φ ; uvažujme zúženie $\varphi|_{E(\varphi, \lambda)}: E(\varphi, \lambda) \rightarrow E(\varphi, \lambda)$ zobrazenia φ na priestor $E(\varphi, \lambda)$. Podľa tvrdenia 4.1.5 je $E(\varphi, \lambda) = \text{Ker}(\lambda \text{Id}_{\mathcal{V}} - \varphi)$, v dôsledku čoho je $\varphi|_{E(\varphi, \lambda)} = \lambda \text{Id}_{E(\varphi, \lambda)}$. Preto

$$\text{ch}_{\varphi|_{E(\varphi, \lambda)}}(z) = (z - \lambda)^{\gamma(\varphi, \lambda)}.$$

Z tvrdenia 4.1.18 vyplýva, že tento polynóm delí $\text{ch}_{\varphi}(z)$. To je však z definície algebraickej násobnosti možné len vtedy, keď platí nerovnosť $\gamma(\varphi, \lambda) \leq \alpha(\varphi, \lambda)$. Odhad pre súčet geometrických násobností potom dostávame z tvrdenia 4.2.5. $\square$

Nasledujúce dôležité tvrdenie vyjadruje charakteristický polynóm lineárneho zobrazenia pomocou jeho spektra a algebraických násobností jeho vlastných čísel.

Tvrdenie 4.2.7. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Potom*

$$\text{ch}_{\varphi}(z) = \prod_{\lambda \in \sigma(\varphi)} (z - \lambda)^{\alpha(\varphi, \lambda)}.$$

Dôkaz. Z definície 4.2.2 a tvrdenia 4.1.16 vyplýva, že pre nejaké $a \in \mathbb{C}$ musí byť

$$\text{ch}_{\varphi}(z) = a \cdot \prod_{\lambda \in \sigma(\varphi)} (z - \lambda)^{\alpha(\varphi, \lambda)}.$$

Pre ľubovoľnú bázu $\mathcal{B}$ priestoru $\mathcal{V}$ ale $\text{ch}_{\varphi}(z) = \det(z\mathbf{I} - A_{\varphi}[\mathcal{B}])$, z čoho ľahko vidieť, že $a = 1 -$ koeficient pri z^n v polynóme $\det(z\mathbf{I} - A_{\varphi}[\mathcal{B}])$ je totiž zrejme rovný jednej. $\square$

Vráťme sa ešte k diagonalizácii matíc. Vďaka vete 4.1.11 už máme k dispozícii asi najdôležitejšiu charakterizáciu diagonalizovateľných matíc: matica typu $n \times n$ nad poľom $\mathbb{C}$ je diagonalizovateľná práve vtedy, keď pre ňu možno nájsť n lineárne nezávislých vlastných vektorov.

Nasledujúca veta hovorí, že vlastné vektory prislúchajúce po dvoch rôznych vlastným číslam lineárneho zobrazenia – a teda aj matice – sú vždy lineárne nezávislé. Z toho vyplýva, že matica $A \in \mathbb{C}^{n \times n}$ s n po dvoch rôznymi vlastnými číslami je nutne diagonalizovateľná.

Veta 4.2.8. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie, $m \in \mathbb{N}$ a $\lambda_1, \dots, \lambda_m$ po dvoch rôzne vlastné čísla zobrazenia φ . Nech pre $k = 1, \dots, m$ je $\mathbf{v}_k \in \mathcal{V}$ vlastný vektor zobrazenia φ prislúchajúci k λ_k . Potom sú vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ lineárne nezávislé.*

Dôkaz. Sporom: nech sú vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ lineárne závislé. Potom nutne $m \geq 2$. Nech $k \in \{2, \dots, m\}$ je *najmenší* index taký, že pre nejaké nie všetky nulové koeficienty $a_1, \dots, a_{k-1} \in \mathbb{C}$ je

$$\mathbf{v}_k = a_1 \mathbf{v}_1 + \dots + a_{k-1} \mathbf{v}_{k-1}.$$

Takýto index k existuje, pretože vlastné vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ musia byť nenulové. Vektory $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ sú potom istotne lineárne nezávislé. Navyše

$$\varphi(\mathbf{v}_k) = \varphi(a_1\mathbf{v}_1 + \dots + a_{k-1}\mathbf{v}_{k-1}) = a_1\varphi(\mathbf{v}_1) + \dots + a_{k-1}\varphi(\mathbf{v}_{k-1}) = a_1\lambda_1\mathbf{v}_1 + \dots + a_{k-1}\lambda_{k-1}\mathbf{v}_{k-1},$$

ako aj

$$\varphi(\mathbf{v}_k) = \lambda_k\mathbf{v}_k = \lambda_k(a_1\mathbf{v}_1 + \dots + a_{k-1}\mathbf{v}_{k-1}) = a_1\lambda_k\mathbf{v}_1 + \dots + a_{k-1}\lambda_k\mathbf{v}_{k-1}.$$

Preto

$$a_1(\lambda_k - \lambda_1)\mathbf{v}_1 + \dots + a_{k-1}(\lambda_k - \lambda_{k-1})\mathbf{v}_{k-1} = \mathbf{0}. \quad (4.5)$$

Medzi koeficientmi $a_1, \dots, a_{k-1}$ existuje aspoň jeden nenulový; rozdiely $(\lambda_k - \lambda_1), \dots, (\lambda_k - \lambda_{k-1})$ sú vďaka predpokladu po dvoch rôznych vlastných čísel $\lambda_1, \dots, \lambda_m$ nenulové všetky. Existuje preto aspoň jeden index $j \in [k-1]$ taký, že $a_j(\lambda_k - \lambda_j) \neq 0$. Z rovnosti (4.5) preto vyplýva, že vektory $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ sú lineárne závislé – čo je spor s ich lineárnou nezávislosťou pozorovanou vyššie. $\square$

Dôsledok 4.2.9. *Nech n je prirodzené číslo a $A \in \mathbb{C}^{n \times n}$ je matica. Ak $|\sigma(A)| = n$, je matica A diagonalizovateľná.*

Dôkaz. Podmienka $|\sigma(A)| = n$ hovorí, že matica A má n po dvoch rôznych vlastných čísel. K nim vďaka vete 4.2.8 prislúcha n lineárne nezávislých vlastných vektorov. Diagonalizovateľnosť matice A je potom dôsledkom vety 4.1.11. $\square$

Dalším dôsledkom vety 4.2.8 je, že súčet vlastných priestorov $E(\varphi, \lambda)$ lineárneho zobrazenia φ je vždy priamy.

Dôsledok 4.2.10. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie, $m \in \mathbb{N}$ a $\lambda_1, \dots, \lambda_m$ po dvoch rôzne vlastné čísla zobrazenia φ . Súčet vlastných priestorov $E(\varphi, \lambda_1), \dots, E(\varphi, \lambda_m)$ je potom priamy:*

$$E(\varphi, \lambda_1) + \dots + E(\varphi, \lambda_m) = E(\varphi, \lambda_1) \oplus \dots \oplus E(\varphi, \lambda_m) = \bigoplus_{k=1}^m E(\varphi, \lambda_k).$$

Každý vektor

$$\mathbf{x} \in E(\varphi, \lambda_1) + \dots + E(\varphi, \lambda_m)$$

tak možno jediným spôsobom vyjadriť ako $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_m$, kde pre $k = 1, \dots, m$ je $\mathbf{x}_k \in E(\varphi, \lambda_k)$. Ak ďalej pre $k = 1, \dots, m$ je $(\mathbf{v}_{k,1}, \dots, \mathbf{v}_{k,s_k})$ bázou $E(\varphi, \lambda_k)$, je $(\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,s_1}, \dots, \mathbf{v}_{m,1}, \dots, \mathbf{v}_{m,s_m})$ bázou priestoru $\bigoplus_{k=1}^m E(\varphi, \lambda_k)$.

Dôkaz. Pre $m = 0$ a $m = 1$ je tvrdenie triviálne pravdivé. Uvažujme teda $m \geq 2$ a za účelom sporu predpokladajme, že súčet $E(\varphi, \lambda_1) + \dots + E(\varphi, \lambda_m)$ nie je priamy. Podľa tvrdenia 4.1.17 potom existuje $k \in [m]$ také, že pre nejaký vektor $\mathbf{v}_k \neq \mathbf{0}$ je

$$\mathbf{v}_k \in E(\varphi, \lambda_k) \cap (E(\varphi, \lambda_1) + \dots + E(\varphi, \lambda_{k-1}) + E(\varphi, \lambda_{k+1}) + \dots + E(\varphi, \lambda_m)).$$

To znamená, že $\mathbf{v}_k$ je vlastným vektorom lineárneho zobrazenia φ prislúchajúcim k λ_k a súčasne existujú vektory $\mathbf{v}_1 \in E(\varphi, \lambda_1), \dots, \mathbf{v}_{k-1} \in E(\varphi, \lambda_{k-1}), \mathbf{v}_{k+1} \in E(\varphi, \lambda_{k+1}), \dots, \mathbf{v}_m \in E(\varphi, \lambda_m)$ také, že

$$\mathbf{v}_k = \mathbf{v}_1 + \dots + \mathbf{v}_{k-1} + \mathbf{v}_{k+1} + \dots + \mathbf{v}_m.$$

Keďže je ale pre $j = 1, \dots, m$ vektor $\mathbf{v}_j$ v prípade svojej nenulovosti vlastným vektorom prislúchajúcim k vlastnému číslu λ_j a keďže je vektor $\mathbf{v}_k$ nenulový, ide o spor s vetou 4.2.8. $\square$

Poznámka 4.2.11. Predpoklady dôsledku 4.2.10 sú špeciálne splnené vtedy, keď $\lambda_1, \dots, \lambda_m$ sú práve všetky po dvoch rôzne vlastné čísla zo $\sigma(\varphi)$. Korektne tak možno písať napríklad aj $\bigoplus_{\lambda \in \sigma(\varphi)} E(\varphi, \lambda)$.

Na záver tohto oddielu ešte vyslovme vetu poskytujúcu niekoľko ekvivalentných charakterizácií diagonalizovateľných matic.

Veta 4.2.12. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$ je matica. Potom sú nasledujúce podmienky ekvivalentné:*

- (i) *Matica A je diagonalizovateľná.*
- (ii) *Existuje n lineárne nezávislých vlastných vektorov $\mathbf{v}_1, \dots, \mathbf{v}_n$ matice A .*
- (iii) *Platí rovnosť $\mathbb{C}^n = \bigoplus_{\lambda \in \sigma(A)} E(A, \lambda)$.*
- (iv) *Platí rovnosť $\sum_{\lambda \in \sigma(A)} \gamma(A, \lambda) = n$.*
- (v) *Pre každé vlastné číslo $\lambda \in \sigma(A)$ je $\gamma(A, \lambda) = \alpha(A, \lambda)$.*

Dôkaz. Podmienky (i) a (ii) sú ekvivalentné podľa vety 4.1.11. Znenie podmienky (iii) dáva zmysel vďaka dôsledku 4.2.10 (a poznámke 4.2.11). Ekvivalencia podmienok (ii) a (iii) je potom dôsledkom skutočnosti, že $\mathbf{v}$ je vlastným vektorom matice A práve vtedy, keď je nenulovým prvkom $E(A, \lambda)$ pre nejaké $\lambda \in \sigma(A)$, ktoré je navyše určené jednoznačne. Ekvivalencia podmienok (iii) a (iv) vyplýva z nasledujúcich rovností:

$$\dim \left(\bigoplus_{\lambda \in \sigma(A)} E(A, \lambda) \right) = \sum_{\lambda \in \sigma(A)} \dim(E(A, \lambda)) = \sum_{\lambda \in \sigma(A)} \gamma(A, \lambda).$$

Ekvivalencia podmienok (iv) a (v) je napokon dôsledkom tvrdení 4.2.5 a 4.2.6. □

4.3 Ukážka aplikácií diagonalizácie

Naznačíme teraz niektoré možnosti využitia diagonalizácie matic. Matica A je diagonalizovateľná práve vtedy, keď pre nejakú diagonálnu maticu Λ a nesusingulárnu maticu P je $P^{-1}AP = \Lambda$ resp. $A = P\Lambda P^{-1}$. Z vety 4.1.11 navyše vieme, že matica A typu $n \times n$ je diagonalizovateľná práve vtedy, keď existuje n lineárne nezávislých vlastných vektorov $\mathbf{v}_1, \dots, \mathbf{v}_n$ matice A prislúchajúcich k nejakým – nie nutne rôznym – vlastným číslam $\lambda_1, \dots, \lambda_n$; v takom prípade

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

a maticu P možno zvoliť ako

$$P = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.$$

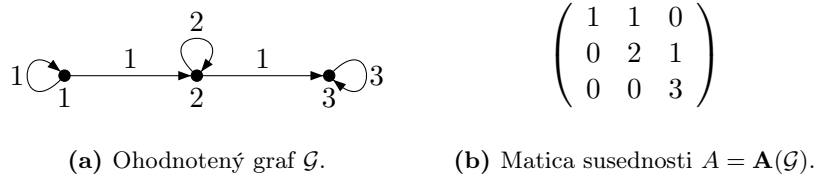
Diagonalizovateľnú maticu A možno okrem iného veľmi jednoducho umocňovať na prirodzený exponent. Pre každé $t \in \mathbb{N}$ totiž

$$A^t = (P\Lambda P^{-1})^t = P\Lambda P^{-1}P\Lambda P^{-1}P \dots P^{-1}P\Lambda P^{-1} = P\Lambda^t P^{-1} = P \begin{pmatrix} \lambda_1^t & 0 & \dots & 0 \\ 0 & \lambda_2^t & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^t \end{pmatrix} P^{-1};$$

nie je teda problém vyjadriť prvky matice A^t v uzavretom tvare. Ako už vieme, mocniny matíc možno prirodzene interpretovať v reči ohodnotených grafov a poukážeme tiež na ich (vcelku priamočiary) súvis so systémami diferenčných rovníc. Na týchto prepojeniach budú založené nasledujúce príklady.

Je nutné podotknúť, že nie všetky matice sú diagonalizovateľné.³ Metódy použité v nasledujúcich príkladoch teda ešte nebudú aplikovateľné univerzálne; obdobný postup bude možné použiť práve vtedy, keď je príslušná matica diagonalizovateľná.⁴ Neskôr ich rozšírime aj na nediagonalizovateľné matice – urobíme tak v kapitolách 5 a 6. Teoretický základ pre tento všeobecný prípad nám však poskytnú až pojmy zovšeobecneného vlastného vektora a Jordanovho kanonického tvaru matice, ktoré zavedieme a preskúame v oddieloch 4.5 a 4.6.

Príklad 4.3.1. Na obrázku 4.1 je znázornený ohodnotený graf $\mathcal{G}$ nad množinou $\mathbb{N} \subseteq \mathbb{C}$. Ten možno chápať ako reprezentáciu multigrafu – ohodnotenie hrany v grafe $\mathcal{G}$ udáva jej násobnosť, čiže počet hrán medzi danou dvojicou vrcholov v zodpovedajúcom multigrafe. Označme $A := \mathbf{A}(\mathcal{G})$ maticu susednosti grafu $\mathcal{G}$.



Obr. 4.1: Ohodnotený graf $\mathcal{G}$ nad $\mathbb{N}$ a matica susednosti $A = \mathbf{A}(\mathcal{G})$.

Predpokladajme, že pre všetky $t \in \mathbb{N}$ potrebujeme nájsť *počet sledov* dĺžky t vedúcich v príslušnom multigrafe z vrcholu 1 do vrcholu 3. Z kapitoly 1 vieme, že pre $A^t = (a_{i,j}[t])_{3 \times 3}$ je hľadaná hodnota rovná $a_{1,3}[t]$. Na vyjadrenie prvku $a_{1,3}[t]$ v uzavretom tvare využijeme diagonalizáciu matice A .

Charakteristickým polynómom matice A je

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z-1 & -1 & 0 \\ 0 & z-2 & -1 \\ 0 & 0 & z-3 \end{pmatrix} = (z-1)(z-2)(z-3).$$

Vlastné čísla matice A sú teda dané nasledovne:

$$\lambda_1 = 3, \quad \lambda_2 = 2, \quad \lambda_3 = 1.$$

Keďže sú všetky tieto vlastné čísla algebraickej násobnosti 1, je matica A diagonalizovateľná vďaka dôsledku 4.2.9 – ľubovoľná trojica vlastných vektorov príslúchajúcich k vlastným číslam $\lambda_1, \lambda_2, \lambda_3$ navyše bude lineárne nezávislá podľa vety 4.2.8.

Vlastné priestory príslúchajúce k jednotlivým vlastným číslam nájdeme riešením homogénnych sústav lineárnych rovníc:

$$\begin{aligned} E(A, \lambda_1) &= \text{Ker}(\lambda_1 \mathbf{I} - A) = \{c(1, 2, 2)^T \mid c \in \mathbb{C}\}, \\ E(A, \lambda_2) &= \text{Ker}(\lambda_2 \mathbf{I} - A) = \{c(1, 1, 0)^T \mid c \in \mathbb{C}\}, \\ E(A, \lambda_3) &= \text{Ker}(\lambda_3 \mathbf{I} - A) = \{c(1, 0, 0)^T \mid c \in \mathbb{C}\}. \end{aligned}$$

K vlastnému číslu $\lambda_1 = 3$ teda príslúcha napríklad vlastný vektor $\mathbf{v}_1 = (1, 2, 2)^T$, k vlastnému číslu $\lambda_2 = 2$ príslúcha napríklad vlastný vektor $\mathbf{v}_2 = (1, 1, 0)^T$ a k vlastnému číslu $\lambda_3 = 1$ príslúcha napríklad vlastný vektor $\mathbf{v}_3 = (1, 0, 0)^T$. Pre maticu

$$P = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 2 & 0 & 0 \end{pmatrix}$$

³Na dôkaz tohto tvrdenia si napríklad stačí dať niektorý z príkladov 4.1.21 a 4.1.22 do súvisu s vetou 4.2.12.

⁴Z dôsledku 4.2.9 ale napríklad vieme, že postupy z nasledujúcich príkladov budú aplikovateľné vždy, keď pre danú maticu typu $n \times n$ existuje n rôznych vlastných čísel (čiže keď je každé vlastné číslo algebraickej násobnosti 1).

s inverznou maticou

$$P^{-1} = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 1 & -1 \\ 1 & -1 & \frac{1}{2} \end{pmatrix}$$

teda nutne

$$A = P \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} P^{-1} = P \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1},$$

v dôsledku čoho pre všetky $t \in \mathbb{N}$ dostávame

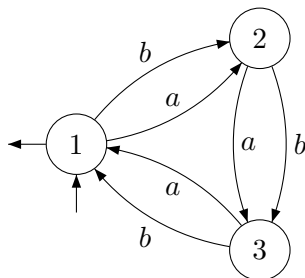
$$A^t = P \begin{pmatrix} 3^t & 0 & 0 \\ 0 & 2^t & 0 \\ 0 & 0 & 1^t \end{pmatrix} P^{-1}$$

a špeciálne

$$a_{1,3}[t] = \frac{1}{2}3^t - 2^t + \frac{1}{2}1^t = \frac{1}{2}3^t - 2^t + \frac{1}{2},$$

čo je hodnota udávajúca hľadaný počet sledov.

Príklad 4.3.2. Uvažujme deterministický konečný automat $\mathcal{A}$ znázornený na obrázku 4.2 a predpokladajme, že potrebujeme nájsť počet slov dĺžky t rozoznávaných automatom $\mathcal{A}$.



Obr. 4.2: Deterministický konečný automat $\mathcal{A}$.

Automat $\mathcal{A}$ je deterministický. Ľahko teda vidieť, že hľadaný počet slov je rovný počtu sledov vedúcich z vrcholu 1 do vrcholu 1 v orientovanom multigrafe zodpovedajúcom automatu $\mathcal{A}$; ten má maticu susednosti

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

Problém sa tak zredukoval na úlohu z predchádzajúceho príkladu – vyriešime ho teda podobným spôsobom. To znamená, že pre maticu $A^t = (a_{i,j}[t])_{3 \times 3}$ vyjadríme v uzavretom tvare hodnotu $a_{1,1}[t]$.

Charakteristický polynóm matice A je daný ako

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z & -2 & 0 \\ 0 & z & -2 \\ -2 & 0 & z \end{pmatrix} = z^3 - 8.$$

Matica A má teda tri vlastné čísla algebraickej násobnosti 1:

$$\lambda_1 = 2, \quad \lambda_2 = -1 + \sqrt{3}i, \quad \lambda_3 = -1 - \sqrt{3}i.$$

V dôsledku toho je diagonalizovateľná a vlastné vektory prislúchajúce k $\lambda_1, \lambda_2, \lambda_3$ sú lineárne nezávislé. Mohli by sme teda postupovať rovnako ako v príklade 4.3.1 nájdením týchto vektorov. Tu si však ukážeme inú metódu: vieme, že pre maticu A^t bude platiť

$$A^t = P \begin{pmatrix} \lambda_1^t & 0 & 0 \\ 0 & \lambda_2^t & 0 \\ 0 & 0 & \lambda_3^t \end{pmatrix} P^{-1}.$$

Preto existujú konštanty $c_1, c_2, c_3 \in \mathbb{C}$ také, že

$$a_{1,1}[t] = c_1 \lambda_1^t + c_2 \lambda_2^t + c_3 \lambda_3^t = c_1 2^t + c_2 \left(-1 + \sqrt{3}i\right)^t + c_3 \left(-1 - \sqrt{3}i\right)^t. \quad (4.6)$$

Zostáva tieto konštanty nájsť. Ľahko ale nahliadneme, že

$$a_{1,1}[0] = 1, \quad a_{1,1}[1] = 0, \quad a_{1,1}[2] = 0.$$

Tieto hodnoty musia byť v súlade s vyjadrením (4.6). Zisťujeme teda, že konštanty c_1, c_2 a c_3 sú riešením sústavy lineárnych rovníc

$$\begin{aligned} 1 &= c_1 2^0 + c_2 \left(-1 + \sqrt{3}i\right)^0 + c_3 \left(-1 - \sqrt{3}i\right)^0, \\ 0 &= c_1 2^1 + c_2 \left(-1 + \sqrt{3}i\right)^1 + c_3 \left(-1 - \sqrt{3}i\right)^1, \\ 0 &= c_1 2^2 + c_2 \left(-1 + \sqrt{3}i\right)^2 + c_3 \left(-1 - \sqrt{3}i\right)^2. \end{aligned}$$

Táto sústava má *jediné* riešenie

$$c_1 = \frac{1}{3}, \quad c_2 = \frac{1}{3}, \quad c_3 = \frac{1}{3}.$$

Môžeme teda uzavrieť, že hľadaný počet slov dĺžky t v jazyku $\|\mathcal{A}\|$ rozoznávanom automatom $\mathcal{A}$ je

$$a_{1,1}[t] = c_1 \lambda_1^t + c_2 \lambda_2^t + c_3 \lambda_3^t = \frac{1}{3} 2^t + \frac{1}{3} \left(-1 + \sqrt{3}i\right)^t + \frac{1}{3} \left(-1 - \sqrt{3}i\right)^t.$$

Ľahko sa možno presvedčiť o tom, že táto hodnota je pre t je deliteľné tromi rovná 2^t ; v opačnom prípade je rovná nule.

Príklad 4.3.3. Predpokladajme napokon, že je naším cieľom vyriešiť nasledujúcu začiatočnú úlohu pre systém lineárnych rekurencií – alebo *diferenčných rovníc*:⁵

$$\begin{aligned} a_{t+1} &= 2a_t - b_t \\ b_{t+1} &= -a_t + 2b_t \end{aligned} \quad \text{pre všetky } t \in \mathbb{N},$$

pričom *začiatočné podmienky* sú dané ako $a_0 = 1$ a $b_0 = 1$. To znamená, že potrebujeme vyjadriť neznáme hodnoty a_t a b_t pre všetky $t \in \mathbb{N}$ – ideálne v uzavretom tvare.

Ak pre všetky $t \in \mathbb{N}$ zavedieme označenie $\mathbf{x}_t = (a_t, b_t)^T$, môžeme uvedenú začiatočnú úlohu prepísať v maticovom tvare ako

$$\mathbf{x}_{t+1} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \mathbf{x}_t \quad \text{pre všetky } t \in \mathbb{N},$$

pričom $\mathbf{x}_0 = (1, 1)^T$. Nie je ťažké nahliadnuť, že *jediné* riešenie tejto úlohy je pre maticu

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

⁵Pomenovanie „diferenčné rovnice“ odôvodníme v kapitole 6, kde sa nimi budeme zaoberať obsiahnejšie.

dané ako

$$A^t \mathbf{x}_0.$$

Potrebuje teda vyjadriť v uzavretom tvare jednotlivé prvky matice A^t . Na tento účel opäť využijeme diagonalizáciu.

Charakteristický polynóm matice A je daný ako

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z-2 & 1 \\ 1 & z-2 \end{pmatrix} = z^2 - 4z + 3 = (z-3)(z-1).$$

Matica A má teda dve vlastné čísla algebraickej násobnosti 1:

$$\lambda_1 = 3 \quad \text{a} \quad \lambda_2 = 1.$$

Vlastné priestory prislúchajúce k týmto vlastným číslam nájdeme riešením homogénnych sústav lineárnych rovníc:

$$\begin{aligned} E(A, \lambda_1) &= \text{Ker}(\lambda_1 \mathbf{I} - A) = \{c(-1, 1)^T \mid c \in \mathbb{C}\}, \\ E(A, \lambda_2) &= \text{Ker}(\lambda_2 \mathbf{I} - A) = \{c(1, 1)^T \mid c \in \mathbb{C}\}. \end{aligned}$$

Dvoma lineárne nezávislými vektormi matice A tak môžu byť napríklad $\mathbf{v}_1 = (-1, 1)^T$ (prislúchajúci k vlastnému číslu $\lambda_1 = 3$) a $\mathbf{v}_2 = (1, 1)^T$ (prislúchajúci k vlastnému číslu $\lambda_2 = 1$). Položme teda

$$P = \begin{pmatrix} \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 \\ \downarrow & \downarrow \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Potom

$$P^{-1} = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

a pre všetky $t \in \mathbb{N}$ dostávame

$$A^t = P \begin{pmatrix} 3^t & 0 \\ 0 & 1^t \end{pmatrix} P^{-1} = \begin{pmatrix} \frac{1}{2}3^t + \frac{1}{2} & -\frac{1}{2}3^t + \frac{1}{2} \\ -\frac{1}{2}3^t + \frac{1}{2} & \frac{1}{2}3^t + \frac{1}{2} \end{pmatrix}.$$

Riešením začiatocnej úlohy teda je vektor

$$\mathbf{x}_t = A^t \mathbf{x}_0 = \begin{pmatrix} \frac{1}{2}3^t + \frac{1}{2} & -\frac{1}{2}3^t + \frac{1}{2} \\ -\frac{1}{2}3^t + \frac{1}{2} & \frac{1}{2}3^t + \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

zodpovedajúci dvojici konštantných postupností $a_t = 1$ a $b_t = 1$ pre všetky $t \in \mathbb{N}$. (Poznamenajme, že k tomuto výsledku možno v tomto konkrétnom prípade prísť aj omnoho jednoduchším spôsobom – stačí si všimnúť, že vektor $\mathbf{x}_0 = (1, 1)^T$ ostáva po prenasobení maticou A nezmenený.)

Príklad 4.3.4. Drobnou zmenou v začiatočných podmienkach úlohy dospejeme pri úlohe z predošlého príkladu k drasticky odlišným výsledkom: napríklad pre začiatocné podmienky $a_0 = 1$ a $b_0 = 0$ máme $\mathbf{x}_0 = (1, 0)^T$ a v dôsledku toho

$$\mathbf{x}_t = A^t \mathbf{x}_0 = \begin{pmatrix} \frac{1}{2}3^t + \frac{1}{2} & -\frac{1}{2}3^t + \frac{1}{2} \\ -\frac{1}{2}3^t + \frac{1}{2} & \frac{1}{2}3^t + \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}3^t + \frac{1}{2} \\ -\frac{1}{2}3^t + \frac{1}{2} \end{pmatrix}.$$

Riešením takejto modifikovanej začiatocnej úlohy sú teda postupnosti dané pre všetky $t \in \mathbb{N}$ ako

$$a_t = \frac{1}{2}3^t + \frac{1}{2}$$

a

$$b_t = -\frac{1}{2}3^t + \frac{1}{2}.$$

4.4 Mocniny lineárnych zobrazení

Kvôli neskoršiemu využitiu pri štúdiu zovšeobecnených vlastných vektorov teraz preskúmame niektoré vlastnosti niekoľkonásobných zložení zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ na konečnorozmernom vektorovom priestore $\mathcal{V}$ samého so sebou, nazývaných aj *mocninami*⁶ zobrazenia φ : kladieme $\varphi^0 = \text{Id}_{\mathcal{V}}$ a $\varphi^{t+1} = \varphi \circ \varphi^t$ pre všetky $t \in \mathbb{N}$. Videli sme, že štvorcové matice nad poľom možno chápať ako špeciálne prípady takýchto lineárnych zobrazení; mocnina lineárneho zobrazenia pritom evidentne zodpovedá mocnine štvorcovej matice. Nasledujúce tvrdenia tak budú aplikovateľné aj na takéto štvorcové matice.

Tvrdenie 4.4.1. *Nech $\mathcal{V}$ je vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Potom*

$$\{\mathbf{0}\} = \text{Ker}(\varphi^0) \subseteq \text{Ker}(\varphi^1) \subseteq \text{Ker}(\varphi^2) \subseteq \text{Ker}(\varphi^3) \subseteq \dots$$

Ak je navyše priestor $\mathcal{V}$ konečnorozmerný, existuje prirodzené číslo $m \leq \dim(\mathcal{V})$ také, že

$$\{\mathbf{0}\} = \text{Ker}(\varphi^0) \subsetneq \dots \subsetneq \text{Ker}(\varphi^m) = \text{Ker}(\varphi^{m+1}) = \text{Ker}(\varphi^{m+2}) = \dots$$

Špeciálne pre $n := \dim(\mathcal{V})$ a všetky $t \in \mathbb{N}$ je $\text{Ker}(\varphi^{n+t}) = \text{Ker}(\varphi^n)$.

Dôkaz. Nech $\mathbf{x} \in \text{Ker}(\varphi^t)$ pre nejaké $t \in \mathbb{N}$. Potom $\varphi^t(\mathbf{x}) = \mathbf{0}$ a v dôsledku toho aj

$$\varphi^{t+1}(\mathbf{x}) = \varphi(\varphi^t(\mathbf{x})) = \varphi(\mathbf{0}) = \mathbf{0}.$$

Preto $\mathbf{x} \in \text{Ker}(\varphi^{t+1})$, čím je dokázaná prvá časť tvrdenia. Predpokladajme teraz, že je priestor $\mathcal{V}$ konečnorozmerný s dimenziou n . Keby bolo

$$\{\mathbf{0}\} = \text{Ker}(\varphi^0) \subsetneq \dots \subsetneq \text{Ker}(\varphi^n) \subsetneq \text{Ker}(\varphi^{n+1}),$$

muselo by byť aj

$$0 = \dim(\text{Ker}(\varphi^0)) < \dots < \dim(\text{Ker}(\varphi^n)) < \dim(\text{Ker}(\varphi^{n+1})),$$

z čoho $\dim(\text{Ker}(\varphi^{n+1})) \geq n+1$; to je ale spor s tým, že ide o podpriestor n -rozmerného priestoru $\mathcal{V}$. Musí teda existovať $m \leq n = \dim(\mathcal{V})$ také, že $\text{Ker}(\varphi^m) = \text{Ker}(\varphi^{m+1})$. Teraz už zjavne stačí ukázať, že pre všetky $t \in \mathbb{N}$ je

$$\text{Ker}(\varphi^{m+t}) \supseteq \text{Ker}(\varphi^{m+t+1}). \quad (4.7)$$

Nech ale $\mathbf{x} \in \text{Ker}(\varphi^{m+t+1})$. Potom

$$\varphi^{m+t+1}(\mathbf{x}) = \varphi^{m+1}(\varphi^t(\mathbf{x})) = \mathbf{0},$$

z čoho $\varphi^t(\mathbf{x}) \in \text{Ker}(\varphi^{m+1}) = \text{Ker}(\varphi^m)$, a teda aj

$$\varphi^{m+t}(\mathbf{x}) = \varphi^m(\varphi^t(\mathbf{x})) = \mathbf{0}.$$

Preto $\mathbf{x} \in \text{Ker}(\varphi^{m+t}(\mathbf{x}))$, čo dokazuje inklúziu (4.7). □

Je dobre známe, že pre všetky lineárne zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ na konečnorozmernom vektorovom priestore $\mathcal{V}$ je

$$\dim(\text{Ker}(\varphi)) + \dim(\text{Im}(\varphi)) = \dim(\mathcal{V}).$$

Napriek tomu ale *nemusí* byť $\mathcal{V} = \text{Ker}(\varphi) \oplus \text{Im}(\varphi)$, ako ukazuje nasledujúci príklad.

⁶Konflikt s mocninou definovanou na základe násobenia funkcií tu nehrozí – na $\mathcal{V}$ totiž vo všeobecnosti nemusí byť daná žiadna multiplikatívna štruktúra.

Príklad 4.4.2. Uvažujme maticu

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

a vektor $\mathbf{x} = (1, 0)^T$. Potom zrejme

$$A\mathbf{x} = \mathbf{0},$$

a teda $\mathbf{x} \in \text{Ker}(A) = \text{Ker}(\varphi_A)$. Súčasne však

$$A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{x},$$

z čoho $\mathbf{x} \in \text{Im}(A) = \text{Im}(\varphi_A)$. Keďže nenulový vektor $\mathbf{x}$ patrí do ich prieniku, súčet podpriestorov

$$\text{Ker}(A) + \text{Im}(A) = \text{Ker}(\varphi_A) + \text{Im}(\varphi_A)$$

podľa tvrdenia 4.1.17 nemôže byť priamy.

V nasledujúcej leme ale dokážeme, že pokiaľ pre nejaké $m \in \mathbb{N}$ je $\text{Ker}(\varphi^m) = \text{Ker}(\varphi^{m+1})$, možno n -rozmerný vektorový priestor $\mathcal{V}$ vždy rozložiť na priamy súčet

$$\mathcal{V} = \text{Ker}(\varphi^m) \oplus \text{Im}(\varphi^m).$$

Podľa tvrdenia 4.4.1 má túto vlastnosť špeciálne aj $m = n$.

Lema 4.4.3. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Pre všetky $m \in \mathbb{N}$ také, že $\text{Ker}(\varphi^m) = \text{Ker}(\varphi^{m+1})$ potom*

$$\mathcal{V} = \text{Ker}(\varphi^m) \oplus \text{Im}(\varphi^m).$$

Špeciálne

$$\mathcal{V} = \text{Ker}(\varphi^n) \oplus \text{Im}(\varphi^n),$$

kde n je dimenzia vektorového priestoru $\mathcal{V}$.

Dôkaz. Na dôkaz prvej časti lemy stačí ukázať, že $\text{Ker}(\varphi^m) \cap \text{Im}(\varphi^m) = \{\mathbf{0}\}$. Uvažujme preto ľubovoľný vektor $\mathbf{x} \in \text{Ker}(\varphi^m) \cap \text{Im}(\varphi^m)$. Potom $\varphi^m(\mathbf{x}) = \mathbf{0}$ a súčasne existuje vektor $\mathbf{y} \in \mathcal{V}$ taký, že $\varphi^m(\mathbf{y}) = \mathbf{x}$. Preto

$$\varphi^{2m}(\mathbf{y}) = \varphi^m(\varphi^m(\mathbf{y})) = \varphi^m(\mathbf{x}) = \mathbf{0},$$

a teda $\mathbf{y} \in \text{Ker}(\varphi^{2m})$. Vďaka tvrdeniu 4.4.1 ale $\text{Ker}(\varphi^{2m}) = \text{Ker}(\varphi^m)$, a teda $\mathbf{y} \in \text{Ker}(\varphi^m)$. Z toho dostávame

$$\mathbf{x} = \varphi^m(\mathbf{y}) = \mathbf{0},$$

čo bolo treba dokázať. Druhá časť lemy je bezprostredným dôsledkom tvrdenia 4.4.1. □

4.5 Zovšeobecnené vlastné vektory

V rámci oddielu 4.2 sme prišli k zisteniu, že súčet vlastných priestorov $E(\varphi, \lambda)$ cez všetky $\lambda \in \sigma(\varphi)$ je pre každé lineárne zobrazenie φ priamy; každý z vlastných priestorov $E(\varphi, \lambda)$ je navyše invariantný vzhľadom na φ . Nech sú teraz $\lambda_1, \dots, \lambda_m$ práve všetky po dvoch rôzne vlastné čísla matice $A \in \mathbb{C}^{n \times n}$ – to znamená, že $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$. Veta 4.2.12 potom hovorí, že matica A je diagonalizovateľná práve vtedy, keď

$$E(A, \lambda_1) \oplus E(A, \lambda_2) \oplus \dots \oplus E(A, \lambda_m) = \mathbb{C}^n.$$

V takom prípade možno bázu priestoru $\mathbb{C}^n$ „vyskladať“ z báz priestorov $E(A, \lambda_1), \dots, E(A, \lambda_m)$ – čiže z vlastných vektorov matice A .

Vo všeobecnosti však matica A diagonalizovateľná byť nemusí; v takom prípade

$$E(A, \lambda_1) \oplus E(A, \lambda_2) \oplus \dots \oplus E(A, \lambda_m) \subsetneq \mathbb{C}^n$$

a priestor $\mathbb{C}^n$ nemá žiadnu bázu zloženú z vlastných vektorov matice A .

Cieľom nasledujúcich úvah bude napraviť tento nedostatok. Definíciu vlastného vektora lineárneho zobrazenia φ prislúchajúceho k nejakému vlastnému číslu λ vhodne zoslabíme, čím získame pojem *zovšeobecneného vlastného vektora* prislúchajúceho k λ . Urobíme to pritom tak, aby priestor $\mathcal{V}$ vždy zaručene mal bázu zloženú zo zovšeobecnených vlastných vektorov zobrazenia φ a aby pre *zovšeobecnené vlastné priestory* $G(\varphi, \lambda_1), \dots, G(\varphi, \lambda_m)$ – tie definujeme ako príslušnú obdobu vlastných priestorov – nutne bolo

$$G(\varphi, \lambda_1) \oplus G(\varphi, \lambda_2) \oplus \dots \oplus G(\varphi, \lambda_m) = \mathcal{V}.$$

Každý z priestorov $G(\varphi, \lambda_1), \dots, G(\varphi, \lambda_m)$ pritom bude invariantný vzhľadom na φ .

Prostredníctvom zobrazenia φ_A tak získame aj pojem zovšeobecneného vlastného vektora *matice* A , ktorý bude samozrejme vykazovať rovnaké vlastnosti. Pre každú maticu typu $n \times n$ teda bude možné z jej zovšeobecnených vlastných vektorov vybrať bázu priestoru $\mathbb{C}^n$. Takáto báza síce isto nezaručí diagonalizovateľnosť matice – ako vieme z vety 4.1.11, na to sú potrebné vlastné vektory – bude však základom pre transformáciu matíc, ktorá je diagonalizácií veľmi blízka: pre prevod matice do takzvaného *Jordanovho kanonického tvaru*. Touto problematikou sa ale budeme zaoberať až v nasledujúcom oddiele: nateraz sa sústreďme na definíciu zovšeobecnených vlastných vektorov a zovšeobecnených vlastných priestorov a na ich vlastnosti.

Vlastný vektor prislúchajúci k vlastnému číslu λ sme definovali ako nenulový vektor $\mathbf{v}$, pre ktorý je $\varphi(\mathbf{v}) = \lambda\mathbf{v}$, čiže $(\varphi - \lambda\text{Id})(\mathbf{v}) = \mathbf{0}$. Malo by byť preto zrejmé, že nasledujúca definícia je skutočne zovšeobecnením definície vlastných vektorov.

Definícia 4.5.1. Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Nech λ je vlastné číslo zobrazenia φ . Vektor $\mathbf{v} \in \mathcal{V}$ nazveme *zovšeobecneným vlastným vektorom* zobrazenia φ prislúchajúcim k λ , ak $\mathbf{v} \neq \mathbf{0}$ a existuje $\ell \in \mathbb{N}$ také, že

$$(\varphi - \lambda\text{Id})^\ell(\mathbf{v}) = \mathbf{0},$$

čiže keď

$$\mathbf{v} \in \text{Ker}\left((\varphi - \lambda\text{Id})^\ell\right) \setminus \{\mathbf{0}\}.$$

Najmenšie také ℓ nazveme *rádom zovšeobecneného vlastného vektora* $\mathbf{v}$.

Namieste je teraz niekoľko poznámok. Predovšetkým: zovšeobecnené vlastné vektory rádu 0 neexistujú – rovnici $(\varphi - \lambda\text{Id})^0(\mathbf{v}) = \text{Id}(\mathbf{v}) = \mathbf{0}$ totiž vyhovuje iba nulový vektor, ktorý sme však v definícii vylúčili. Zovšeobecnený vlastný vektor rádu 1 je ďalej to isté ako vlastný vektor. Všimnime si tiež, že zovšeobecnené vlastné vektory sme definovali iba pre *vlastné čísla* λ . Táto voľba však nie je nijak obmedzujúca – ak λ nie je vlastné číslo, je $\text{Ker}(\varphi - \lambda\text{Id}) = \{\mathbf{0}\} = \text{Ker}((\varphi - \lambda\text{Id})^0)$ a z tvrdenia 4.4.1 potom aj pre všetky $\ell \in \mathbb{N}$ dostávame $\text{Ker}((\varphi - \lambda\text{Id})^\ell) = \text{Ker}((\varphi - \lambda\text{Id})^0) = \{\mathbf{0}\}$.

Definícia 4.5.2. Nech $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$ je štvorcová matica. *Zovšeobecneným vlastným vektorom* matice A prislúchajúcim k vlastnému číslu λ nazveme ľubovoľný zovšeobecnený vlastný vektor $\mathbf{v} \in \mathbb{C}^n$ zobrazenia φ_A prislúchajúci λ . *Rádom zovšeobecneného vlastného vektora* $\mathbf{v}$ matice A nazveme rád vektora $\mathbf{v}$ chápaného ako zovšeobecnený vlastný vektor zobrazenia φ_A .

Vektor $\mathbf{v} \in \mathbb{C}^n$ je teda *zovšeobecneným vlastným vektorom* matice A prislúchajúcim k vlastnému číslu λ práve vtedy, keď $\mathbf{v} \neq \mathbf{0}$ a existuje $\ell \in \mathbb{N}$ také, že

$$(A - \lambda\mathbf{I})^\ell \mathbf{v} = \mathbf{0},$$

čiže keď

$$\mathbf{v} \in \text{Ker} \left((A - \lambda \mathbf{I})^\ell \right) \setminus \{\mathbf{0}\}.$$

Najmenšie také ℓ je potom *rádom zovšeobecneného vlastného vektora* $\mathbf{v}$.

Vlastný priestor $E(\varphi, \lambda)$ lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ prislúchajúci k jeho vlastnému číslu λ sme v oddiele 4.1 definovali ako podpriestor priestoru $\mathcal{V}$ generovaný množinou všetkých vlastných vektorov zobrazenia φ prislúchajúcich k λ . Videli sme však, že $E(\varphi, \lambda)$ je v skutočnosti množina všetkých vlastných vektorov φ prislúchajúcich k λ , do ktorej je ešte pridaný nulový vektor; ukázali sme tiež, že platí $E(\varphi, \lambda) = \text{Ker}(\varphi - \lambda \text{Id})$. Nasledujúca definícia *zovšeobecnených vlastných priestorov* je priamočiarym zovšeobením všetkých troch týchto vlastností.

Definícia 4.5.3. Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Nech λ je vlastné číslo zobrazenia φ . *Zovšeobeným vlastným priestorom* zobrazenia φ prislúchajúcim k vlastnému číslu λ nazveme podpriestor $G(\varphi, \lambda)$ priestoru $\mathcal{V}$ generovaný množinou všetkých zovšeobených vlastných vektorov zobrazenia φ prislúchajúcich k λ .

Tvrdenie 4.5.4. Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Nech λ je vlastné číslo zobrazenia φ . Potom $G(\varphi, \lambda)$ pozostáva z práve všetkých zovšeobených vlastných vektorov zobrazenia φ prislúchajúcich k λ a z nulového vektora. Navyše

$$G(\varphi, \lambda) = \text{Ker}((\varphi - \lambda \text{Id})^n),$$

kde n je dimenzia vektorového priestoru $\mathcal{V}$.

Dôkaz. Nenulový vektor $\mathbf{v} \in \mathcal{V}$ je zovšeobeným vlastným vektorom zobrazenia φ prislúchajúcim k λ práve vtedy, keď je prvkom jadra zobrazenia $(\varphi - \lambda \text{Id})^\ell$ pre nejaké $\ell \in \mathbb{N}$, čo podľa tvrdenia 4.4.1 nastane práve vtedy, keď $\mathbf{v} \in \text{Ker}((\varphi - \lambda \text{Id})^n)$. Priestor $\text{Ker}((\varphi - \lambda \text{Id})^n)$ teda pozostáva z práve všetkých zovšeobených vlastných vektorov φ prislúchajúcich k λ a z nulového vektora – a ako taký musí byť týmito zovšeobenými vlastnými vektormi generovaný: čiže $\text{Ker}((\varphi - \lambda \text{Id})^n) = G(\varphi, \lambda)$. $\square$

Definícia 4.5.5. Nech n je prirodzené číslo, $A \in \mathbb{C}^{n \times n}$ je štvorcová matica a $\lambda \in \mathbb{C}$ je vlastné číslo matice A . *Zovšeobeným vlastným priestorom* matice A prislúchajúcim k vlastnému číslu λ nazveme priestor $G(A, \lambda) := G(\varphi_A, \lambda)$.

Je zrejmé, že $G(A, \lambda)$ je podpriestor priestoru $\mathbb{C}^n$ generovaný množinou zovšeobených vlastných vektorov matice A prislúchajúcich k λ . Pozostáva tak zo všetkých vlastných vektorov matice A prislúchajúcich k λ a z nulového vektora a je rovný podpriestoru $\text{Ker}((A - \lambda \mathbf{I})^n)$ priestoru $\mathbb{C}^n$.

Než prejdeme k výsledkom o zovšeobených vlastných vektoroch a priestoroch, dokážeme lemu o „komutatívnosti“ zobrazení typu $(\varphi - a \text{Id})$ pre $a \in \mathbb{C}$ a fixné lineárne zobrazenie φ .

Lema 4.5.6. Nech $\mathcal{V}$ je vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Pre všetky $a, b \in \mathbb{C}$ potom

$$(\varphi - a \text{Id}) \circ (\varphi - b \text{Id}) = (\varphi - b \text{Id}) \circ (\varphi - a \text{Id}).$$

Dôkaz. Pre všetky $\mathbf{x} \in \mathcal{V}$ je

$$((\varphi - a \text{Id}) \circ (\varphi - b \text{Id}))(\mathbf{x}) = \varphi(\varphi(\mathbf{x}) - b\mathbf{x}) - a(\varphi(\mathbf{x}) - b\mathbf{x}) = \varphi^2(\mathbf{x}) - a\varphi(\mathbf{x}) - b\varphi(\mathbf{x}) + ab\mathbf{x},$$

čo je to isté ako

$$((\varphi - b \text{Id}) \circ (\varphi - a \text{Id}))(\mathbf{x}) = \varphi(\varphi(\mathbf{x}) - a\mathbf{x}) - b(\varphi(\mathbf{x}) - a\mathbf{x}) = \varphi^2(\mathbf{x}) - a\varphi(\mathbf{x}) - b\varphi(\mathbf{x}) + ab\mathbf{x}. \quad \square$$

O niečo zrejmejšia je predchádzajúca lema v reči matíc: každý súčin matíc typu $(A - a\mathbf{I})$ pre fixnú maticu A totiž možno chápať ako „polynóm o premennej A “.

Po zvyšok oddielu budeme dokazovať už avizované tvrdenie, podľa ktorého možno pre každé lineárne zobrazenie $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ na konečnorozmernom vektorovom priestore $\mathcal{V}$ tento priestor rozložiť na priamy súčet zovšeobecnených vlastných priestorov zobrazenia φ . Prvým krokom k tomuto výsledku bude zosilnenie vety 4.2.8 hovoriace o lineárnej nezávislosti vlastných vektorov prislúchajúcich k po dvoch rôznym vlastným číslam: dokážeme, že rovnaká vlastnosť platí aj pre zovšeobecnené vlastné vektory.

Veta 4.5.7. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie, $m \in \mathbb{N}$ a $\lambda_1, \dots, \lambda_m$ po dvoch rôzne vlastné čísla zobrazenia φ . Nech pre $k = 1, \dots, m$ je $\mathbf{v}_k \in \mathcal{V}$ zovšeobecnený vlastný vektor zobrazenia φ prislúchajúci k λ_k . Potom sú vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ lineárne nezávislé.*

Dôkaz. Pre $m = 0$ je tvrdenie triviálne pravdivé; nech teda $m \geq 1$. Nech $a_1, \dots, a_m$ sú komplexné koeficienty také, že

$$a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m = \mathbf{0}. \quad (4.8)$$

Dokážeme, že pre $j = 1, \dots, m$ je $a_j = 0$. Nech $j \in [m]$ je dané pevne a predpokladajme, že zovšeobecnený vlastný vektor $\mathbf{v}_j$ je rádu ℓ . V takom prípade nutne $\ell \geq 1$,

$$(\varphi - \lambda_j \text{Id})^\ell(\mathbf{v}_j) = \mathbf{0}, \quad (\varphi - \lambda_j \text{Id})^{\ell-1}(\mathbf{v}_j) = \mathbf{v}'_j \neq \mathbf{0} \quad \text{a} \quad (\varphi - \lambda_j \text{Id})(\mathbf{v}'_j) = \mathbf{0}.$$

Z poslednej z týchto rovností teda vidíme, že

$$\varphi(\mathbf{v}'_j) = \lambda_j \mathbf{v}'_j$$

a $\mathbf{v}'_j$ je vlastný vektor zobrazenia φ prislúchajúci k λ_j . Z toho vyplýva, že pre všetky $a \in \mathbb{C}$ je $(\varphi - a\text{Id})(\mathbf{v}'_j) = (\lambda_j - a)\mathbf{v}'_j$ a aplikovaním tejto identity t -krát pre $t \in \mathbb{N}$ dostávame aj

$$(\varphi - a\text{Id})^t(\mathbf{v}'_j) = (\lambda_j - a)^t \mathbf{v}'_j. \quad (4.9)$$

Označme teraz ako n dimenziu priestoru $\mathcal{V}$ a ako ψ zobrazenie $\psi: \mathcal{V} \rightarrow \mathcal{V}$ dané ako

$$\psi = (\varphi - \lambda_1 \text{Id})^n \circ \dots \circ (\varphi - \lambda_{j-1} \text{Id})^n \circ (\varphi - \lambda_{j+1} \text{Id})^n \circ \dots \circ (\varphi - \lambda_m \text{Id})^n.$$

Aplikovaním zobrazenia ψ na obidve strany rovnosti (4.8) potom s využitím lemy 4.5.6 a rovnosti $G(\varphi, \lambda_k) = \text{Ker}((\varphi - \lambda_k \text{Id})^n)$ pre $k = 1, \dots, j-1, j+1, \dots, m$ dostávame

$$a_j \psi(\mathbf{v}_j) = \mathbf{0},$$

z čoho po aplikovaní zobrazenia $(\varphi - \lambda_j \text{Id})^{\ell-1}$ na obe strany rovnosti a opätovnom využití lemy 4.5.6 vyplýva

$$a_j \psi \left((\varphi - \lambda_j \text{Id})^{\ell-1}(\mathbf{v}_j) \right) = a_j \psi(\mathbf{v}'_j) = \mathbf{0}.$$

To následne možno s použitím identity (4.9) prepísať ako

$$a_j (\lambda_j - \lambda_1)^n \dots (\lambda_j - \lambda_{j-1})^n (\lambda_j - \lambda_{j+1})^n \dots (\lambda_j - \lambda_m)^n \mathbf{v}'_j = \mathbf{0}.$$

Keďže sú ale vlastné čísla $\lambda_1, \dots, \lambda_m$ po dvoch rôzne a vektor $\mathbf{v}'_j$ je nenulový, platnosť tejto rovnosti implikuje $a_j = 0$, čo bolo treba dokázať. $\square$

Rovnako ako pre vlastné priestory tak prichádzame k zisteniu, že súčet zovšeobecnených vlastných priestorov prislúchajúcich k vlastným číslam lineárneho zobrazenia je vždy priamy.

Dôsledok 4.5.8. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie, $m \in \mathbb{N}$ a $\lambda_1, \dots, \lambda_m$ po dvoch rôzne vlastné čísla zobrazenia φ . Súčet zovšeobecnených vlastných priestorov $G(\varphi, \lambda_1), \dots, G(\varphi, \lambda_m)$ je potom priamy:*

$$G(\varphi, \lambda_1) + \dots + G(\varphi, \lambda_m) = G(\varphi, \lambda_1) \oplus \dots \oplus G(\varphi, \lambda_m) = \bigoplus_{k=1}^m G(\varphi, \lambda_k).$$

Každý vektor

$$\mathbf{x} \in G(\varphi, \lambda_1) + \dots + G(\varphi, \lambda_m)$$

tak možno jediným spôsobom vyjadriť ako $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_m$, kde pre $k = 1, \dots, m$ je $\mathbf{x}_k \in G(\varphi, \lambda_k)$. Ak ďalej pre $k = 1, \dots, m$ je $(\mathbf{v}_{k,1}, \dots, \mathbf{v}_{k,s_k})$ bázou $G(\varphi, \lambda_k)$, je $(\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,s_1}, \dots, \mathbf{v}_{m,1}, \dots, \mathbf{v}_{m,s_m})$ bázou priestoru $\bigoplus_{k=1}^m G(\varphi, \lambda_k)$.

Dôkaz. Pre $m = 0$ a $m = 1$ je tvrdenie triviálne pravdivé. Uvažujme teda $m \geq 2$ a za účelom sporu predpokladajme, že súčet $G(\varphi, \lambda_1) + \dots + G(\varphi, \lambda_m)$ nie je priamy. Podľa tvrdenia 4.1.17 potom existuje $k \in [m]$ také, že pre nejaký vektor $\mathbf{v}_k \neq \mathbf{0}$ je

$$\mathbf{v}_k \in G(\varphi, \lambda_k) \cap (G(\varphi, \lambda_1) + \dots + G(\varphi, \lambda_{k-1}) + G(\varphi, \lambda_{k+1}) + \dots + G(\varphi, \lambda_m)).$$

To znamená, že $\mathbf{v}_k$ je zovšeobecneným vlastným vektorom zobrazenia φ prislúchajúcim k λ_k a súčasne existujú vektory $\mathbf{v}_1 \in G(\varphi, \lambda_1), \dots, \mathbf{v}_{k-1} \in G(\varphi, \lambda_{k-1}), \mathbf{v}_{k+1} \in G(\varphi, \lambda_{k+1}), \dots, \mathbf{v}_m \in G(\varphi, \lambda_m)$ také, že

$$\mathbf{v}_k = \mathbf{v}_1 + \dots + \mathbf{v}_{k-1} + \mathbf{v}_{k+1} + \dots + \mathbf{v}_m.$$

Keďže je ale pre $j = 1, \dots, m$ vektor $\mathbf{v}_j$ v prípade svojej nenulovosti zovšeobecneným vlastným vektorom prislúchajúcim k λ_j a keďže je vektor $\mathbf{v}_k$ nenulový, ide o spor s vetou 4.5.7. $\square$

Lineárny súčet zovšeobecnených vlastných priestorov prislúchajúcich k jednotlivým vlastným číslam lineárneho zobrazenia je teda vždy priamy. K avizovanému cieľu tohto oddielu nám tak chýba už len pozorovanie, že súčtom týchto priestorov cez *všetky* vlastné čísla lineárneho zobrazenia $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je vždy kompletný priestor $\mathcal{V}$. Nasledujúca lema nám poskytne predprípravu pre toto pozorovanie.

Lema 4.5.9. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Pre $n := \dim(\mathcal{V})$ a každé vlastné číslo $\lambda \in \sigma(\varphi)$ sú potom pravdivé nasledujúce tvrdenia:*

- (i) *Priestor $G(\varphi, \lambda) = \text{Ker}((\varphi - \lambda \text{Id})^n)$ je invariantný vzhľadom na φ .*
- (ii) *Priestor $\text{Im}((\varphi - \lambda \text{Id})^n)$ je invariantný vzhľadom na φ .*
- (iii) *Platí $\mathcal{V} = G(\varphi, \lambda) \oplus \text{Im}((\varphi - \lambda \text{Id})^n) = \text{Ker}((\varphi - \lambda \text{Id})^n) \oplus \text{Im}((\varphi - \lambda \text{Id})^n)$.*

Dôkaz. Nech $\mathbf{x} \in G(\varphi, \lambda) = \text{Ker}((\varphi - \lambda \text{Id})^n)$, čo znamená, že $(\varphi - \lambda \text{Id})^n(\mathbf{x}) = \mathbf{0}$. Podľa lemy 4.5.6 potom

$$(\varphi - \lambda \text{Id})^n(\varphi(\mathbf{x})) = \varphi((\varphi - \lambda \text{Id})^n(\mathbf{x})) = \varphi(\mathbf{0}) = \mathbf{0},$$

a teda nutne $\varphi(\mathbf{x}) \in \text{Ker}((\varphi - \lambda \text{Id})^n) = G(\varphi, \lambda)$. Priestor $G(\varphi, \lambda)$ je teda invariantný vzhľadom na φ , čo dokazuje tvrdenie (i). Nech ďalej $\mathbf{x} \in \text{Im}((\varphi - \lambda \text{Id})^n)$. Potom existuje $\mathbf{y} \in \mathcal{V}$ také, že

$$(\varphi - \lambda \text{Id})^n(\mathbf{y}) = \mathbf{x}.$$

S pomocou lemy 4.5.6 ale v takom prípade dostávame

$$(\varphi - \lambda \text{Id})^n(\varphi(\mathbf{y})) = \varphi((\varphi - \lambda \text{Id})^n(\mathbf{y})) = \varphi(\mathbf{x}),$$

z čoho $\varphi(\mathbf{x}) \in \text{Im}((\varphi - \lambda \text{Id})^n)$. Priestor $\text{Im}((\varphi - \lambda \text{Id})^n)$ je teda tiež invariantný vzhľadom na φ a dokázané je aj tvrdenie (ii). Tvrdenie (iii) je napokon priamym dôsledkom lemy 4.4.3. $\square$

Môžeme konečne vysloviť a dokázať hlavnú vetu tohto oddielu: pre každé lineárne zobrazenie φ na konečnorozmernom priestore $\mathcal{V}$ je tento priamym súčtom zovšeobecnených vlastných priestorov prislúchajúcich k jednotlivým vlastným číslam zobrazenia φ . Dimenzia každého z týchto zovšeobecnených vlastných priestorov je navyše rovná algebraickej násobnosti príslušného vlastného čísla.

Veta 4.5.10. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Ak $\sigma(\varphi) = \{\lambda_1, \dots, \lambda_m\}$ pre po dvoch rôzne $\lambda_1, \dots, \lambda_m$, tak*

$$\mathcal{V} = G(\varphi, \lambda_1) \oplus G(\varphi, \lambda_2) \oplus \dots \oplus G(\varphi, \lambda_m) = \bigoplus_{\lambda \in \sigma(\varphi)} G(\varphi, \lambda).$$

Pre $k = 1, \dots, m$ navyše $\dim(G(\varphi, \lambda_k)) = \alpha(\varphi, \lambda_k)$.

Dôkaz. Prvú časť vety dokážeme indukciou vzhľadom na $n = \dim(\mathcal{V})$. Pre $n = 0$ je tvrdenie triviálne a pre $n = 1$ má zobrazenie φ zrejme jediné vlastné číslo λ_1 , pre ktoré je

$$\mathcal{V} = G(\varphi, \lambda_1) = E(\varphi, \lambda_1).$$

Predpokladajme teraz, že prvá časť vety platí pre $n = 0, \dots, q$, kde $q \in \mathbb{N} \setminus \{0\}$; uvažujme $n = q + 1$. Z lemy 4.5.9 potom

$$\mathcal{V} = G(\varphi, \lambda_1) \oplus \mathcal{I}, \quad (4.10)$$

kde $\mathcal{I} = \text{Im}((\varphi - \lambda_1 \text{Id})^n)$ a obidva priestory $G(\varphi, \lambda_1)$ a $\mathcal{I}$ sú invariantné vzhľadom na φ . Keďže všetky vlastné vektory zobrazenia φ prislúchajúce k λ_1 patria do $G(\varphi, \lambda_1)$ a keďže pre $k = 2, \dots, m$ je $G(\varphi, \lambda_1) \cap E(\varphi, \lambda_k) \subseteq G(\varphi, \lambda_1) \cap G(\varphi, \lambda_k) = \{\mathbf{0}\}$, nutne $\sigma(\varphi|_{\mathcal{I}}) = \{\lambda_2, \dots, \lambda_m\}$. Zrejme navyše $\dim(G(\varphi, \lambda_1)) \geq 1$, v dôsledku čoho $\dim(\mathcal{I}) \leq q$. Ak $\dim(\mathcal{I}) = 0$, nie je čo dokazovať – λ_1 je totiž v takom prípade zjavne jediným vlastným číslom, pričom $\mathcal{V} = G(\varphi, \lambda_1)$. Pre $1 \leq \dim(\mathcal{I}) \leq q$ z indukčného predpokladu dostávame

$$\mathcal{I} = G(\varphi|_{\mathcal{I}}, \lambda_2) \oplus \dots \oplus G(\varphi|_{\mathcal{I}}, \lambda_m). \quad (4.11)$$

Potrebuje dokázať, že pre $k = 2, \dots, m$ je $G(\varphi|_{\mathcal{I}}, \lambda_k) = G(\varphi, \lambda_k)$. Inklúzia $G(\varphi|_{\mathcal{I}}, \lambda_k) \subseteq G(\varphi, \lambda_k)$ je zrejma: pre každé $\mathbf{v} \in G(\varphi|_{\mathcal{I}}, \lambda_k) \subseteq \mathcal{I}$ totiž

$$(\varphi - \lambda \text{Id}_{\mathcal{V}})^n(\mathbf{v}) = (\varphi|_{\mathcal{I}} - \lambda \text{Id}_{\mathcal{I}})^n(\mathbf{v}) = \mathbf{0}.$$

Na dôkaz opačnej inklúzie uvažujme ľubovoľné $\mathbf{v} \in G(\varphi, \lambda_k)$. Zo vzťahov (4.10) a (4.11) vyplýva, že existujú vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ také, že $\mathbf{v}_1 \in G(\varphi, \lambda_1)$, pre $\ell = 2, \dots, m$ je $\mathbf{v}_\ell \in G(\varphi|_{\mathcal{I}}, \lambda_\ell)$ a

$$\mathbf{v} = \mathbf{v}_1 + \dots + \mathbf{v}_m.$$

Z toho $\mathbf{v}_1 + \dots + \mathbf{v}_{k-1} + (\mathbf{v}_k - \mathbf{v}) + \mathbf{v}_{k+1} + \dots + \mathbf{v}_m = \mathbf{0}$. Nenulové spomedzi vektorov tohto súčtu sú nutne zovšeobecnenými vlastnými vektormi prislúchajúcimi k po dvoch rôznym vlastným číslam zobrazenia φ . Z vety 4.5.7 preto vyplýva, že sú lineárne nezávislé a ich lineárnou kombináciou tak nemožno získať nulový vektor. Pre $\ell = 1, \dots, k-1, k+1, \dots, m$ teda nutne $\mathbf{v}_\ell = \mathbf{0}$ a $\mathbf{v} = \mathbf{v}_k \in G(\varphi|_{\mathcal{I}}, \lambda_k)$.

Skombinovaním práve dokázanej skutočnosti so vzťahmi (4.10) a (4.11) napokon dostávame

$$\mathcal{V} = G(\varphi, \lambda_1) \oplus G(\varphi, \lambda_2) \oplus \dots \oplus G(\varphi, \lambda_m),$$

čím je prvá časť vety dokázaná. Zostáva ukázať, že pre $k = 1, \dots, m$ je $n_k := \dim(G(\varphi, \lambda_k)) = \alpha(\varphi, \lambda_k)$. Označme pre $k = 1, \dots, m$ ako φ_k zúženie zobrazenia φ na $G(\varphi, \lambda_k)$; zrejme potom $\text{ch}_{\varphi_k}(z) = (z - \lambda_k)^{n_k}$, lebo inak by priestor $G(\varphi, \lambda_k)$ musel obsahovať vlastný vektor prislúchajúci k vlastnému číslu rôznemu od λ_k . Keďže navyše

$$\mathcal{V} = G(\varphi, \lambda_1) \oplus G(\varphi, \lambda_2) \oplus \dots \oplus G(\varphi, \lambda_m)$$

a priestory $G(\varphi, \lambda_k)$ sú pre $k = 1, \dots, m$ invariantné vzhľadom na φ , z tvrdenia 4.1.18 dostávame

$$\text{ch}_{\varphi}(z) = \text{ch}_{\varphi_1}(z) \cdot \text{ch}_{\varphi_2}(z) \cdot \dots \cdot \text{ch}_{\varphi_m}(z) = (z - \lambda_1)^{n_1} (z - \lambda_2)^{n_2} \dots (z - \lambda_m)^{n_m}.$$

Z tvrdenia 4.2.7 však na druhej strane vyplýva

$$\text{ch}_\varphi(z) = (z - \lambda_1)^{\alpha(\varphi, \lambda_1)} (z - \lambda_2)^{\alpha(\varphi, \lambda_2)} \dots (z - \lambda_m)^{\alpha(\varphi, \lambda_m)};$$

pre $k = 1, \dots, m$ teda skutočne $n_k = \dim(G(\varphi, \lambda_k)) = \alpha(\varphi, \lambda_k)$ a veta je dokázaná. $\square$

Dôsledok 4.5.11. *Nech $n \in \mathbb{N}$, $\mathcal{V}$ je n -rozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Potom existuje n lineárne nezávislých zovšeobecnených vlastných vektorov $\mathbf{v}_1, \dots, \mathbf{v}_n$ zobrazenia φ tvoriacich bázu priestoru $\mathcal{V}$.*

4.6 Jordanov kanonický tvar matice

Videli sme, že existujú matice, ktoré nie sú diagonalizovateľné. Nie sú teda podobné žiadnej diagonálnej matici, čo okrem iného zabraňuje univerzálnemu použitiu metód podobných tým z oddielu 4.3.

S využitím teórie z oddielu 4.5 teraz negatívny dopad tohto nedostatku zmiernime. Ukážeme, že aj nediagonalizovateľné štvorcové matice sú vždy podobné matici v tvare, ktorý je „veľmi blízky“ diagonálnemu: tento tvar nazveme *Jordanovým kanonickým tvarom* matice. Pre matice v Jordanovom kanonickom tvare budeme vedieť, podobne ako pre diagonálne matice, vyjadriť prvky na jednotlivých pozíciách ich t -tej mocniny v uzavretom tvare. To nám otvorí cestu k adaptácii metód z oddielu 4.3 pre všeobecné matice a k aplikáciám, ktorým venujeme podstatnú časť nasledujúcich dvoch kapitol.

Definícia 4.6.1. Nech n je nenulové prirodzené číslo a λ je komplexné číslo. *Jordanovou bunkou* rádu n prislúchajúcou k číslu λ nazveme štvorcovú maticu $J_n(\lambda) \in \mathbb{C}^{n \times n}$ danú ako

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Definícia 4.6.2. Nech n je prirodzené číslo a $A \in \mathbb{C}^{n \times n}$ je matica. Budeme hovoriť, že matica A je v *Jordanovom kanonickom tvare*, ak existuje prirodzené číslo k , komplexné čísla $\lambda_1, \dots, \lambda_k$ a nenulové prirodzené čísla $n_1, \dots, n_k$ také, že $n_1 + \dots + n_k = n$ a matica A má blokový rozklad

$$A = \begin{pmatrix} J_{n_1}(\lambda_1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & J_{n_2}(\lambda_2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & J_{n_k}(\lambda_k) \end{pmatrix}.$$

Matica v Jordanovom kanonickom tvare tak môže obsahovať nenulové prvky nanajvýš na hlavnej diagonále a na diagonále nad ňou, pričom v druhom prípade môžu byť takýmito nenulovými prvkami iba jednotky.

V nasledujúcom postupne odôvodníme pomenovanie *Jordanov kanonický tvar* – ukážeme, že každá štvorcová matica je podobná nejakej matici v tomto tvare. Urobíme to tak, že pre ľubovoľné lineárne zobrazenie $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ na konečnorozmernom vektorovom priestore $\mathcal{V}$ nad $\mathbb{C}$ nájdeme bázu $\mathcal{B}$, pre ktorú je matica $A_\varphi[\mathcal{B}]$ v Jordanovom kanonickom tvare.⁷

Z vety 4.5.10 vieme, že priestor $\mathcal{V}$ možno rozložiť na priamy súčet

$$\mathcal{V} = G(\varphi, \lambda_1) \oplus G(\varphi, \lambda_2) \oplus \dots \oplus G(\varphi, \lambda_m)$$

⁷Vieme, že každá matica je maticou nejakého lineárneho zobrazenia, pričom matice toho istého lineárneho zobrazenia nad rôznymi bázami sú nutne podobné.

podpriestorov invariantných vzhľadom na φ , kde $\lambda_1, \dots, \lambda_m$ sú práve všetky (po dvoch rôzne) vlastné čísla zobrazenia φ . Z oddielu 4.1⁸ ďalej vieme, že pre bázy priestorov $G(\varphi, \lambda_j)$ dané pre $j = 1, \dots, m$ ako $\mathcal{B}_j = (\mathbf{v}_{j,1}, \dots, \mathbf{v}_{j,s_j})$ možno maticu zobrazenia φ vzhľadom na bázu

$$\mathcal{B} = (\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,s_1}, \dots, \mathbf{v}_{m,1}, \dots, \mathbf{v}_{m,s_m})$$

vyjadriť ako blokovo diagonálnu maticu, ktorej diagonálne bloky sú maticami zúžení zobrazenia φ na jednotlivé invariantné podpriestory $G(\varphi, \lambda_1), \dots, G(\varphi, \lambda_m)$:

$$A_\varphi[\mathcal{B}] = \begin{pmatrix} A_{\varphi|_{G(\varphi, \lambda_1)}}[\mathcal{B}_1] & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & A_{\varphi|_{G(\varphi, \lambda_2)}}[\mathcal{B}_2] & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & A_{\varphi|_{G(\varphi, \lambda_m)}}[\mathcal{B}_m] \end{pmatrix}.$$

Stačí sa teda zamerať na fixný zovšeobecnený vlastný priestor $G(\varphi, \lambda_j)$ a na zúženie $\varphi_j := \varphi|_{G(\varphi, \lambda_j)}$. Keďže navyše $G(\varphi, \lambda_j) = \text{Ker}((\varphi_j - \lambda_j \text{Id})^{s_j})$, kde $s_j = \alpha(\varphi, \lambda_j)$ je dimenzia priestoru $G(\varphi, \lambda_j)$, musí pre zobrazenie $\nu := \varphi_j - \lambda_j \text{Id}$ existovať nenulové prirodzené číslo $r \leq s_j$ také, že $\nu^r = \mathbf{0}$, kde $\mathbf{0}$ označuje nulové zobrazenie na $G(\varphi, \lambda_j)$. Takéto lineárne zobrazenia, z ktorých po umocnení na dostatočne veľký exponent vznikne nulové zobrazenie, nazývame *nilpotentnými*. Je zrejmé, že ak $\nu = \varphi_j - \lambda_j \text{Id}$ má pri nejakej báze $\mathcal{B}_j$ priestoru $G(\varphi, \lambda_j)$ maticu $A_\nu[\mathcal{B}_j]$ v Jordanovom kanonickom tvare, je v Jordanovom kanonickom tvare aj matica $A_{\varphi_j}[\mathcal{B}_j] = A_\nu[\mathcal{B}_j] + \lambda_j \mathbf{I}$. Stačí sa teda zamerať na nilpotentné zobrazenia, pričom priestor $G(\varphi, \lambda_j)$ už v takom prípade prestáva hrať rolu; možno teda namiesto neho pracovať nad ľubovoľným netriviálnym konečnorozmerným priestorom $\mathcal{W}$.

Definícia 4.6.3. Nech $\mathcal{W}$ je netriviálny konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\psi: \mathcal{W} \rightarrow \mathcal{W}$ je lineárne zobrazenie. Zobrazenie ψ nazveme *nilpotentným*, ak existuje prirodzené číslo r také, že $\psi^r = \mathbf{0}_{\mathcal{W}}$, kde $\mathbf{0}_{\mathcal{W}}: \mathcal{W} \rightarrow \mathcal{W}$ je lineárne zobrazenie dané pre všetky $\mathbf{x} \in \mathcal{W}$ ako $\mathbf{0}_{\mathcal{W}}(\mathbf{x}) = \mathbf{0}$. Najmenšie r s touto vlastnosťou – to musí byť zjavne nenulové – nazveme *rádom nilpotentného zobrazenia* ψ .

Definícia 4.6.4. Nech n je nenulové prirodzené číslo a $A \in \mathbb{C}^{n \times n}$ je matica. Maticu A nazveme *nilpotentnou*, ak je φ_A nilpotentné lineárne zobrazenie na $\mathbb{C}^n$.

Štvorcová matica $A \in \mathbb{C}^{n \times n}$ je teda očividne nilpotentná práve vtedy, keď existuje prirodzené číslo r také, že $A^r = \mathbf{0}$. Ľahko tiež vidieť, že jediným vlastným číslom nilpotentného lineárneho zobrazenia $\psi: \mathcal{W} \rightarrow \mathcal{W}$ – či nilpotentnej štvorcovej matice $A \in \mathbb{C}^{n \times n}$ – môže byť nula. Ak je totiž ψ nilpotentné zobrazenie rádu r , k ľubovoľnému jeho vlastnému číslu λ musí prislúchať vlastný vektor $\mathbf{w} \neq \mathbf{0}$ taký, že $\psi(\mathbf{w}) = \lambda \mathbf{w}$, z čoho $\psi^r(\mathbf{w}) = \lambda^r \mathbf{w}$ – avšak súčasne $\psi^r(\mathbf{w}) = \mathbf{0}$, pričom splnenie oboch rovností je pre nenulový vektor $\mathbf{w}$ možné iba v prípade, že $\lambda = 0$.

Ukážeme, že ku každému nilpotentnému lineárnemu zobrazeniu $\psi: \mathcal{W} \rightarrow \mathcal{W}$ rádu r na netriviálnom konečnorozmernom vektorovom priestore $\mathcal{W}$ možno nájsť bázu $\mathcal{B}$ priestoru $\mathcal{W}$, vzhľadom na ktorú má zobrazenie ψ maticu v Jordanovom kanonickom tvare, pričom každá Jordanova bunka tejto matice je rovná $J_k(0)$ pre nejaké $k \in [r]$. Z argumentácie vyššie vyplýva, že tým bude prakticky dokázané aj analogické tvrdenie pre všeobecné lineárne zobrazenia na konečnorozmerných priestoroch.

Za účelom nájdenia takejto bázy si najprv všimnime pôsobenie lineárneho zobrazenia prislúchajúceho k Jordanovej bunke

$$J_k(0) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

⁸Presnejšie zo vzťahu (4.4) skombinovaného s indukciou vzhľadom na m .

na vektoroch štandardnej bázy $\mathbf{e}_1 = (1, 0, \dots, 0)^T, \dots, \mathbf{e}_k = (0, \dots, 0, 1)^T$ priestoru $\mathbb{C}^k$. Zrejme

$$J_k(0) \mathbf{e}_k = \mathbf{e}_{k-1}, \quad J_k(0) \mathbf{e}_{k-1} = \mathbf{e}_{k-2}, \quad \dots, \quad J_k(0) \mathbf{e}_2 = \mathbf{e}_1, \quad J_k(0) \mathbf{e}_1 = \mathbf{0}.$$

Zúženie $\psi|_{\mathcal{W}'}$ zobrazenia ψ na k -rozmerný podpriestor $\mathcal{W}'$ priestoru $\mathcal{W}$ invariantný vzhľadom na ψ má teda vzhľadom na nejakú bázu $\mathcal{B}' = (\mathbf{w}'_1, \dots, \mathbf{w}'_k)$ tohto podpriestoru maticu $A_{\psi|_{\mathcal{W}'}}[\mathcal{B}'] = J_k(0)$ práve vtedy, keď

$$\psi(\mathbf{w}'_k) = \mathbf{w}'_{k-1}, \quad \psi(\mathbf{w}'_{k-1}) = \mathbf{w}'_{k-2}, \quad \dots, \quad \psi(\mathbf{w}'_2) = \mathbf{w}'_1, \quad \psi(\mathbf{w}'_1) = \mathbf{0}, \quad (4.12)$$

čo je to isté ako

$$\psi(\mathbf{w}'_k) = \mathbf{w}'_{k-1}, \quad \psi^2(\mathbf{w}'_k) = \mathbf{w}'_{k-2}, \quad \dots, \quad \psi^{k-1}(\mathbf{w}'_k) = \mathbf{w}'_1, \quad \psi^k(\mathbf{w}'_k) = \mathbf{0}.$$

V takom prípade budeme hovoriť, že vektory $\mathbf{w}'_1, \dots, \mathbf{w}'_k$ tvoria *Jordanov reťazec* a rovnosti (4.12) budeme skrátene zapisovať ako

$$\psi: \mathbf{w}'_k \mapsto \mathbf{w}'_{k-1} \mapsto \mathbf{w}'_{k-2} \mapsto \dots \mapsto \mathbf{w}'_1 \mapsto \mathbf{0}.$$

Podpriestor priestoru $\mathcal{W}$ s bázou tvoriacou Jordanov reťazec pre ψ musí zjavne byť vzhľadom na ψ invariantný. Nilpotentné zobrazenie ψ tak má vzhľadom na nejakú bázu $\mathcal{B}$ maticu $A_\psi[\mathcal{B}]$ v Jordanovom kanonickom tvare práve vtedy, keď možno bázu $\mathcal{B}$ rozložiť na niekoľko takýchto Jordanových reťazcov; ku každému reťazcu o k vektoroch bude zodpovedať jedna Jordanova bunka rádu k , pričom nutne $k \leq r$, kde r je rád nilpotentného zobrazenia ψ . Bázy s touto vlastnosťou nazveme *Jordanovými bázami*.

Definícia 4.6.5. Nech n je nenulové prirodzené číslo, $\mathcal{W}$ je n -rozmerný vektorový priestor nad $\mathbb{C}$ a $\psi: \mathcal{W} \rightarrow \mathcal{W}$ je nilpotentné lineárne zobrazenie. *Jordanovou bázou* priestoru $\mathcal{W}$ prislúchajúcou k ψ nazveme ľubovoľnú n -ticu lineárne nezávislých vektorov

$$\mathcal{B}_\psi = (\mathbf{w}_1[1], \dots, \mathbf{w}_{n_1}[1], \mathbf{w}_1[2], \dots, \mathbf{w}_{n_2}[2], \dots, \mathbf{w}_1[p], \dots, \mathbf{w}_{n_p}[p]),$$

kde $p \in \mathbb{N}$, $n_1, \dots, n_p \in \mathbb{N} \setminus \{0\}$ sú také, že $n_1 + \dots + n_p = n$ a pre $j = 1, \dots, p$ tvoria vektory $\mathbf{w}_1[j], \dots, \mathbf{w}_{n_j}[j]$ Jordanov reťazec – čiže môžeme písať

$$\psi: \mathbf{w}_{n_j}[j] \mapsto \dots \mapsto \mathbf{w}_1[j] \mapsto \mathbf{0}.$$

Než prejdeme ku konštrukcii Jordanovej bázy k danému nilpotentnému zobrazeniu, zavedme pojem *lineárnej (ne)závislosti vzhľadom na podpriestor*, ktorý sa nám pri neskorších úvahách zide. Prezentácia nasledujúcich partií tohto textu je z veľkej časti založená na [105].

Definícia 4.6.6. Nech $\mathcal{W}$ je vektorový priestor nad $\mathbb{C}$ a nech $\mathcal{S}$ je jeho podpriestor. Nech $q \in \mathbb{N}$ a $\mathbf{x}_1, \dots, \mathbf{x}_q \in \mathcal{W}$. Hovoríme, že vektory $\mathbf{x}_1, \dots, \mathbf{x}_q$ sú *lineárne závislé vzhľadom na $\mathcal{S}$* , ak existujú koeficienty $a_1, \dots, a_q \in \mathbb{C}$ také, že aspoň jeden z nich je nenulový a

$$a_1 \mathbf{x}_1 + \dots + a_q \mathbf{x}_q \in \mathcal{S}.$$

V opačnom prípade hovoríme, že $\mathbf{x}_1, \dots, \mathbf{x}_q$ sú *lineárne nezávislé vzhľadom na $\mathcal{S}$* .

Vektory $\mathbf{x}_1, \dots, \mathbf{x}_q$ sú teda lineárne nezávislé práve vtedy, keď sú lineárne nezávislé vzhľadom na triviálny podpriestor $\{\mathbf{0}\}$. Lineárna nezávislosť vzhľadom na netriviálny podpriestor je silnejšou požiadavkou, než „obyčajná“ lineárna nezávislosť.

Tvrdenie 4.6.7. Nech $\mathcal{W}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$, nech $\mathcal{S}$ je jeho podpriestor, $q \in \mathbb{N}$ a $\mathbf{x}_1, \dots, \mathbf{x}_q \in \mathcal{W}$. Potom sú nasledujúce tvrdenia ekvivalentné:

- (i) Vektory $\mathbf{x}_1, \dots, \mathbf{x}_q$ sú lineárne nezávislé vzhľadom na $\mathcal{S}$.
- (ii) Pre ľubovoľné $s \in \mathbb{N}$ a lineárne nezávislé vektory $\mathbf{v}_1, \dots, \mathbf{v}_s \in \mathcal{S}$ sú vektory $\mathbf{x}_1, \dots, \mathbf{x}_q, \mathbf{v}_1, \dots, \mathbf{v}_s$ lineárne nezávislé.
- (iii) Pre ľubovoľnú bázu $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_t)$ priestoru $\mathcal{S}$ sú $\mathbf{x}_1, \dots, \mathbf{x}_q, \mathbf{v}_1, \dots, \mathbf{v}_t$ lineárne nezávislé.
- (iv) Existuje báza $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_t)$ priestoru $\mathcal{S}$ taká, že $\mathbf{x}_1, \dots, \mathbf{x}_q, \mathbf{v}_1, \dots, \mathbf{v}_t$ sú lineárne nezávislé.

Dôkaz. Začnime *implikáciou* „ak (i), tak (ii)“. Predpokladajme, že $\mathbf{x}_1, \dots, \mathbf{x}_q$ sú lineárne nezávislé vzhľadom na $\mathcal{S}$ a uvažujme lineárne nezávislé vektory $\mathbf{v}_1, \dots, \mathbf{v}_s \in \mathcal{S}$. Nech $a_1, \dots, a_q, b_1, \dots, b_s \in \mathbb{C}$ sú také, že

$$a_1 \mathbf{x}_1 + \dots + a_q \mathbf{x}_q + b_1 \mathbf{v}_1 + \dots + b_s \mathbf{v}_s = \mathbf{0}.$$

Potom

$$a_1 \mathbf{x}_1 + \dots + a_q \mathbf{x}_q = -b_1 \mathbf{v}_1 - \dots - b_s \mathbf{v}_s \in \mathcal{S},$$

pričom z lineárnej nezávislosti vektorov $\mathbf{x}_1, \dots, \mathbf{x}_q$ vzhľadom na $\mathcal{S}$ vyplýva $a_1 = \dots = a_q = 0$. Z toho

$$b_1 \mathbf{v}_1 + \dots + b_s \mathbf{v}_s = \mathbf{0}$$

a z lineárnej nezávislosti vektorov $\mathbf{v}_1, \dots, \mathbf{v}_s$ aj $b_1 = \dots = b_s = 0$. Vektory $\mathbf{x}_1, \dots, \mathbf{x}_q, \mathbf{v}_1, \dots, \mathbf{v}_s$ sú teda skutočne lineárne nezávislé.

Implikácie „ak (ii), tak (iii)“ a „ak (iii), tak (iv)“ sú zrejmé. Dokážeme implikáciu „ak (iv), tak (i)“. Nech $(\mathbf{v}_1, \dots, \mathbf{v}_t)$ je báza priestoru $\mathcal{S}$ taká, že vektory $\mathbf{x}_1, \dots, \mathbf{x}_q, \mathbf{v}_1, \dots, \mathbf{v}_t$ sú lineárne nezávislé. Nech $a_1, \dots, a_q \in \mathbb{C}$ sú také, že

$$a_1 \mathbf{x}_1 + \dots + a_q \mathbf{x}_q \in \mathcal{S}.$$

Potom

$$a_1 \mathbf{x}_1 + \dots + a_q \mathbf{x}_q = b_1 \mathbf{v}_1 + \dots + b_t \mathbf{v}_t$$

pre nejaké $b_1, \dots, b_t \in \mathbb{C}$, z čoho

$$a_1 \mathbf{x}_1 + \dots + a_q \mathbf{x}_q - b_1 \mathbf{v}_1 - \dots - b_t \mathbf{v}_t = \mathbf{0}$$

a z lineárnej nezávislosti vektorov $\mathbf{x}_1, \dots, \mathbf{x}_q, \mathbf{v}_1, \dots, \mathbf{v}_t$ tak okrem iného dostávame $a_1 = \dots = a_q = 0$. Vektory $\mathbf{x}_1, \dots, \mathbf{x}_q$ sú teda lineárne nezávislé vzhľadom na $\mathcal{S}$. $\square$

Pristúpme teraz k samotnej konštrukcii Jordanovej bázy prislúchajúcej k danému nilpotentnému lineárnemu zobrazeniu.

Lema 4.6.8. *Nech $\mathcal{W}$ je netriviálny konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\psi: \mathcal{W} \rightarrow \mathcal{W}$ je nilpotentné lineárne zobrazenie. Potom existuje Jordanova báza $\mathcal{B}_\psi$ priestoru $\mathcal{W}$ prislúchajúca k ψ .*

Dôkaz. Nech $r \in \mathbb{N} \setminus \{0\}$ je rád nilpotentného zobrazenia ψ . Skonštruujeme Jordanovu bázu

$$\mathcal{B}_\psi = (\mathbf{w}_1[1], \dots, \mathbf{w}_{n_1}[1], \mathbf{w}_1[2], \dots, \mathbf{w}_{n_2}[2], \dots, \mathbf{w}_1[p], \dots, \mathbf{w}_{n_p}[p])$$

priestoru $\mathcal{W}$ prislúchajúcu k zobrazeniu ψ takú, že $p \in \mathbb{N}$, $n_1 \geq n_2 \geq \dots \geq n_p$ a pre $j = 1, \dots, p$ tvoria bázi vektory $\mathbf{w}_1[j], \dots, \mathbf{w}_{n_j}[j]$ Jordanov reťazec:

$$\psi: \mathbf{w}_{n_j}[j] \mapsto \dots \mapsto \mathbf{w}_1[j] \mapsto \mathbf{0}.$$

Jednotlivé Jordanove reťazce tak budú v konštruovanej Jordanovej báze vystupovať v nerastúcom poradí podľa počtu vektorov, ktoré ich tvoria.

Keďže je ψ nilpotentné zobrazenie rádu r , každý Jordanov reťazec konštruovanej Jordanovej bázy môže byť dĺžky najviac r . Pre $s = 1, \dots, r$ budeme označovať ako $\ell(s)$ počet tých reťazcov konštruovanej Jordanovej bázy, ktoré sú dĺžky *aspoň* s . Samotnú konštrukciu bázičských vektorov pritom zrealizujeme v nasledujúcom poradí: postupne pre $s = r, \dots, 1$ špecifikujeme počet $\ell(s)$ všetkých reťazcov dĺžky *aspoň* s a následne aj všetky vektory $\mathbf{w}_s[j]$ pre $j = 1, \dots, \ell(s)$ – čiže pre všetky j také, že j -ty reťazec

iste $\mathbf{w}_s[j] \in \text{Ker}(\psi^s)$. Musíme dokázať, že takto definované vektory $\mathbf{w}_s[1], \dots, \mathbf{w}_s[\ell(s+1)]$ sú lineárne nezávislé vzhľadom na $\text{Ker}(\psi^{s-1})$. Uvažujme teda koeficienty $a_1, \dots, a_{\ell(s+1)} \in \mathbb{C}$ také, že

$$a_1 \mathbf{w}_s[1] + \dots + a_{\ell(s+1)} \mathbf{w}_s[\ell(s+1)] \in \text{Ker}(\psi^{s-1}).$$

Je

$$\begin{aligned} a_1 \mathbf{w}_s[1] + \dots + a_{\ell(s+1)} \mathbf{w}_s[\ell(s+1)] &= a_1 \psi(\mathbf{w}_{s+1}[1]) + \dots + a_{\ell(s+1)} \psi(\mathbf{w}_{s+1}[\ell(s+1)]) = \\ &= \psi(a_1 \mathbf{w}_{s+1}[1] + \dots + a_{\ell(s+1)} \mathbf{w}_{s+1}[\ell(s+1)]) \end{aligned}$$

a z príslušnosti tohto vektora do $\text{Ker}(\psi^{s-1})$ vyplýva, že

$$a_1 \mathbf{w}_{s+1}[1] + \dots + a_{\ell(s+1)} \mathbf{w}_{s+1}[\ell(s+1)] \in \text{Ker}(\psi^s).$$

Z lineárnej nezávislosti vektorov $\mathbf{w}_{s+1}[1], \dots, \mathbf{w}_{s+1}[\ell(s+1)]$ vzhľadom na $\text{Ker}(\psi^s)$ teda dostávame

$$a_1 = \dots = a_{\ell(s+1)} = 0$$

a vektory $\mathbf{w}_s[1], \dots, \mathbf{w}_s[\ell(s+1)] \in \text{Ker}(\psi^s)$ sú skutočne lineárne nezávislé vzhľadom na $\text{Ker}(\psi^{s-1})$.

Ak už samotné tieto vektory dopĺňajú nejakú bázu priestoru $\text{Ker}(\psi^{s-1})$ na bázu priestoru $\text{Ker}(\psi^s)$, stačí položiť $\ell(s) := \ell(s+1)$ a zrejme $\ell(s) = \dim(\text{Ker}(\psi^s)) - \dim(\text{Ker}(\psi^{s-1}))$. V opačnom prípade existuje nenulové prirodzené číslo ℓ a vektory $\mathbf{w}[1], \dots, \mathbf{w}[\ell]$ príslušné do $\text{Ker}(\psi^s) \setminus \text{Ker}(\psi^{s-1})$ také, že vektory

$$\mathbf{w}_s[1], \dots, \mathbf{w}_s[\ell(s+1)], \mathbf{w}[1], \dots, \mathbf{w}[\ell]$$

dopĺňajú nejakú bázu priestoru $\text{Ker}(\psi^{s-1})$ na bázu priestoru $\text{Ker}(\psi^s)$. V takom prípade sú tieto vektory lineárne nezávislé vzhľadom na $\text{Ker}(\psi^{s-1})$ a môžeme položiť $\ell(s) := \ell(s+1) + \ell$ a $\mathbf{w}_s[\ell(s+1) + j] := \mathbf{w}[j]$ pre $j = 1, \dots, \ell$. Zrejme tiež potom $\ell(s) = \dim(\text{Ker}(\psi^s)) - \dim(\text{Ker}(\psi^{s-1}))$. Týmto je konštrukcia vektorov z $\mathcal{B}_\psi$ dokončená.

Zostáva dokázať, že takto skonštruované vektory skutočne tvoria Jordanovu bázu priestoru $\mathcal{W}$ prislúchajúcu k zobrazeniu ψ . Z konštrukcie je zrejmé, že $\mathcal{B}_\psi$ možno rozložiť na niekoľko Jordanových reťazcov; zostáva teda ukázať, že skutočne ide o bázu priestoru $\mathcal{W}$. Keďže sú však pre $j = 1, \dots, r$ vektory

$$\mathbf{w}_j[1], \dots, \mathbf{w}_j[\ell(j)] \in \text{Ker}(\psi^j)$$

lineárne nezávislé vzhľadom na podpriestor $\text{Ker}(\psi^{j-1})$, z tvrdenia 4.6.7 ľahko vidieť, že vektory z $\mathcal{B}_\psi$ musia byť lineárne nezávislé. Skonštruovali sme navyše presne

$$\sum_{s=1}^r \ell(s) = \sum_{s=1}^r (\dim(\text{Ker}(\psi^s)) - \dim(\text{Ker}(\psi^{s-1}))) = \dim(\text{Ker}(\psi^r)) - \dim(\text{Ker}(\psi^0)) = \dim(\mathcal{W})$$

vektorov $\mathcal{B}_\psi$; musí teda naozaj ísť o bázu priestoru $\mathcal{W}$. □

Dôsledok 4.6.9. *Nech $\mathcal{W}$ je netriviálny konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\psi: \mathcal{W} \rightarrow \mathcal{W}$ je nilpotentné lineárne zobrazenie. Potom existuje báza $\mathcal{B}$ priestoru $\mathcal{W}$ taká, že matica $A_\psi[\mathcal{B}]$ zobrazenia ψ vzhľadom na $\mathcal{B}$ je v Jordanovom kanonickom tvare, pričom každá Jordanova bunka matice $A_\psi[\mathcal{B}]$ je rovná $J_s(0)$ pre nejaké $s \in \mathbb{N}$ spĺňajúce $1 \leq s \leq \dim(\mathcal{W})$.*

Dôkaz. Podľa lemy 4.6.8 existuje Jordanova báza $\mathcal{B}_\psi$ priestoru $\mathcal{W}$ prislúchajúca k zobrazeniu ψ . Z úvah predchádzajúcich definícii 4.6.5 zisťujeme, že matica $A_\psi[\mathcal{B}_\psi]$ je v Jordanovom kanonickom tvare, pričom jej Jordanove bunky sú všetky uvedeného typu. □

Môžeme prejsť k hlavnému výsledku tohto oddielu, ktorého dôkaz už bude len zozbieraním faktov známych z predchádzajúceho textu. Opätovne sa vrátíme k všeobecným – čiže nie nutne nilpotentným – lineárnym zobrazeniam φ na konečnorozmernom vektorovom priestore $\mathcal{V}$ a ukážeme, že ku každému takémuto φ možno nájsť bázu, vzhľadom na ktorú má φ maticu v Jordanovom kanonickom tvare. V dôsledku toho je každá štvorcová matica nad $\mathbb{C}$ podobná nejakej matici v Jordanovom kanonickom tvare.

Veta 4.6.10. *Nech $\mathcal{V}$ je konečnorozmerný vektorový priestor nad $\mathbb{C}$ a $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie. Potom existuje báza $\mathcal{B}$ priestoru $\mathcal{V}$ pozostávajúca zo zovšeobecnených vlastných vektorov zobrazenia φ taká, že matica $A_\varphi[\mathcal{B}]$ zobrazenia φ vzhľadom na $\mathcal{B}$ je v Jordanovom kanonickom tvare, pričom každá jej Jordanova bunka je rovná $J_s(\lambda)$ pre nejaké vlastné číslo λ zobrazenia φ a $s \in \mathbb{N}$ také, že $1 \leq s \leq \alpha(\varphi, \lambda)$.*

Dôkaz. Nech $\lambda_1, \dots, \lambda_m$ sú práve všetky po dvoch rôzne vlastné čísla zobrazenia φ . Lineárne zobrazenie

$$\nu_j := \left(\varphi|_{G(\varphi, \lambda_j)} - \lambda_j \text{Id}_{G(\varphi, \lambda_j)} \right) : G(\varphi, \lambda_j) \rightarrow G(\varphi, \lambda_j)$$

na netriviálnom priestore $G(\varphi, \lambda_j)$ je potom pre $j = 1, \dots, m$ nilpotentné. Podľa dôsledku 4.6.9 teda má, vzhľadom na nejakú bázu $\mathcal{B}_j$ priestoru $G(\varphi, \lambda_j)$, maticu $A_{\nu_j}[\mathcal{B}_j]$ v Jordanovom kanonickom tvare. V Jordanovom kanonickom tvare je tak aj matica $A_{\varphi|_{G(\varphi, \lambda_j)}}[\mathcal{B}_j]$ zobrazenia

$$\varphi|_{G(\varphi, \lambda_j)} = \nu_j + \lambda_j \text{Id}_{G(\varphi, \lambda_j)},$$

ktorá je daná ako

$$A_{\varphi|_{G(\varphi, \lambda_j)}}[\mathcal{B}_j] = A_{\nu_j}[\mathcal{B}_j] + \lambda_j \mathbf{I}.$$

Podľa dôsledku 4.6.9 sú všetky Jordanove bunky matice $A_{\nu_j}[\mathcal{B}_j]$ rovné $J_s(0)$ pre nejaké $s \in \mathbb{N}$ spĺňajúce $1 \leq s \leq \dim(G(\varphi, \lambda_j))$. Z vety 4.5.10 navyše vieme, že $\dim(G(\varphi, \lambda_j)) = \alpha(\varphi, \lambda_j)$. Jordanove bunky matice $A_{\varphi|_{G(\varphi, \lambda_j)}}[\mathcal{B}_j]$ sú teda všetky typu $J_s(\lambda_j)$ pre $1 \leq s \leq \alpha(\varphi, \lambda_j)$.

Predpokladajme teraz, že pre $j = 1, \dots, m$ je $\mathcal{B}_j = (\mathbf{v}_{j,1}, \dots, \mathbf{v}_{j,s_j})$. Vďaka vete 4.5.10 možno priestor $\mathcal{V}$ rozložiť na priamy súčet

$$\mathcal{V} = G(\varphi, \lambda_1) \oplus G(\varphi, \lambda_2) \oplus \dots \oplus G(\varphi, \lambda_m)$$

a každý z podpriestorov $G(\varphi, \lambda_1), \dots, G(\varphi, \lambda_m)$ je invariantný vzhľadom na φ . Pre bázu

$$\mathcal{B} = (\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,s_1}, \dots, \mathbf{v}_{m,1}, \dots, \mathbf{v}_{m,s_m})$$

priestoru $\mathcal{V}$ teda nutne dostávame

$$A_\varphi[\mathcal{B}] = \begin{pmatrix} A_{\varphi|_{G(\varphi, \lambda_1)}}[\mathcal{B}_1] & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & A_{\varphi|_{G(\varphi, \lambda_2)}}[\mathcal{B}_2] & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & A_{\varphi|_{G(\varphi, \lambda_m)}}[\mathcal{B}_m] \end{pmatrix}.$$

To je však opäť matica v Jordanovom kanonickom tvare a jej Jordanove bunky sú typu uvedeného v znení vety. Báza $\mathcal{B}$ navyše pozostáva z nenulových vektorov z množiny

$$G(\varphi, \lambda_1) \cup G(\varphi, \lambda_2) \cup \dots \cup G(\varphi, \lambda_m),$$

a teda musí byť tvorená zovšeobecnenými vlastnými vektormi zobrazenia φ . □

Dôsledok 4.6.11. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{C}^{n \times n}$ je matica. Potom je A podobná nejakej matici J v Jordanovom kanonickom tvare, ktorej Jordanove bunky sú všetky typu $J_s(\lambda)$ pre $\lambda \in \sigma(A)$ a $1 \leq s \leq \alpha(A, \lambda)$. Stĺpce nesingulárnej matice P spĺňajúcej*

$$P^{-1}AP = J$$

sú navyše lineárne nezávislými zovšeobecnenými vlastnými vektormi matice A tvoriacimi bázu $\mathbb{C}^n$.

Dôkaz. Podľa vety 4.6.10 existuje báza $\mathcal{B} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ priestoru $\mathbb{C}^n$ tvorená zovšeobecnenými vlastnými vektormi zobrazenia φ_A taká, že zobrazenie φ_A má vzhľadom na bázu $\mathcal{B}$ maticu $J = A_{\varphi_A}[\mathcal{B}]$ v Jordanovom kanonickom tvare. Pre maticu $A = A_{\varphi_A}[\mathcal{B}_{st}(\mathbb{C}^n)]$ ale potom

$$J = A_{\varphi_A}[\mathcal{B}] = P[\mathcal{B} \leftarrow \mathcal{B}_{st}(\mathbb{C}^n)]A_{\varphi_A}[\mathcal{B}_{st}(\mathbb{C}^n)]P[\mathcal{B}_{st}(\mathbb{C}^n) \leftarrow \mathcal{B}] = P^{-1}AP,$$

kde P je matica so stĺpcami $\mathbf{v}_1, \dots, \mathbf{v}_n$. Ak je naopak P nesingulárna matica taká, že $J = P^{-1}AP$ je v Jordanovom kanonickom tvare, musí pre bázu $\mathcal{B}$ priestoru $\mathbb{C}^n$ tvorenú stĺpcami matice P nutne byť $J = A_{\varphi_A}[\mathcal{B}]$ a stĺpce matice P sú tak evidentne zovšeobecnenými vlastnými vektormi matice A . $\square$

Pri väčšine aplikácií Jordanovho kanonického tvaru, s ktorými sa v rámci tohto textu stretneme, nám bude popri nájdenom spektre matice stačiť poznanie, že každá štvorcová matica je nejakej matici v Jordanovom kanonickom tvare podobná – samotný výpočet Jordanovho kanonického tvaru matice, ktorý môže byť občas trochu prácný, budeme obyčajne obchádzať jednoduchšími metódami. Aplikácie Jordanovho kanonického tvaru, ktorými sa budeme zaoberať, tak budú trochu atypického charakteru – napriek tomu, že nám umožnia získať celkom hmatateľné výsledky, bude povedomie o možnosti realizovať práve opísané konštrukcie pri nich podstatnejšie, než tieto konštrukcie samotné.

Občas však môže byť transformácia matice do Jordanovho kanonického tvaru nevyhnutná – je preto užitočné ujasniť si, ako ju možno realizovať. Postup, ktorého správnosť vyplýva z dôkazov lemy 4.6.8 a vety 4.6.10, môže byť pre maticu $A \in \mathbb{C}^{n \times n}$ nasledujúci:

1. Nájdi všetky po dvoch rôzne vlastné čísla $\lambda_1, \dots, \lambda_m$ matice A – čiže korene charakteristického polynómu $\chi_A(z)$.
2. Opakuj pre $\lambda = \lambda_1, \dots, \lambda_m$:
 - 2.1. Pomocou Gaussovej eliminácie konštruuj matice $B_1, B_2, B_3, \dots$ v redukovanom stupňovitom tvare reprezentujúce bázy priestorov $\text{Ker}(A - \lambda \mathbf{I}), \text{Ker}((A - \lambda \mathbf{I})^2), \text{Ker}((A - \lambda \mathbf{I})^3), \dots$, až kým pre nejaké $r \in \mathbb{N} \setminus \{0\}$ bude $B_r = B_{r+1}$.
 - 2.2. Pomocou Gaussovej eliminácie nájdi vektory $\mathbf{w}[1], \dots, \mathbf{w}[\ell] \in \text{Ker}((A - \lambda \mathbf{I})^r)$, ktoré dopĺňajú bázu priestoru $\text{Ker}((A - \lambda \mathbf{I})^{r-1})$ reprezentovanú maticou B_{r-1} na bázu $\text{Ker}((A - \lambda \mathbf{I})^r)$. Polož $\ell(r) := \ell$ a $\mathbf{w}_r[j] := \mathbf{w}[j]$ pre $j = 1, \dots, \ell(r)$.
 - 2.3. Opakuj pre $s = r - 1, \dots, 1$:
 - 2.3.1. Pre $j = 1, \dots, \ell(s + 1)$ polož $\mathbf{w}_s[j] := (A - \lambda \mathbf{I})\mathbf{w}_{s+1}[j]$.
 - 2.3.2. Pomocou Gaussovej eliminácie nájdi vektory $\mathbf{w}[1], \dots, \mathbf{w}[\ell] \in \text{Ker}((A - \lambda \mathbf{I})^s)$ také, že vektory $\mathbf{w}_s[1], \dots, \mathbf{w}_s[\ell(s + 1)], \mathbf{w}[1], \dots, \mathbf{w}[\ell]$ dopĺňajú bázu priestoru $\text{Ker}((A - \lambda \mathbf{I})^{s-1})$ reprezentovanú maticou B_{s-1} na bázu priestoru $\text{Ker}((A - \lambda \mathbf{I})^s)$. Polož $\ell(s) := \ell(s + 1) + \ell$ a $\mathbf{w}_s[\ell(s + 1) + j] := \mathbf{w}[j]$ pre $j = 1, \dots, \ell$.
 - 2.4. Nech $p = \ell(1)$ a n_j je pre $j = 1, \dots, p$ najväčšie prirodzené číslo také, že $j \leq \ell(n_j)$. Nech $\mathcal{B}_\lambda$ je Jordanova báza $\mathcal{B}_\lambda = (\mathbf{w}_1[1], \dots, \mathbf{w}_{n_1}[1], \dots, \mathbf{w}_1[p], \dots, \mathbf{w}_{n_p}[p])$ a

$$J_\lambda = \begin{pmatrix} J_{n_1}(\lambda) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & J_{n_2}(\lambda) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & J_{n_p}(\lambda) \end{pmatrix}.$$

3. Skonštruuj maticu P s blokovým rozkladom $P = (\mathcal{B}_{\lambda_1} \quad \mathcal{B}_{\lambda_2} \quad \dots \quad \mathcal{B}_{\lambda_m})$. Matica J v Jordanovom kanonickom tvare podobná matici A je potom daná ako

$$J = P^{-1}AP = \begin{pmatrix} J_{\lambda_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & J_{\lambda_2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & J_{\lambda_m} \end{pmatrix}.$$

4.7 Mocniny matíc v Jordanovom kanonickom tvare

Veľká časť aplikácií diagonalizácie matíc spočíva na skutočnosti, že pre diagonálne matice je ľahké vyjadriť v uzavretom tvare prvky ich t -tych mocnín pre $t \in \mathbb{N}$ – o tom sme sa už koniec koncov mali možnosť presvedčiť v oddiele 4.3. Ukážeme teraz, že prvky t -tej mocniny možno v uzavretom tvare bez väčších ťažkostí vyjadriť aj pre každú maticu v Jordanovom kanonickom tvare. To nám v nasledujúcich dvoch kapitolách otvorí cestu k použitiu metód z oddielu 4.3 pre ľubovoľnú maticu.

Veta 4.7.1. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $\lambda \in \mathbb{C}$. Pre Jordanovu bunku*

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

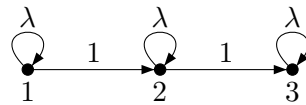
a všetky $t \in \mathbb{N}$ potom, pri konvencii $\binom{t}{k} 0^{t-k} = 0$ pre všetky prirodzené čísla $t < k$,

$$J_n(\lambda)^t = \begin{pmatrix} \lambda^t & \binom{t}{1}\lambda^{t-1} & \binom{t}{2}\lambda^{t-2} & \dots & \binom{t}{n-1}\lambda^{t-n+1} \\ 0 & \lambda^t & \binom{t}{1}\lambda^{t-1} & \dots & \binom{t}{n-2}\lambda^{t-n+2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda^t & \binom{t}{1}\lambda^{t-1} \\ 0 & 0 & \dots & 0 & \lambda^t \end{pmatrix}.$$

Pre $J_n(\lambda)^t = (a_{i,j}[t])_{n \times n}$ a $i, j = 1, \dots, n$ je teda

$$a_{i,j}[t] = \begin{cases} \binom{t}{j-i}\lambda^{t-(j-i)} & \text{ak } i \leq j, \\ 0 & \text{ak } i > j. \end{cases}$$

Dôkaz. Predpokladajme najprv, že $\lambda \neq 0$ a uvažujme graf $\text{Gr}(J_n(\lambda))$ Jordanovej bunky $J_n(\lambda)$. Ide o ohodnotený graf nad $\mathbb{C}$ s množinou vrcholov $[n]$, ktorý je pre $n = 3$ znázornený na obrázku 4.3. Pozostáva z hrán (k, k) pre $k = 1, \dots, n$ ohodnotených číslom λ a z hrán $(k, k+1)$ pre $k = 1, \dots, n-1$ ohodnotených číslom 1. Hrany prvého typu budeme v súlade s bežnou grafovou terminológiou volať *slučkami* a hrany druhého typu nazveme *doprednými hranami*.



Obr. 4.3: Graf $\text{Gr}(J_3(\lambda))$ Jordanovej bunky $J_3(\lambda)$.

Z vety 1.3.10 vyplýva, že $a_{i,j}[t]$ je pre $i, j = 1, \dots, n$ súčtom ohodnotení sledov dĺžky t vedúcich v grafe $\text{Gr}(J_n(\lambda))$ z vrcholu i do vrcholu j . Ak $i > j$, zjavne neexistuje žiaden takýto sled, a teda $a_{i,j}[t] = 0$. Ak $i \leq j$, pozostáva každý takýto sled dĺžky $t \geq j - i$ z presne $j - i$ dopredných hrán a z $t - (j - i)$ slučiek. Ohodnotením každého takéhoto sledu je teda $\lambda^{t-(j-i)}$ a každý takýto sled je jednoznačne určený pozíciami svojich dopredných hrán – dohromady ich teda je $\binom{t}{j-i}$. Pre $i \leq j$ a všetky $t \in \mathbb{N}$ tak dostávame

$$a_{i,j}[t] = \binom{t}{j-i}\lambda^{t-(j-i)},$$

čo dokazuje vetu pre $\lambda \neq 0$. Pre $\lambda = 0$ možno argumentovať analogicky s tým, že graf $\text{Gr}(J_n(0))$ obsahuje iba dopredné hrany; netreba pritom zabúdať na to, že $0^0 = 1$. $\square$

Každá matica

$$A = \begin{pmatrix} J_{n_1}(\lambda_1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & J_{n_2}(\lambda_2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & J_{n_k}(\lambda_k) \end{pmatrix}$$

v Jordanovom kanonickom tvare je blokovo diagonálna a všetky jej diagonálne bloky sú Jordanove bunky. Jej t -tu mocninu tak možno vyjadriť ako

$$A^t = \begin{pmatrix} J_{n_1}(\lambda_1)^t & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & J_{n_2}(\lambda_2)^t & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & J_{n_k}(\lambda_k)^t \end{pmatrix},$$

kde $J_{n_1}(\lambda_1)^t, \dots, J_{n_k}(\lambda_k)^t$ sú dané podľa vety 4.7.1.

Podobné myšlienky možno využiť aj pri výpočte niektorých maticových funkcií definovaných pomocou Maclaurinových radov, do ktorých sa pri ich vyhodnocovaní za premennú namiesto čísla dosadí matica – pre ľubovoľnú maticu $A \in \mathbb{C}^{n \times n}$ napríklad možno hovoriť o maticiach ako e^A , $\cos A$, alebo $\sin A$, pričom podobne ako pri umocňovaní sa pre matice v Jordanovom kanonickom tvare výpočet týchto funkcií značne zjednoduší. Viac sa možno dočítať napríklad v [105].

4.8 Spektrálne vlastnosti jednoduchých operácií na maticiach

Preskúmame teraz vplyv niekoľkých jednoduchých maticových operácií na charakteristické polynómy a spektrá matíc, ako aj na algebraické a geometrické násobnosti jednotlivých vlastných čísel. Tieto vlastnosti sa nám zídu v neskorších kapitolách.

Veta 4.8.1. *Nech n je prirodzené číslo a $A \in \mathbb{C}^{n \times n}$ je matica. Potom*

$$\text{ch}_{A^T}(z) = \text{ch}_A(z),$$

$\sigma(A^T) = \sigma(A)$ a pre každé $\lambda \in \sigma(A)$ je $\alpha(A^T, \lambda) = \alpha(A, \lambda)$ a $\gamma(A^T, \lambda) = \gamma(A, \lambda)$.

Dôkaz. Platí $\text{ch}_{A^T}(z) = \det(z\mathbf{I} - A^T) = \det((z\mathbf{I} - A)^T) = \det(z\mathbf{I} - A) = \text{ch}_A(z)$. Z toho bezprostredne vyplývajú aj tvrdenia o rovnosti spektier a algebraických násobností vlastných čísel. Pri geometrických násobnostiach si stačí uvedomiť, že pre každé vlastné číslo λ je

$$\begin{aligned} \gamma(A^T, \lambda) &= \dim(E(A^T, \lambda)) = \dim(\text{Ker}(A^T - \lambda\mathbf{I})) = \dim(\text{Ker}((A - \lambda\mathbf{I})^T)) = \\ &= \dim(\text{Ker}(A - \lambda\mathbf{I})) = \dim(E(A, \lambda)) = \gamma(A, \lambda), \end{aligned}$$

keďže hodnosť štvorcovej matice a jej transpozície si musia byť rovné. □

Veta 4.8.2. *Nech n je prirodzené číslo, $A \in \mathbb{C}^{n \times n}$ je matica a $q \in \mathbb{C} \setminus \{0\}$. Potom*

$$\text{ch}_{qA}(z) = q^n \text{ch}_A(z/q),$$

$\sigma(qA) = \{q\lambda \mid \lambda \in \sigma(A)\}$ a pre každé $\lambda \in \sigma(A)$ je $\alpha(qA, q\lambda) = \alpha(A, \lambda)$ a $\gamma(qA, q\lambda) = \gamma(A, \lambda)$.

Dôkaz. Platí $\text{ch}_{qA}(z) = \det(z\mathbf{I} - qA) = q^n \det((z/q)\mathbf{I} - A) = q^n \text{ch}_A(z/q)$. Tým sú dokázané aj tvrdenia o spektre a algebraických násobnostiach. Pri geometrických násobnostiach zisťujeme, že

$$\begin{aligned} \gamma(qA, q\lambda) &= \dim(E(qA, q\lambda)) = \dim(\text{Ker}(qA - q\lambda\mathbf{I})) = \dim(\text{Ker}(q(A - \lambda\mathbf{I}))) = \\ &= \dim(\text{Ker}(A - \lambda\mathbf{I})) = \dim(E(A, \lambda)) = \gamma(A, \lambda), \end{aligned}$$

keďže hodnosť matice je rovnaká ako hodnosť jej ľubovoľného nenulového skalárneho násobku. □

Veta 4.8.3. *Nech n je prirodzené číslo, $A \in \mathbb{C}^{n \times n}$ je matica a $k \in \mathbb{C}$. Potom*

$$\text{ch}_{A+k\mathbf{I}}(z) = \text{ch}_A(z - k),$$

$\sigma(A + k\mathbf{I}) = \{\lambda + k \mid \lambda \in \sigma(A)\}$ a pre každé vlastné číslo $\lambda \in \sigma(A)$ je $\alpha(A + k\mathbf{I}, \lambda + k) = \alpha(A, \lambda)$ a $\gamma(A + k\mathbf{I}, \lambda + k) = \gamma(A, \lambda)$.

Dôkaz. Evidentne $\text{ch}_{A+k\mathbf{I}}(z) = \det(z\mathbf{I} - (A + k\mathbf{I})) = \det((z - k)\mathbf{I} - A) = \text{ch}_A(z - k)$, čím sú dokázané aj tvrdenia o spektre a algebraických násobnostiach. Pre geometrické násobnosti navyše

$$\begin{aligned} \gamma(A + k\mathbf{I}, \lambda + k) &= \dim(E(A + k\mathbf{I}, \lambda + k)) = \dim(\text{Ker}(A + k\mathbf{I} - (\lambda + k)\mathbf{I})) = \\ &= \dim(\text{Ker}(A - \lambda\mathbf{I})) = \dim(E(A, \lambda)) = \gamma(A, \lambda). \end{aligned}$$

□

Veta 4.8.4. *Nech n je prirodzené číslo, $A \in \mathbb{C}^{n \times n}$ je matica a $t \in \mathbb{N}$. Potom*

$$\text{ch}_{A^t}(z) = \prod_{\lambda \in \sigma(A)} (z - \lambda^t)^{\alpha(A, \lambda)},$$

$\sigma(A^t) = \{\lambda^t \mid \lambda \in \sigma(A)\}$ a pre všetky $\kappa \in \sigma(A^t)$ je

$$\alpha(A^t, \kappa) = \sum_{\substack{\lambda \in \sigma(A) \\ \lambda^t = \kappa}} \alpha(A, \lambda).$$

Dôkaz. Nech $J = (b_{i,j})_{n \times n}$ je matica v Jordanovom kanonickom tvare taká, že pre nejakú nesingulárnu štvorcovú maticu P je $A = PJP^{-1}$. Označme diagonálne prvky matice J pre $k = 1, \dots, n$ postupne ako $\lambda_k := b_{k,k}$. Potom

$$\text{ch}_A(z) = \text{ch}_J(z) = (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_n),$$

a teda $\sigma(A) = \sigma(J) = \{\lambda_1, \dots, \lambda_n\}$. Ďalej $A^t = PJ^tP^{-1}$, pričom diagonálne prvky hornej trojuholníkovvej matice J^t sú vďaka vete 4.7.1 dané ako $\lambda_1^t, \lambda_2^t, \dots, \lambda_n^t$. Preto

$$\text{ch}_{A^t}(z) = \text{ch}_{J^t}(z) = (z - \lambda_1^t)(z - \lambda_2^t) \dots (z - \lambda_n^t) = \prod_{\lambda \in \sigma(A)} (z - \lambda^t)^{\alpha(A, \lambda)},$$

z čoho $\sigma(A^t) = \{\lambda_1^t, \dots, \lambda_n^t\} = \{\lambda^t \mid \lambda \in \sigma(A)\}$ a algebraická násobnosť každého vlastného čísla κ matice A^t je daná súčtom algebraických násobností vlastných čísel λ matice A takých, že $\lambda^t = \kappa$. □

4.9 Vlastné čísla a stopa matice

Našu exkurziu do elementárnej lineárnej algebry teraz zakončíme dôkazom jednoduchého pozorovania, dávajúceho do súvisu vlastné čísla štvorcovej matice A so súčtom jej diagonálnych prvkov – ten nazývame aj *stopou* matice A .

Definícia 4.9.1. Nech $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ je štvorcová matica. *Stopou* matice A nazveme číslo

$$\text{tr}(A) = a_{1,1} + a_{2,2} + \dots + a_{n,n}.$$

Tvrdenie 4.9.2. *Nech $n \in \mathbb{N}$ a nech $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ a $B = (b_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ sú štvorcové matice. Potom $\text{tr}(AB) = \text{tr}(BA)$.*

Dôkaz. Priamo z definícií stopy a maticového násobenia je

$$\text{tr}(AB) = \sum_{j=1}^n \sum_{k=1}^n a_{j,k} b_{k,j} = \sum_{k=1}^n \sum_{j=1}^n b_{k,j} a_{j,k} = \text{tr}(BA).$$

□

Dokážeme teraz, že stopa štvorcovej matice je vždy rovná súčtu všetkých jej vlastných čísel, kde každé z nich je zarátané aj viackrát podľa jeho algebraickej násobnosti.

Veta 4.9.3. *Nech $n \in \mathbb{N}$ a $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ je štvorcová matica. Potom*

$$\operatorname{tr}(A) = \sum_{\lambda \in \sigma(A)} \alpha(A, \lambda) \lambda.$$

Dôkaz. Tvrdenie je zrejmé pre matice v Jordanovom kanonickom tvare. Vďaka dôsledku 4.6.11 je ale každá matica A s nejakou maticou $J \in \mathbb{C}^{n \times n}$ v Jordanovom kanonickom tvare podobná – pre nejakú neregulárnu maticu $P \in \mathbb{C}^{n \times n}$ je $A = PJP^{-1}$. S použitím tvrdenia 4.9.2 teda dostávame

$$\operatorname{tr}(A) = \operatorname{tr}(PJP^{-1}) = \operatorname{tr}(JP^{-1}P) = \operatorname{tr}(J) = \sum_{\lambda \in \sigma(J)} \alpha(J, \lambda) \lambda = \sum_{\lambda \in \sigma(A)} \alpha(A, \lambda) \lambda. \quad \square$$

Cvičenia

1. Dokážte tvrdenie 4.1.8.
2. Dokážte, že podobnosť matíc je reláciou ekvivalencie na $\mathbb{C}^{n \times n}$.
3. Existuje pre každú konečnú množinu $X \subseteq \mathbb{C}$ štvorcová matica A nad $\mathbb{C}$ taká, že $\sigma(A) = X$? Zvoľme ďalej pre každé $\lambda \in X$ nejaké nenulové prirodzené čísla $\gamma(\lambda) \leq \alpha(\lambda)$. Možno nájsť maticu A nad $\mathbb{C}$ takú, že $\sigma(A) = X$ a pre všetky $\lambda \in \sigma(A)$ je $\alpha(A, \lambda) = \alpha(\lambda)$ a $\gamma(A, \lambda) = \gamma(\lambda)$?
4. Nájdite všetky prirodzené čísla $n \geq 2$, pre ktoré existuje matica $A \in \mathbb{R}^{n \times n}$ taká, že $\sigma(A) \subseteq \mathbb{C} \setminus \mathbb{R}$.
5. Nech $n \in \mathbb{N}$. Dokážte, že matica $A \in \mathbb{C}^{n \times n}$ je singulárna práve vtedy, keď $0 \in \sigma(A)$. Dajte v takom prípade do súvisu geometrickú násobnosť $\gamma(A, 0)$ s hodnotou matice A .
6. Nech $n \in \mathbb{N}$. Nech $A \in \mathbb{C}^{n \times n}$ je matica s vlastnými číslami $\lambda_1, \dots, \lambda_n$, kde každé $\lambda \in \sigma(A)$ je uvedené $\alpha(A, \lambda)$ -krát. Dokážte, že $\det(A) = \lambda_1 \cdot \dots \cdot \lambda_n$.
7. Nech $\mathcal{V}$ je netriviálny konečnorozmerný vektorový priestor, $\varphi: \mathcal{V} \rightarrow \mathcal{V}$ je lineárne zobrazenie, λ je jeho vlastné číslo a ℓ je nenulové prirodzené číslo. Označme ako $G_{\leq \ell}(\varphi, \lambda)$ množinu pozostávajúcu z nulového vektora a práve všetkých zovšeobecnených vlastných vektorov zobrazenia φ rádu najviac ℓ prislúchajúcich k λ . Musí $G_{\leq \ell}(\varphi, \lambda)$ tvoriť vektorový podpriestor priestoru $\mathcal{V}$? Ak áno, je tento podpriestor invariantný vzhľadom na φ ? Ako by to bolo s množinou $G_{=\ell}(\varphi, \lambda)$, do ktorej okrem nulového vektora patria iba zovšeobecnené vlastné vektory rádu práve ℓ ?
8. Nech $n \in \mathbb{N} \setminus \{0\}$. Videli sme, že pre každú nilpotentnú štvorcovú maticu $A \in \mathbb{C}^{n \times n}$ musí byť $\sigma(A) = \{0\}$. Dokážte obrátené tvrdenie: ak $\sigma(A) = \{0\}$, je matica A nutne nilpotentná.
9. Nech n je nenulové prirodzené číslo a $A \in \mathbb{C}^{n \times n}$. Bez odvolávania sa na Jordanov kanonický tvar matice dokážte, že pre každý zovšeobecnený vlastný vektor $\mathbf{v}$ matice A rádu 2 prislúchajúci k vlastnému číslu λ a pre všetky $t \in \mathbb{N}$ je

$$A^t \mathbf{v} = \lambda^t \mathbf{v} + t \lambda^{t-1} \mathbf{v}',$$

kde $\mathbf{v}'$ je nejaký vlastný vektor matice A prislúchajúci k λ . Sformulujte a dokážte analogické tvrdenie pre zovšeobecnené vlastné vektory vyššieho rádu.

10. Nech n je nenulové prirodzené číslo a $A \in \mathbb{C}^{n \times n}$. Dokážte, že matica J v Jordanovom kanonickom tvare podobná matici A je určená jednoznačne až na poradie Jordanových buniek na diagonále.

Poznámky a odkazy na samoštúdium

Spomedzi množstva učebníc lineárnej algebry spomeňme predovšetkým tú Zlatošovu [105], ktorou je čiastočne ovplyvnená aj prezentácia tejto kapitoly. V [105] je však teória okolo vlastných čísel a vlastných vektorov spracovaná o niečo všeobecnejšie, pre matice nad poľom. Z tohto dôvodu sa v [105] namiesto o vlastných číslach hovorí o vlastných *hodnotách*; ďalším terminologickým rozdielom je, že namiesto o zovšeobecnených vlastných vektorech [105] hovorí o *koreňových* vektorech. Naznačené sú tam aj možnosti rozšírenia teórie na lineárne zobrazenia (alebo *operátory*) na nekonečnorozmerných priestoroch, ktoré sú však skôr predmetom skúmania funkcionálnej analýzy [72].

Prezentácia tejto kapitoly je do istej miery ovplyvnená aj Axlerovou učebnicou [7]. Do veľkej miery originálnu prezentáciu základov lineárnej algebry zas možno nájsť v Laxovej učebnici [77]. Dôrazom na geometrickú intuíciu a aplikácie sa vyznačujú Hladíkove skriptá [55]. Spomedzi učebníc maticového počtu spomeňme aspoň [57].

Kapitola 5

Vlastné čísla a grafy I

Teóriu z predošlej kapitoly aplikujeme na problémy súvisiace s grafmi. Podobnosť každej štvorcovej matice komplexných čísel s nejakou maticou v Jordanovom kanonickom tvare využijeme – spoločne s poznatkom o tvare mocnín takýchto matíc – na riešenie problému enumerácie sledov dĺžky t medzi danou dvojicou vrcholov orientovaného multigrafu. Ako dôsledok tak získame aj relatívne jednoduchú metódu enumerácie slov v racionálnom jazyku. V oboch prípadoch pôjde o zovšeobecnenia metód z oddielu 4.3 na prípad všeobecných – teda nie nutne diagonalizovateľných – matíc.

Matice susednosti multigrafov sú špeciálnym prípadom nezáporných matíc – čiže matíc, ktorých prvkami sú nezáporné reálne čísla. Ukazuje sa, že o vlastných číslach nezáporných štvorcových matíc toho možno povedať omnoho viac, než o vlastných číslach všeobecných komplexných alebo reálnych štvorcových matíc. Štruktúra spektra takýchto matíc navyše úzko súvisí s ich grafovou interpretáciou – viacero vlastností nezáporných reálnych matíc možno ekvivalentne sformulovať ako vlastnosti príslušných ohodnotených grafov a naopak. Spektrá nezáporných matíc sú hlavným objektom záujmu takzvanej *Perronovej-Frobeniovej teórie*, najdôležitejšie zistenia ktorej si priblížime v rámci oddielu 5.4. Prechod od matíc ku grafom a naopak tam bude zásadného významu.

K súvisu vlastných čísel s grafmi sa ešte vrátíme v kapitole 7, v ktorej načrieme do vybraných partií spektrálnej teórie grafov. Tam už ale budeme skúmať neohodnotené neorientované grafy.

5.1 Vandermondova a zovšeobecnená Vandermondova matica

Než prejdeme k spomínaným kombinatorickým aplikáciám vlastných čísel, zameriame sa na dve triedy špeciálnych štvorcových matíc, variácie ktorých sa v našich neskorších úvahách budú prirodzene vynárať – najprv zavedieme relatívne dobre známy pojem *Vandermondovej matice*, od ktorého sa následne posunieme k definícii o niečo komplikovanejších *zovšeobecnených Vandermondových matíc* [60].

Podstatné pre nás budú predovšetkým vzorce pre determinanty „obyčajných“ a zovšeobecnených Vandermondových matíc,¹ ktoré v nasledujúcom dokážeme.

Definícia 5.1.1. Nech $\mathbb{F}$ je pole. *Vandermondovou maticou* rádu $n \in \mathbb{N}$ nad $\mathbb{F}$ nazveme pre ľubovoľné $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ maticu

$$V(\alpha_1, \dots, \alpha_n) = \begin{pmatrix} \alpha_1^0 & \alpha_2^0 & \dots & \alpha_n^0 \\ \alpha_1^1 & \alpha_2^1 & \dots & \alpha_n^1 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_n^{n-1} \end{pmatrix}.$$

¹Presnejšie pre nás bude podstatná najmä skutočnosť, že sú tieto determinanty pre „nede degenerované“ Vandermondove a zovšeobecnené Vandermondove matice nenulové – takéto matice teda budú vždy nesingulárne. O niečo jednoduchší alternatívny dôkaz tejto skutočnosti, pri ktorom je samotný výpočet determinantu nahradený elementárnym argumentom z teórie diferenciálnych rovníc, neskôr v hrubých rysoch opíšeme v rámci poznámky 6.5.1.

Vandermondova matica je teda matica typu $n \times n$, ktorej k -ty stĺpec pozostáva pre $k = 1, \dots, n$ z prvých n členov geometrickej postupnosti s počiatočným členom 1 a s kvocientom α_k . V literatúre sa často možno stretnúť aj s definíciou „po riadkoch“ [57] – pod Vandermondovou maticou sa v takom prípade rozumie transpozícia Vandermondovej matice v zmysle definície vyššie. Táto nekonzistencia nemá prakticky žiadne vážne dôsledky, pretože najdôležitejší parameter Vandermondovej matice – jej determinant – zostáva pri oboch alternatívnych definíciách nezmenený.

Veta 5.1.2. *Nech $\mathbb{F}$ je pole, $n \in \mathbb{N}$ a $\alpha_1, \dots, \alpha_n \in \mathbb{F}$. Potom*

$$\det(V(\alpha_1, \dots, \alpha_n)) = \prod_{i=1}^n \prod_{j=i+1}^n (\alpha_j - \alpha_i).$$

Ak sú teda prvky $\alpha_1, \dots, \alpha_n$ po dvoch rôzne, je matica $V(\alpha_1, \dots, \alpha_n)$ nesingulárna.

Dôkaz. Ak $\alpha_1, \dots, \alpha_n$ nie sú po dvoch rôzne, obsahuje matica $V(\alpha_1, \dots, \alpha_n)$ dva rovnaké stĺpce – je teda singulárna. Jej determinant je v takom prípade rovný nule, čo súhlasí so znením vety.

Predpokladajme teda, že $\alpha_1, \dots, \alpha_n$ sú po dvoch rôzne a postupujme indukciou vzhľadom na n . Pre $n = 0$ je tvrdenie triviálne pravdivé, pretože determinant matice typu 0×0 , ako aj prázdny súčin sú rovné jednej. Pre $n = 1$ a ľubovoľné $\alpha_1 \in \mathbb{F}$ je $V(\alpha_1) = (1)$, v dôsledku čoho

$$\det(V(\alpha_1)) = 1 = \prod_{i=1}^1 \prod_{j=2}^1 (\alpha_j - \alpha_i).$$

Tým je dokázaná báza indukcie. Nech teraz tvrdenie platí pre $n = q$, kde q je nenulové prirodzené číslo: pre všetky po dvoch rôzne $\alpha_1, \dots, \alpha_q \in \mathbb{F}$ teda

$$\det(V(\alpha_1, \dots, \alpha_q)) = \prod_{i=1}^q \prod_{j=i+1}^q (\alpha_j - \alpha_i).$$

Dokážeme, že tvrdenie platí aj pre $n = q + 1$. Nech sú $\alpha_1, \dots, \alpha_{q+1} \in \mathbb{F}$ dané a po dvoch rôzne; uvažujme polynóm $D(x) \in \mathbb{F}[x]$ daný ako

$$D(x) = \det(V(\alpha_1, \dots, \alpha_q, x)) = \det \begin{pmatrix} \alpha_1^0 & \alpha_2^0 & \dots & \alpha_q^0 & x^0 \\ \alpha_1^1 & \alpha_2^1 & \dots & \alpha_q^1 & x^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_1^{q-1} & \alpha_2^{q-1} & \dots & \alpha_q^{q-1} & x^{q-1} \\ \alpha_1^q & \alpha_2^q & \dots & \alpha_q^q & x^q \end{pmatrix}.$$

Stupeň polynómu $D(x)$ je evidentne najviac q . Z Laplaceovho rozvoja podľa posledného stĺpca navyše vyplýva, že koeficient pri x^q je v polynóme $D(x)$ rovný $\det(V(\alpha_1, \dots, \alpha_q))$, pričom podľa indukčného predpokladu je toto číslo nutne nenulové; polynóm $D(x)$ je teda stupňa práve q . Je tiež zrejmé, že pre $k = 1, \dots, q$ je hodnota $D(\alpha_k)$ rovná determinantu matice s dvoma rovnakými stĺpcami, ktorá je tým pádom singulárna. Pre $k = 1, \dots, q$ je teda $D(\alpha_k) = 0$, v dôsledku čoho sú $\alpha_1, \dots, \alpha_q$ práve všetky korene polynómu $D(x)$. Preto

$$D(x) = \det(V(\alpha_1, \dots, \alpha_q)) \prod_{k=1}^q (x - \alpha_k)$$

a z indukčného predpokladu dostávame

$$\det(V(\alpha_1, \dots, \alpha_{q+1})) = D(\alpha_{q+1}) = \prod_{i=1}^q \prod_{j=i+1}^q (\alpha_j - \alpha_i) \prod_{k=1}^q (\alpha_{q+1} - \alpha_k) = \prod_{i=1}^{q+1} \prod_{j=i+1}^{q+1} (\alpha_j - \alpha_i),$$

čo bolo treba dokázať. □

Príklad 5.1.3. Pre $n \in \mathbb{N}$ uvažujme $n+1$ po dvoch rôznych uzlov $z_0, \dots, z_n \in \mathbb{C}$ a k nim prislúchajúce – nie nutne po dvoch rôzne – hodnoty $w_0, \dots, w_n \in \mathbb{C}$. Dôsledkom nesingulárnosti Vandermonovej matice je existencia *práve jedného* polynómu $p(z) \in \mathbb{C}[z]$ stupňa najviac n takého, že pre $k = 0, \dots, n$ je $p(z_k) = w_k$. Pre každú $(n+1)$ -ticu bodov $\mathbb{C}^2$ teda existuje práve jeden *interpolačný polynóm*.

Každý polynóm $p(z) \in \mathbb{C}[z]$ stupňa n totiž možno zapísať ako $p(z) = a_0 z^0 + a_1 z^1 + \dots + a_n z^n$ pre nejaké koeficienty $a_0, \dots, a_n \in \mathbb{C}$. Tento polynóm vyhovuje rovnostiam $p(z_k) = w_k$ pre $k = 0, \dots, n$ práve vtedy, keď

$$(a_0, a_1, \dots, a_n) \begin{pmatrix} z_0^0 & z_1^0 & \dots & z_n^0 \\ z_0^1 & z_1^1 & \dots & z_n^1 \\ \vdots & \vdots & \ddots & \vdots \\ z_0^n & z_1^n & \dots & z_n^n \end{pmatrix} = (w_0, w_1, \dots, w_n),$$

čiže keď

$$(a_0, a_1, \dots, a_n)V(z_0, \dots, z_n) = (w_0, w_1, \dots, w_n).$$

Vektor koeficientov $(a_0, a_1, \dots, a_n)$ teda musí byť riešením nehomogénnej sústavy lineárnych rovníc s maticou $V(z_0, \dots, z_n)$. Keďže sú však uzly interpolácie $z_0, \dots, z_n \in \mathbb{C}$ po dvoch rôzne, je podľa vety 5.1.2 Vandermondova matica $V(z_0, \dots, z_n)$ nesingulárna a uvedená sústava tak má práve jedno riešenie. Koeficienty polynómu $p(z)$ sú teda dané jednoznačne a tým pádom musí byť daný jednoznačne aj samotný polynóm $p(z)$.

Rovnakú argumentáciu možno, samozrejme, použiť aj na dôkaz existencie práve jedného *reálneho* interpolačného polynómu pre dané po dvoch rôzne uzly interpolácie $x_0, \dots, x_n \in \mathbb{R}$ a k nim prislúchajúce hodnoty $y_0, \dots, y_n \in \mathbb{R}$.

Prejdime postupne od Vandermondových matíc k *zovšeobecneným Vandermondovým maticiam*. Hoci samotná definícia týchto matíc je vonkoncom elementárna, zide sa pri uvažovaní o nich aj pohľad, pri ktorom sú jednotlivé stĺpce matice vyjadrené pomocou derivácií istých polynómov. Kým nad poľom komplexných čísel $\mathbb{C}$ a ďalšími poľami charakteristiky nula by na tieto účely postačovala formálna obdoba bežnej derivácie, nad všeobecným poľom by tento prístup znamenal nepríjemnosti spojené s možným delením nulou. Zavedieme preto jemne modifikovanú verziu formálnych derivácií polynómov – tzv. *Hasseho deriváciu*.

Definícia 5.1.4. Nech $\mathbb{F}$ je pole, $n \in \mathbb{N}$, $a_0, \dots, a_n \in \mathbb{F}$ a $p(x) \in \mathbb{F}[x]$ je polynóm

$$p(x) = \sum_{t=0}^n a_t x^t.$$

Pre ľubovoľné $k \in \mathbb{N}$ potom k -tou *Hasseho deriváciou* – prípadne *Hasseho deriváciou rádu k* – polynómu $p(x)$ nazveme polynóm

$$D^{(k)}p(x) = \sum_{t=k}^n a_t \binom{t}{k} x^{t-k}.$$

Poznámka 5.1.5. Bežná k -ta formálna derivácia polynómu

$$p(x) = \sum_{t=0}^n a_t x^t \in \mathbb{F}[x]$$

je daná ako

$$\frac{d^k}{dx^k} p(x) = \sum_{t=k}^n a_t t^k x^{t-k},$$

kde $t^{\underline{k}} := t(t-1)\dots(t-k+1)$ je k -ty klesajúci faktoriál. Vzťah

$$k! \binom{t}{k} = t^{\underline{k}}$$

možno v poli komplexných čísel $\mathbb{C}$ a ďalších poliach charakteristiky nula prepísať ako

$$\binom{t}{k} = \frac{t^{\underline{k}}}{k!}.$$

Nad takýmito poľami teda môžeme k -tu Hasseho deriváciu polynómu $p(x)$ vyjadriť vzťahom

$$D^{(k)}p(x) = \frac{1}{k!} \frac{d^k}{dx^k} p(x);$$

ide teda o k -tu deriváciu normalizovanú tak, aby bolo $D^{(k)}x^k = 1$. V poliach kladnej charakteristiky naopak môže byť $t^{\underline{k}}$ aj $k!$ rovné nule, čím sa takéto vyjadrenie Hasseho derivácie pomocou bežnej derivácie stáva nemožným.

Niektoré základné vlastnosti Hasseho derivácie, ktoré budeme pri úvahách spojených so zovšeobecnenými Vandermondovými maticami využívať, sú sformulované v nasledujúcich tvrdeniach. Začnime jednoduchým pozorovaním linearitu operátora $D^{(k)}$.

Tvrdenie 5.1.6. *Nech $\mathbb{F}$ je pole, $p(x), q(x) \in \mathbb{F}[x]$ polynómy, $a, b \in \mathbb{F}$ a $k \in \mathbb{N}$. Potom*

$$D^{(k)}(ap(x) + bq(x)) = aD^{(k)}p(x) + bD^{(k)}q(x).$$

Dôkaz. Nech $n \in \mathbb{N}$, $a_0, \dots, a_n \in \mathbb{F}$ a $b_0, \dots, b_n \in \mathbb{F}$ sú také, že

$$p(x) = \sum_{t=0}^n a_t x^t \quad \text{a} \quad q(x) = \sum_{t=0}^n b_t x^t.$$

Potom skutočne

$$\begin{aligned} D^{(k)}(ap(x) + bq(x)) &= D^{(k)}\left(\sum_{t=0}^n (aa_t + bb_t)x^t\right) = \sum_{t=k}^n (aa_t + bb_t) \binom{t}{k} x^{t-k} = \\ &= a \sum_{t=k}^n a_t \binom{t}{k} x^{t-k} + b \sum_{t=k}^n b_t \binom{t}{k} x^{t-k} = aD^{(k)}p(x) + bD^{(k)}q(x). \quad \square \end{aligned}$$

Zídu sa nám aj varianty *Leibnizovho* a *zovšeobecného Leibnizovho vzorca* pre Hasseho derivácie, ktoré postupne sformulujeme v nasledujúcich dvoch tvrdeniach.

Tvrdenie 5.1.7. *Nech $\mathbb{F}$ je pole, $p_1(x), p_2(x) \in \mathbb{F}[x]$ polynómy a $k \in \mathbb{N}$. Potom*

$$D^{(k)}(p_1(x)p_2(x)) = \sum_{j=0}^k \left(D^{(j)}p_1(x)\right) \left(D^{(k-j)}p_2(x)\right).$$

Dôkaz. Uvažujme $n \in \mathbb{N}$, $a_0, \dots, a_n \in \mathbb{F}$ a $b_0, \dots, b_n \in \mathbb{F}$ také, že

$$p_1(x) = \sum_{t=0}^n a_t x^t \quad \text{a} \quad p_2(x) = \sum_{t=0}^n b_t x^t.$$

Pre všetky prirodzené čísla $t > n$ ďalej položíme $a_t = b_t = 0$. V takom prípade je

$$p_1(x)p_2(x) = \sum_{t=0}^{2n} \left(\sum_{\ell=0}^t a_\ell b_{t-\ell}\right) x^t,$$

z čoho

$$D^{(k)}(p_1(x)p_2(x)) = \sum_{t=k}^{2n} \left(\sum_{\ell=0}^t a_\ell b_{t-\ell} \right) \binom{t}{k} x^{t-k} = \sum_{t=k}^{2n} \left(\sum_{\ell=0}^t \binom{t}{k} a_\ell b_{t-\ell} \right) x^{t-k}. \quad (5.1)$$

Zisťujeme teda, že

$$\begin{aligned} \sum_{j=0}^k \left(D^{(j)} p_1(x) \right) \left(D^{(k-j)} p_2(x) \right) &= \sum_{j=0}^k \left(\sum_{u=j}^n a_u \binom{u}{j} x^{u-j} \right) \left(\sum_{v=k-j}^n b_v \binom{v}{k-j} x^{v-k+j} \right) = \\ &= \sum_{j=0}^k \sum_{t=k}^{2n} \left(\sum_{\ell=0}^t a_\ell \binom{\ell}{j} b_{t-\ell} \binom{t-\ell}{k-j} \right) x^{t-k} = \\ &= \sum_{t=k}^{2n} \left(\sum_{\ell=0}^t \left(\sum_{j=0}^k \binom{\ell}{j} \binom{t-\ell}{k-j} \right) a_\ell b_{t-\ell} \right) x^{t-k} = \\ &= \sum_{t=k}^{2n} \left(\sum_{\ell=0}^t \binom{t}{k} a_\ell b_{t-\ell} \right) x^{t-k} = D^{(k)}(p_1(x)p_2(x)) \end{aligned}$$

podľa (5.1). V poslednom kroku sme pritom použili elementárny variant *Vandermondovej identity*:

$$\binom{m+n}{k} = \sum_{j=0}^k \binom{m}{j} \binom{n}{k-j}$$

pre všetky $m, n, k \in \mathbb{N}$.²

□

Tvrdenie 5.1.8. *Nech $\mathbb{F}$ je pole, $m \in \mathbb{N}$, $p_1(x), \dots, p_m(x) \in \mathbb{F}[x]$ polynómy a $k \in \mathbb{N}$. Potom*

$$D^{(k)}(p_1(x) \dots p_m(x)) = \sum_{\substack{k_1, \dots, k_m \in \mathbb{N} \\ k_1 + \dots + k_m = k}} \prod_{j=1}^m D^{(k_j)} p_j(x).$$

Dôkaz. Indukciou vzhľadom na m . Pre $m = 0$ a $m = 1$ je tvrdenie zrejmé. Predpokladajme teraz platnosť tvrdenia pre $m = s \in \mathbb{N} \setminus \{0\}$ a uvažujme $m = s + 1$. Pre ľubovoľné polynómy $p_1(x), \dots, p_{s+1}(x) \in \mathbb{F}[x]$ potom s použitím Leibnizovho vzorca z tvrdenia 5.1.7 dostávame

$$\begin{aligned} D^{(k)}(p_1(x) \dots p_{s+1}(x)) &= D^{(k)}(p_{s+1}(x) (p_1(x) \dots p_s(x))) = \\ &= \sum_{k_{s+1}=0}^k \left(D^{(k_{s+1})} p_{s+1}(x) \right) \left(D^{(k-k_{s+1})} (p_1(x) \dots p_s(x)) \right). \end{aligned}$$

S využitím indukčného predpokladu teda vidíme, že

$$\begin{aligned} D^{(k)}(p_1(x) \dots p_{s+1}(x)) &= \sum_{k_{s+1}=0}^k \left(D^{(k_{s+1})} p_{s+1}(x) \right) \left(\sum_{\substack{k_1, \dots, k_s \in \mathbb{N} \\ k_1 + \dots + k_s = k - k_{s+1}}} \prod_{j=1}^s D^{(k_j)} p_j(x) \right) = \\ &= \sum_{\substack{k_1, \dots, k_{s+1} \in \mathbb{N} \\ k_1 + \dots + k_{s+1} = k}} \prod_{j=1}^{s+1} D^{(k_j)} p_j(x), \end{aligned}$$

čo bolo treba dokázať.

□

²Výber k prvkov z množiny $[m+n]$ môžeme realizovať tak, že určíme počet prvkov j , ktoré budeme vyberať z $[m]$ a následne vyberieme j prvkov z $[m]$ a $k-j$ prvkov z $\{m+1, \dots, m+n\}$.

Dokážme napokon ešte jedno pozorovanie o Hasseho deriváciách, podľa ktorého možno – podobne ako pri bežných deriváciách – polynómy $(x+a)^t$ derivovať rovnako ako v prípade polynómov x^t .

Tvrdenie 5.1.9. *Nech $\mathbb{F}$ je pole, $a \in \mathbb{F}$ a $k, t \in \mathbb{N}$ sú také, že $t \geq k$. Potom*

$$D^{(k)}(x+a)^t = \binom{t}{k}(x+a)^{t-k}.$$

Dôkaz. Uvažujme najprv prípad, keď $t = k$. Potom

$$D^{(k)}(x+a)^t = D^{(k)}(x+a)^k = D^{(k)} \sum_{j=0}^k \binom{k}{j} x^j a^{k-j} = \binom{k}{k} x^{k-k} = \binom{t}{k}(x+a)^{t-k}.$$

Predpokladajme teraz platnosť tvrdenia pre $t = s \geq k$ a uvažujme $t = s + 1$. S použitím Leibnizovho vzorca z tvrdenia 5.1.7 a indukčného predpokladu potom zisťujeme, že

$$\begin{aligned} D^{(k)}(x+a)^{s+1} &= D^{(k)}((x+a)^s(x+a)) = \sum_{j=0}^k \left(D^{(j)}(x+a)^s \right) \left(D^{(k-j)}(x+a) \right) = \\ &= D^{(k-1)}(x+a)^s + (x+a)D^{(k)}(x+a)^s = \\ &= \binom{s}{k-1}(x+a)^{s+1-k} + (x+a)\binom{s}{k}(x+a)^{s-k} = \\ &= \left(\binom{s}{k-1} + \binom{s}{k} \right) (x+a)^{s+1-k} = \binom{s+1}{k}(x+a)^{s+1-k}, \end{aligned}$$

čím je tvrdenie dokázané. □

Vráťme sa teraz opäť k maticiam. Vandermondovu maticu $V(\alpha_1, \dots, \alpha_n)$ nad $\mathbb{F}$ možno chápať aj tak, že jej k -ty stĺpec je rovný $\mathbf{f}(\alpha_k)$, kde $\mathbf{f}: \mathbb{F} \rightarrow \mathbb{F}^n$ je zobrazenie dané pre všetky $x \in \mathbb{F}$ ako

$$\mathbf{f}(x) = (x^0, x^1, x^2, \dots, x^{n-1})^T.$$

Pre $m = 0, \dots, n-1$ teraz uvažujme analogicky definované zobrazenie $\mathbf{f}^{(m)}: \mathbb{F} \rightarrow \mathbb{F}^n$, ktorého výstupmi sú vektory hodnôt polynomických funkcií zodpovedajúcich m -tým Hasseho deriváciám polynómov $x^0, \dots, x^{n-1}$. Keďže pre $r \geq m$ je

$$D^{(m)}x^r = \binom{r}{m}x^{r-m},$$

pre $m = 0, \dots, n-1$ a všetky $x \in \mathbb{F}$ dostávame

$$\mathbf{f}^{(m)}(x) = \left(\underbrace{0, 0, \dots, 0}_m, \binom{m}{m}x^0, \binom{m+1}{m}x^1, \dots, \binom{n-1}{m}x^{n-1-m} \right)^T.$$

Zovšeobecnená Vandermondova matica rádu n bude pre $\alpha_1, \dots, \alpha_s \in \mathbb{F}$ a násobnosti $n_1, \dots, n_s \in \mathbb{N} \setminus \{0\}$ spĺňajúce $n_1 + \dots + n_s = n$ okrem stĺpcov $\mathbf{f}(\alpha_k)$ pre $k = 1, \dots, s$ obsahovať aj príslušné stĺpce zodpovedajúce prvým $n_k - 1$ Hasseho deriváciám – ku každému α_k tak budú prislúchať stĺpce

$$\mathbf{f}(\alpha_k) = \mathbf{f}^{(0)}(\alpha_k), \quad \mathbf{f}^{(1)}(\alpha_k), \quad \dots, \quad \mathbf{f}^{(n_k-1)}(\alpha_k).$$

Definícia 5.1.10. Nech $\mathbb{F}$ je pole. Zovšeobecnenou Vandermondovou maticou rádu $n \in \mathbb{N}$ nad $\mathbb{F}$ nazveme pre ľubovoľné $s \in \mathbb{N}$, $\alpha_1, \dots, \alpha_s \in \mathbb{F}$ a $n_1, \dots, n_s \in \mathbb{N} \setminus \{0\}$ spĺňajúce $n_1 + \dots + n_s = n$ štvorcovú maticu $G = G(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)$ typu $n \times n$ danú ako³

$$G = \left(\begin{array}{cccc|cccc} \alpha_1^0 & 0 & \dots & 0 & \dots & \alpha_s^0 & 0 & \dots & 0 \\ \alpha_1^1 & \binom{1}{1}\alpha_1^0 & \dots & 0 & \dots & \alpha_s^1 & \binom{1}{1}\alpha_s^0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n_1-1} & \binom{n_1-1}{1}\alpha_1^{n_1-2} & \dots & \binom{n_1-1}{n_1-1}\alpha_1^0 & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_1^{n-1} & \binom{n-1}{1}\alpha_1^{n-2} & \dots & \binom{n-1}{n_1-1}\alpha_1^{n-n_1} & \dots & \alpha_s^{n-1} & \binom{n-1}{1}\alpha_s^{n-2} & \dots & \binom{n-1}{n_s-1}\alpha_s^{n-n_s} \end{array} \right).$$

Veta 5.1.11. Nech $\mathbb{F}$ je pole, $n, s \in \mathbb{N}$, $\alpha_1, \dots, \alpha_s \in \mathbb{F}$ a $n_1, \dots, n_s \in \mathbb{N} \setminus \{0\}$ sú také, že $n_1 + \dots + n_s = n$. Potom

$$\det(G(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)) = \prod_{i=1}^s \prod_{j=i+1}^s (\alpha_j - \alpha_i)^{n_i n_j}.$$

Ak sú teda prvky $\alpha_1, \dots, \alpha_s$ po dvoch rôzne, je matica $G(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)$ nesingulárna.

Dôkaz. Pre $n = 0$ je dokazované tvrdenie triviálne. Predpokladajme teda, že je n nenulové a dané pevne. Znenie vety dokážeme pre všetky prípustné $\mathbb{F}$, s , $\alpha_1, \dots, \alpha_s$ a $n_1, \dots, n_s$ indukciou vzhľadom na počet stĺpcov zodpovedajúcich Hasseho deriváciám nenulového rádu – čiže vzhľadom na

$$(n_1 - 1) + \dots + (n_s - 1) = \sum_{k=1}^s (n_k - 1) = n - s.$$

Zrejme musí byť $s \in \mathbb{N} \setminus \{0\}$ a $0 \leq (n_1 - 1) + \dots + (n_s - 1) \leq n - 1$. Ak $(n_1 - 1) + \dots + (n_s - 1) = 0$, nutne $n_1 = \dots = n_s = 1$ a matica $G(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)$ je Vandermondovou maticou:

$$G(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s) = V(\alpha_1, \dots, \alpha_s).$$

Báza indukcie tak vyplýva priamo z vety 5.1.2.

Predpokladajme teraz, že je veta dokázaná pre $(n_1 - 1) + \dots + (n_s - 1) = q$, kde $q < n - 1$ je prirodzené číslo. Dokážeme jej platnosť pre $(n_1 - 1) + \dots + (n_s - 1) = q + 1$.

Nech $\mathbb{F}$ je pole, $s \in \mathbb{N} \setminus \{0\}$, $\alpha_1, \dots, \alpha_s \in \mathbb{F}$ a $n_1, \dots, n_s \in \mathbb{N} \setminus \{0\}$ sú také, že $n_1 + \dots + n_s = n$ a

$$(n_1 - 1) + \dots + (n_s - 1) = q + 1.$$

Maticu $G = G(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)$ potom možno v blokovom tvare zapísať ako

$$G = (B(\alpha_1, n_1) \mid B(\alpha_2, n_2) \mid \dots \mid B(\alpha_s, n_s)),$$

kde blok $B(\alpha_k, n_k)$ je pre $k = 1, \dots, s$ daný ako

$$B(\alpha_k, n_k) = \left(\begin{array}{cccc|cccc} \alpha_k^0 & 0 & \dots & 0 & 0 \\ \alpha_k^1 & \binom{1}{1}\alpha_k^0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_k^{n_k-2} & \binom{n_k-2}{1}\alpha_k^{n_k-3} & \dots & \binom{n_k-2}{n_k-2}\alpha_k^0 & 0 \\ \alpha_k^{n_k-1} & \binom{n_k-1}{1}\alpha_k^{n_k-2} & \dots & \binom{n_k-1}{n_k-2}\alpha_k^1 & \binom{n_k-1}{n_k-1}\alpha_k^0 \\ \vdots & \vdots & & \vdots & \vdots \\ \alpha_k^{n-1} & \binom{n-1}{1}\alpha_k^{n-2} & \dots & \binom{n-1}{n_k-2}\alpha_k^{n-(n_k-1)} & \binom{n-1}{n_k-1}\alpha_k^{n-n_k} \end{array} \right).$$

³Zvislé čiary slúžia v nasledujúcom zápise matice $G = G(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)$ len na vizuálne oddelenie blokov prislúchajúcich k jednotlivým prvkom $\alpha_1, \dots, \alpha_s$.

Nech teraz $k \in [s]$ je ľubovoľný index, pre ktorý je $n_k > 1$; aspoň jeden taký index určite musí existovať, pretože $(n_1 - 1) + \dots + (n_s - 1) = q + 1 > 0$. Položme

$$\overline{B}(\alpha_k, n_k, x) := \begin{pmatrix} \alpha_k^0 & 0 & \dots & 0 & x^0 \\ \alpha_k^1 & \binom{1}{1} \alpha_k^0 & \dots & 0 & x^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_k^{n_k-2} & \binom{n_k-2}{1} \alpha_k^{n_k-3} & \dots & \binom{n_k-2}{n_k-2} \alpha_k^0 & x^{n_k-2} \\ \alpha_k^{n_k-1} & \binom{n_k-1}{1} \alpha_k^{n_k-2} & \dots & \binom{n_k-1}{n_k-2} \alpha_k^1 & x^{n_k-1} \\ \vdots & \vdots & & \vdots & \vdots \\ \alpha_k^{n-1} & \binom{n-1}{1} \alpha_k^{n-2} & \dots & \binom{n-1}{n_k-2} \alpha_k^{n-(n_k-1)} & x^{n-1} \end{pmatrix}$$

a uvažujme maticu

$$\overline{G} = (B(\alpha_1, n_1) \mid \dots \mid B(\alpha_{k-1}, n_{k-1}) \mid \overline{B}(\alpha_k, n_k, x) \mid B(\alpha_{k+1}, n_{k+1}) \mid \dots \mid B(\alpha_s, n_s)).$$

To možno pre $B(\alpha_k, n_k - 1)$ pozostávajúce z prvých $n_k - 1$ stĺpcov bloku $B(\alpha_k, n_k)$ a pre jednotlípový blok $B(x, 1) = (x^0, x^1, \dots, x^{n-1})^T$ očividne prepísať ako

$$\overline{G} = (B(\alpha_1, n_1) \mid \dots \mid B(\alpha_{k-1}, n_{k-1}) \mid B(\alpha_k, n_k - 1) \mid B(x, 1) \mid B(\alpha_{k+1}, n_{k+1}) \mid \dots \mid B(\alpha_s, n_s)),$$

pričom $(n_1 - 1) + \dots + (n_{k-1} - 1) + (n_k - 1 - 1) + (1 - 1) + (n_{k+1} - 1) + \dots + (n_s - 1) = q$. Na maticu $\overline{G}$ nad podielovým poľom $\mathbb{F}(x)$ okruhu polynómov $\mathbb{F}[x]$ sa teda vzťahuje indukčný predpoklad, z ktorého dostávame

$$\det(\overline{G}) = \prod_{i=1}^s \prod_{j=i+1}^s (\alpha_j - \alpha_i)^{m_i m_j} \prod_{i=1}^k (x - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - x)^{m_j},$$

kde m_i je pre $i = 1, \dots, s$ dané ako

$$m_i = \begin{cases} n_i - 1 & \text{ak } i = k, \\ n_i & \text{inak.} \end{cases} \quad (5.2)$$

Ľahko potom vidieť, že

$$\begin{aligned} \det(G) &= \left[D^{(n_k-1)} \det(\overline{G}) \right]_{x=\alpha_k} = \\ &= \left[D^{(n_k-1)} \left(\prod_{i=1}^s \prod_{j=i+1}^s (\alpha_j - \alpha_i)^{m_i m_j} \prod_{i=1}^k (x - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - x)^{m_j} \right) \right]_{x=\alpha_k} = \\ &= \left(\prod_{i=1}^s \prod_{j=i+1}^s (\alpha_j - \alpha_i)^{m_i m_j} \right) \left[D^{(n_k-1)} \left(\prod_{i=1}^k (x - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - x)^{m_j} \right) \right]_{x=\alpha_k}. \end{aligned}$$

Zo vzťahov (5.2) teda vyplýva, že dôkaz bude hotový, keď ukážeme, že

$$\begin{aligned} &\left[D^{(n_k-1)} \left(\prod_{i=1}^k (x - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - x)^{m_j} \right) \right]_{x=\alpha_k} = \\ &= \prod_{i=1}^{k-1} (\alpha_k - \alpha_i)^{n_i} \prod_{j=k+1}^s (\alpha_j - \alpha_k)^{n_j}. \end{aligned} \quad (5.3)$$

Ak ale na výpočet hodnoty

$$\left[D^{(n_k-1)} \left(\prod_{i=1}^k (x - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - x)^{m_j} \right) \right]_{x=\alpha_k}$$

aplikujeme zovšeobecnený Leibnizov vzorec z tvrdenia 5.1.8, vďaka vzťahu $m_k = n_k - 1$ a tvrdeniu 5.1.9 zisťujeme, že po dosadení $x = \alpha_k$ môže byť nenulový iba ten člen výslednej sumy, v ktorom aplikujeme Hasseho deriváciu rádu $(n_k - 1)$ na $(x - \alpha_k)^{m_k} = (x - \alpha_k)^{n_k-1}$. V dôsledku toho

$$\begin{aligned} & \left[D^{(n_k-1)} \left(\prod_{i=1}^k (x - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - x)^{m_j} \right) \right]_{x=\alpha_k} = \\ &= \left[\prod_{i=1}^{k-1} (x - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - x)^{m_j} \right]_{x=\alpha_k} = \\ &= \prod_{i=1}^{k-1} (\alpha_k - \alpha_i)^{m_i} \prod_{j=k+1}^s (\alpha_j - \alpha_k)^{m_j} = \\ &= \prod_{i=1}^{k-1} (\alpha_k - \alpha_i)^{n_i} \prod_{j=k+1}^s (\alpha_j - \alpha_k)^{n_j}, \end{aligned}$$

čím je dôkaz rovnosti (5.3) – a teda aj celej vety – dokončený. $\square$

Na záver tohto oddielu ešte uvedme jednoduchý dôsledok vety 5.1.11 hovoriaci o nesusingulárnosti určitých matíc *komplexných čísel* veľmi blízkych zovšeobecneným Vandermondovým. Rovnaké pozorovanie by sme mohli urobiť aj nad ostatnými poľami *charakteristiky nula*; nad poľami kladnej charakteristiky už ale matice z nasledujúceho dôsledku vo všeobecnosti nemusia byť nesusingulárne.

Dôsledok 5.1.12. *Nech $n, s \in \mathbb{N}$, $\alpha_1, \dots, \alpha_s \in \mathbb{C}$ a $n_1, \dots, n_s \in \mathbb{N} \setminus \{0\}$ sú také, že $n_1 + \dots + n_s = n$. Nech $H = H(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)$ je matica*

$$H = \left(\begin{array}{cccc|cccc} \uparrow & \uparrow & & \uparrow & & \uparrow & \uparrow & \uparrow \\ \mathbf{h}(\alpha_1, 0) & \mathbf{h}(\alpha_1, 1) & \dots & \mathbf{h}(\alpha_1, n_1 - 1) & \dots & \mathbf{h}(\alpha_s, 0) & \mathbf{h}(\alpha_s, 1) & \dots & \mathbf{h}(\alpha_s, n_s - 1) \\ \downarrow & \downarrow & & \downarrow & & \downarrow & \downarrow & & \downarrow \end{array} \right),$$

kde pre $k = 1, \dots, s$ a $r = 0, \dots, n_k - 1$ je stĺpec $\mathbf{h}(\alpha_k, r)$ daný v prípade $\alpha_k \neq 0$ ako

$$\mathbf{h}(\alpha_k, r) = (0^r \alpha_k^0, 1^r \alpha_k^1, \dots, (n-1)^r \alpha_k^{n-1})^T$$

a v prípade $\alpha_k = 0$ ako

$$\mathbf{h}(\alpha_k, r) = \left(\underbrace{0, \dots, 0}_r, 1, \underbrace{0, \dots, 0}_{n-r-1} \right)^T.$$

Ak sú čísla $\alpha_1, \dots, \alpha_s$ po dvoch rôzne, je matica $H(\alpha_1, \dots, \alpha_s; n_1, \dots, n_s)$ nesusingulárna.

Dôkaz. Nech sú $\alpha_1, \dots, \alpha_s$ po dvoch rôzne. Podľa vety 5.1.11 je v takom prípade nesusingulárna matica $G = G(\alpha_1, \dots, \alpha_s, n_1, \dots, n_s)$, ktorá je daná ako

$$G = \left(\begin{array}{cccc|cccc} \alpha_1^0 & 0 & \dots & 0 & \dots & \alpha_s^0 & 0 & \dots & 0 \\ \alpha_1^1 & \binom{1}{1} \alpha_1^0 & \dots & 0 & \dots & \alpha_s^1 & \binom{1}{1} \alpha_s^0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n_1-1} & \binom{n_1-1}{1} \alpha_1^{n_1-2} & \dots & \binom{n_1-1}{n_1-1} \alpha_1^0 & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_1^{n-1} & \binom{n-1}{1} \alpha_1^{n-2} & \dots & \binom{n-1}{n_1-1} \alpha_1^{n-n_1} & \dots & \alpha_s^{n-1} & \binom{n-1}{1} \alpha_s^{n-2} & \dots & \binom{n-1}{n_s-1} \alpha_s^{n-n_s} \end{array} \right).$$

Položme teraz pre jej stĺpce

$$G = \left(\begin{array}{ccc|ccc} \uparrow & \uparrow & & \uparrow & & \\ \mathbf{g}(\alpha_1, 0) & \mathbf{g}(\alpha_1, 1) & \dots & \mathbf{g}(\alpha_1, n_1 - 1) & \dots & \mathbf{g}(\alpha_s, 0) \\ \downarrow & \downarrow & & \downarrow & & \downarrow \end{array} \right),$$

Pre $k = 1, \dots, s$ a $r = 0, \dots, n_k - 1$ teda

$$\mathbf{g}(\alpha_k, r) = \frac{1}{\alpha_k^r} \left(\binom{0}{r} \alpha_k^0, \binom{1}{r} \alpha_k^1, \dots, \binom{n-1}{r} \alpha_k^{n-1} \right)^T,$$

ak $\alpha_k \neq 0$ a pre $\alpha_k = 0$ je

$$\mathbf{g}(\alpha_k, r) = \left(\underbrace{0, \dots, 0}_r, 1, \underbrace{0, \dots, 0}_{n-r-1} \right)^T.$$

Opíšeme teraz postup, ktorým možno pre každé $k \in [s]$ získať zo stĺpcov $\mathbf{g}(\alpha_k, r)$ pre $r = 0, \dots, n_k - 1$ pomocou elementárnych stĺpcových operácií stĺpce $\mathbf{h}(\alpha_k, r)$ pre $r = 0, \dots, n_k - 1$. Po vykonaní tejto transformácie pre $k = 1, \dots, s$ tak získame z nesusingularnej matice G iba s použitím elementárnych stĺpcových operácií maticu H , ktorá v dôsledku toho bude musieť byť tiež nesusingularná.

Nech je teda $k \in [s]$ pevné. Ak $\alpha_k = 0$, je pre $r = 0, \dots, n_k - 1$ priamo $\mathbf{g}(\alpha_k, r) = \mathbf{h}(\alpha_k, r)$. Uvažujme teda prípad, keď $\alpha_k \neq 0$. Evidentne je $\mathbf{g}(\alpha_k, 0) = \mathbf{h}(\alpha_k, 0)$. Predpokladajme, že sme už pre nejaké $q \in \{0, \dots, n_k - 2\}$ stĺpce $\mathbf{g}(\alpha_k, 0), \dots, \mathbf{g}(\alpha_k, q)$ transformovali s použitím elementárnych stĺpcových operácií na $\mathbf{h}(\alpha_k, 0), \dots, \mathbf{h}(\alpha_k, q)$, pričom zvyšné stĺpce matice ostali nezmenené. Ukážeme, ako možno transformovať stĺpec $\mathbf{g}(\alpha_k, q+1)$ na $\mathbf{h}(\alpha_k, q+1)$, a to iba s pomocou elementárnych stĺpcových operácií a bez toho, aby sme pozmenili zvyšné stĺpce matice.

Pre $j = 0, \dots, n-1$ je v riadku $j+1$ stĺpcového vektora $\mathbf{g}(\alpha_k, q+1)$ číslo

$$\begin{aligned} \frac{1}{\alpha_k^{q+1}} \binom{j}{q+1} \alpha_k^j &= \frac{j(j-1) \dots (j-q)}{\alpha_k^{q+1} (q+1)!} \alpha_k^j = \\ &= \frac{1}{\alpha_k^{q+1} (q+1)!} \left(j^{q+1} \alpha_k^j + a_q j^q \alpha_k^j + a_{q-1} j^{q-1} \alpha_k^j + \dots + a_0 j^0 \alpha_k^j \right), \end{aligned}$$

kde $a_0, \dots, a_q \in \mathbb{Z}$ sú konštanty nezávislé od j . Je teda zrejmé, že

$$\mathbf{h}(\alpha_k, q+1) = \left(\alpha_k^{q+1} (q+1)! \right) \mathbf{g}(\alpha_k, q+1) - a_q \mathbf{h}(\alpha_k, q) - a_{q-1} \mathbf{h}(\alpha_k, q-1) - \dots - a_0 \mathbf{h}(\alpha_k, 0).$$

Stĺpec $\mathbf{h}(\alpha_k, q+1)$ tak získame prenásobením stĺpca $\mathbf{g}(\alpha_k, q+1)$ konštantou a následným odpočítaním niekoľkých skalárnych násobkov iných stĺpcov matice – čiže postupnosťou elementárnych stĺpcových operácií, čo sme chceli dokázať. $\square$

5.2 Počet sledov dĺžky t v multigrafe

Pripomeňme si, že pod *orientovaným multigrafom* rozumieme – neformálne povedané – orientovaný graf, ktorý môže obsahovať násobné hrany medzi vrcholmi. Podobne ako v príklade 1.1.9 budeme multigrafy reprezentovať ohodnotenými grafmi nad polokruhom prirodzených čísel. Ohodnotenie hrany vedúcej medzi nejakou dvojicou vrcholov tak možno chápať ako jej multiplicitu – čiže počet hrán spájajúcich tú istú dvojicu vrcholov v pôvodnom multigrafe. V takýchto prípadoch budeme hovoriť, že ohodnotený graf $\mathcal{G}$ nad $\mathbb{N}$ je *reprezentáciou* príslušného multigrafu.

Veta 5.2.1. *Nech n je prirodzené číslo, $\mathcal{G} = ([n], \varphi)$ je ohodnotený graf nad $\mathbb{N}$ a $i, j \in V(\mathcal{G})$ sú jeho vrcholy. Nech $A := \mathbf{A}(\mathcal{G})$ je matica susednosti grafu $\mathcal{G}$. Pre všetky $t \in \mathbb{N}$ je potom počet sledov $S_{i,j}(t)$ dĺžky t z vrcholu i do vrcholu j v multigrafe reprezentovanom grafom $\mathcal{G}$ daný ako*

$$S_{i,j}(t) = \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} c_{\lambda,k} t^k \lambda^t + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} c_{\lambda,k} \delta_{t,k},$$

kde pre všetky $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$ je $c_{\lambda,k} \in \mathbb{C}$ konštanta nezávislá od t a Kroneckerova delta $\delta_{r,s}$ je pre všetky $r, s \in \mathbb{N}$ daná ako

$$\delta_{r,s} = \begin{cases} 1 & \text{ak } r = s, \\ 0 & \text{ak } r \neq s. \end{cases}$$

Dôkaz. Z dôsledku 4.6.11 vyplýva, že existuje matica $J \in \mathbb{C}^{n \times n}$ v Jordanovom kanonickom tvare taká, že pre nejakú nesingulárnu maticu P je

$$A = PJP^{-1}.$$

Pre všetky $t \in \mathbb{N}$ preto

$$A^t = PJ^tP^{-1}. \quad (5.4)$$

Matica J je blokovo diagonálna a pozostáva z niekoľkých Jordanových buniek, kde každá z nich je rovná $J_m(\lambda)$ pre nejaké $\lambda \in \sigma(A)$ a $m \in \mathbb{N}$ také, že $1 \leq m \leq \alpha(A, \lambda)$. Každý prvok matice J^t je teda buď nulový, alebo je prvkom t -tej mocniny niektorej z týchto Jordanových buniek. Na základe vety 4.7.1 potom zisťujeme, že pre ľubovoľné $i, j \in V(\mathcal{G})$ je prvok $q_{i,j}[t]$ matice $J^t = (q_{i,j}[t])_{n \times n}$ alebo nulový pre všetky $t \in \mathbb{N}$, alebo existuje $\lambda \in \sigma(A)$ také, že pre nejaké $k \in \mathbb{N}$ spĺňajúce $0 \leq k \leq \alpha(A, \lambda) - 1$ a všetky $t \in \mathbb{N}$ je⁴

$$q_{i,j}[t] = \binom{t}{k} \lambda^{t-k}. \quad (5.5)$$

Keďže binomický koeficient $\binom{t}{k}$ možno interpretovať aj ako polynomickeú funkciu stupňa k o premennej t a keďže pre $\lambda = 0$ možno rovnosť (5.5) prepísať ako $q_{i,j}[t] = \delta_{t,k}$, na završenie dôkazu sa už len stačí odvolať na rovnosť (5.4) a na vetu 1.3.10 poskytujúcu grafovú interpretáciu matice A^t . $\square$

Z tvrdenia 4.2.5 vieme, že pre ľubovoľné nenulové prirodzené číslo n a štvorcovú maticu $A \in \mathbb{C}^{n \times n}$ je

$$\sum_{\lambda \in \sigma(A)} \alpha(A, \lambda) = n.$$

Vo výraze

$$\sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} c_{\lambda,k} t^k \lambda^t + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} c_{\lambda,k} \delta_{t,k},$$

ktorý podľa predchádzajúcej vety vyjadruje počet sledov dĺžky t medzi dvoma danými vrcholmi daného multigrafu, tak vystupuje práve n neznámych konštánt $c_{\lambda,k}$. Obvykle navyše nie je žiaden problém vypočítať počty sledov $S_{i,j}(t)$ dĺžky t vedúcich z vrcholu i do vrcholu j pre $t = 0, \dots, n-1$; jednotlivé

⁴Pri konvencii, že pre $t < k$ je $\binom{t}{k} 0^{t-k} = 0$.

konštanty $c_{\lambda,k}$ pre $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$ sú potom riešením sústavy

$$\begin{aligned} \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A, \lambda)-1} x_{\lambda,k} 0^k \lambda^0 + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A, \lambda)-1} x_{\lambda,k} \delta_{0,k} &= S_{i,j}(0), \\ \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A, \lambda)-1} x_{\lambda,k} 1^k \lambda^1 + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A, \lambda)-1} x_{\lambda,k} \delta_{1,k} &= S_{i,j}(1), \\ &\vdots \\ \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A, \lambda)-1} x_{\lambda,k} (n-1)^k \lambda^{n-1} + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A, \lambda)-1} x_{\lambda,k} \delta_{n-1,k} &= S_{i,j}(n-1), \end{aligned}$$

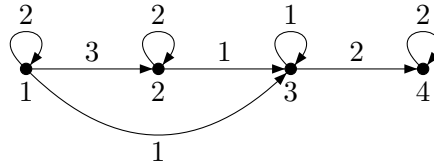
kde neznámymi sú $x_{\lambda,k}$ pre $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$. Ak $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$, kde vlastné čísla $\lambda_1, \dots, \lambda_m$ sú po dvoch rôzne, možno uvedenú sústavu rovníc v reči vektorov a matíc prepísať ako

$$H(\lambda_1, \dots, \lambda_m; \alpha(A, \lambda_1), \dots, \alpha(A, \lambda_m)) \begin{pmatrix} x_{\lambda_1,0} \\ \vdots \\ x_{\lambda_1, \alpha(A, \lambda_1)-1} \\ \vdots \\ x_{\lambda_m,0} \\ \vdots \\ x_{\lambda_m, \alpha(A, \lambda_m)-1} \end{pmatrix} = \begin{pmatrix} S_{i,j}(0) \\ S_{i,j}(1) \\ \vdots \\ S_{i,j}(n-1) \end{pmatrix},$$

kde $H(\lambda_1, \dots, \lambda_m; \alpha(A, \lambda_1), \dots, \alpha(A, \lambda_m))$ označuje maticu z dôsledku 5.1.12, podľa ktorého je táto matica nesingulárna. Uvedená sústava rovníc má teda *jediné* riešenie, ktorým sú hľadané konštanty $c_{\lambda,k}$ pre $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$.

Pozorovania učinené vyššie dohromady poskytujú relatívne jednoduchú metódu enumerácie sledov medzi dvoma danými vrcholmi multigrafu. Tú teraz ilustrujeme na príklade.

Príklad 5.2.2. Uvažujme orientovaný multigraf reprezentovaný ohodnoteným grafom $\mathcal{G}$ nad $\mathbb{N}$, ktorý je znázornený na obrázku 5.1.



Obr. 5.1: Ohodnotený graf $\mathcal{G}$ nad $\mathbb{N}$ reprezentujúci uvažovaný multigraf.

Predpokladajme, že sa zaujímate o počet sledov dĺžky t vedúcich v tomto multigrafe z vrcholu 1 do vrcholu 4. Označme túto kvantitu ako $S_{1,4}(t)$. Z učinенých pozorovaní vyplýva, že na vyjadrenie tejto hodnoty pre všetky $t \in \mathbb{N}$ v uzavretom tvare je možné využiť vlastné čísla matice susednosti

$$A := \mathbf{A}(\mathcal{G}) = \begin{pmatrix} 2 & 3 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Zjavne $\text{ch}_A(z) = (z - 2)^3(z - 1)$. Pre maticu A je teda $\sigma(A) = \{\lambda_1, \lambda_2\}$, kde

$$\lambda_1 = 2 \quad \text{a} \quad \lambda_2 = 1,$$

pričom algebraické násobnosti týchto dvoch vlastných čísel sú $\alpha(A, 2) = 3$ a $\alpha(A, 1) = 1$. Existujú preto konštanty $c_{2,0}, c_{2,1}, c_{2,2}$ a $c_{1,0}$ také, že pre všetky $t \in \mathbb{N}$ je

$$S_{1,4}(t) = c_{2,0} \cdot 2^t + c_{2,1} \cdot t \cdot 2^t + c_{2,2} \cdot t^2 \cdot 2^t + c_{1,0}. \quad (5.6)$$

Ľahko pritom nájdeme hodnoty $S_{1,4}(t)$ pre $t = 0, 1, 2, 3$:

$$S_{1,4}(0) = 0, \quad S_{1,4}(1) = 0, \quad S_{1,4}(2) = 2, \quad S_{1,4}(3) = 16.$$

Konštanty $c_{2,0}, c_{2,1}, c_{2,2}$ a $c_{1,0}$ teda získame ako *jediné* riešenie sústavy lineárnych rovníc

$$\begin{aligned} x_{2,0} \cdot 2^0 + x_{2,1} \cdot 0 \cdot 2^0 + x_{2,2} \cdot 0^2 \cdot 2^0 + x_{1,0} &= 0, \\ x_{2,0} \cdot 2^1 + x_{2,1} \cdot 1 \cdot 2^1 + x_{2,2} \cdot 1^2 \cdot 2^1 + x_{1,0} &= 0, \\ x_{2,0} \cdot 2^2 + x_{2,1} \cdot 2 \cdot 2^2 + x_{2,2} \cdot 2^2 \cdot 2^2 + x_{1,0} &= 2, \\ x_{2,0} \cdot 2^3 + x_{2,1} \cdot 3 \cdot 2^3 + x_{2,2} \cdot 3^2 \cdot 2^3 + x_{1,0} &= 16, \end{aligned}$$

čiže po úprave

$$\begin{aligned} x_{2,0} + x_{1,0} &= 0, \\ 2 \cdot x_{2,0} + 2 \cdot x_{2,1} + 2 \cdot x_{2,2} + x_{1,0} &= 0, \\ 4 \cdot x_{2,0} + 8 \cdot x_{2,1} + 16 \cdot x_{2,2} + x_{1,0} &= 2, \\ 8 \cdot x_{2,0} + 24 \cdot x_{2,1} + 72 \cdot x_{2,2} + x_{1,0} &= 16. \end{aligned}$$

Jediné riešenie tejto sústavy je dané ako

$$\begin{aligned} x_{2,0} &= 4 = c_{2,0}, \\ x_{2,1} &= -\frac{11}{4} = c_{2,1}, \\ x_{2,2} &= \frac{3}{4} = c_{2,2}, \\ x_{1,0} &= -4 = c_{1,0}. \end{aligned}$$

Dosadením do (5.6) tak napokon zisťujeme, že pre všetky $t \in \mathbb{N}$ je

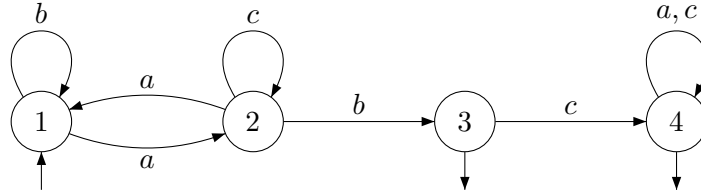
$$S_{1,4}(t) = \frac{3}{4} \cdot t^2 \cdot 2^t - \frac{11}{4} \cdot t \cdot 2^t + 4 \cdot 2^t - 4.$$

Skúškou správnosti tu môže byť napríklad dosadenie niekoľkých hodnôt $t = 4, 5, \dots$ a konfrontácia takto získaných výsledkov s hodnotami „nameranými“ prehľadávaním grafu.

5.3 Počet slov dĺžky t v racionálnom jazyku

Už dávnejšie sme prišli k pozorovaniu, že problém enumerácie slov dĺžky t v danom racionálnom jazyku L možno previesť na problém enumerácie sledov vedúcich medzi niekoľkými dvojicami vrcholov multigrafu získaného z diagramu *deterministického* (alebo všeobecnejšie: *jednoznačného*) konečného automatu $\mathcal{A}$ takého, že $\|\mathcal{A}\| = L$. V nasledujúcom príklade budeme opäť postupovať týmto spôsobom; na enumeráciu sledov ale použijeme metódu z predchádzajúceho oddielu. Pri troche úsilia nie je ťažké odhaliť úzky súvis nižšie opísanej enumeračnej metódy založenej na lineárnej algebre s metódou založenou na vytvárajúcich funkciách, ktorú sme si predstavili v príklade 3.5.14.

Príklad 5.3.1. Uvažujme deterministický⁵ konečný automat $\mathcal{A}$ na obrázku 5.2 a predpokladajme, že potrebujeme pre všetky $t \in \mathbb{N}$ vyčísliť počet slov dĺžky t v racionálnom jazyku $\|\mathcal{A}\|$. Zjavne ide o jazyk nad abecedou $\Sigma = \{a, b, c\}$; zaujíma nás teda pre každé $t \in \mathbb{N}$ hodnota $|\|\mathcal{A}\| \cap \Sigma^t|$.



Obr. 5.2: Deterministický konečný automat $\mathcal{A}$.

Lahko vidieť, že pre všetky $t \in \mathbb{N}$ je

$$|\|\mathcal{A}\| \cap \Sigma^t| = S_{1,3}(t) + S_{1,4}(t),$$

kde $S_{1,3}(t)$ je počet sledov dĺžky t v multigrafe prislúchajúcom k diagramu automatu $\mathcal{A}$ vedúcich z vrcholu (stavu) 1 do vrcholu (stavu) 3 a $S_{1,4}(t)$ má rovnaký význam pre vrcholy (stavy) 1 a 4.

Hodnoty $S_{1,3}(t)$ a $S_{1,4}(t)$ teraz môžeme vyčísliť rovnako ako v predchádzajúcom oddiele – to jest prostredníctvom výpočtu vlastných čísel príslušnej matice susednosti

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

Charakteristický polynóm matice A je daný ako

$$\text{ch}_A(z) = z^4 - 4z^3 + 4z^2 = (z - 2)^2 z^2.$$

Preto $\sigma(A) = \{2, 0\}$, pričom $\alpha(A, 2) = 2$ a $\alpha(A, 0) = 2$. Existujú teda konštanty $c_{2,0}, c_{2,1}, c_{0,0}, c_{0,1} \in \mathbb{C}$ také, že pre všetky $t \in \mathbb{N}$ je

$$S_{1,3}(t) = c_{2,0} \cdot 2^t + c_{2,1} \cdot t \cdot 2^t + c_{0,0} \cdot \delta_{t,0} + c_{0,1} \cdot \delta_{t,1}, \quad (5.7)$$

ako aj konštanty $d_{2,0}, d_{2,1}, d_{0,0}, d_{0,1} \in \mathbb{C}$ také, že pre všetky $t \in \mathbb{N}$ je

$$S_{1,4}(t) = d_{2,0} \cdot 2^t + d_{2,1} \cdot t \cdot 2^t + d_{0,0} \cdot \delta_{t,0} + d_{0,1} \cdot \delta_{t,1}. \quad (5.8)$$

Riešením sústavy lineárnych rovníc

$$\begin{aligned} x_{2,0} + x_{0,0} &= 0 \\ 2 \cdot x_{2,0} + 2 \cdot x_{2,1} + x_{0,1} &= 0 \\ 4 \cdot x_{2,0} + 8 \cdot x_{2,1} &= 1 \\ 8 \cdot x_{2,0} + 24 \cdot x_{2,1} &= 2 \end{aligned}$$

potom zistujeme, že $c_{2,0} = \frac{1}{4}$, $c_{2,1} = 0$, $c_{0,0} = -\frac{1}{4}$ a $c_{0,1} = -\frac{1}{2}$. Podobne riešením sústavy

$$\begin{aligned} y_{2,0} + y_{0,0} &= 0 \\ 2 \cdot y_{2,0} + 2 \cdot y_{2,1} + y_{0,1} &= 0 \\ 4 \cdot y_{2,0} + 8 \cdot y_{2,1} &= 0 \\ 8 \cdot y_{2,0} + 24 \cdot y_{2,1} &= 1 \end{aligned}$$

⁵Presnejšie ide o deterministický konečný automat s čiastočnou prechodovou funkciou. To je v poriadku – ako sme už spomenuli, rovnaká argumentácia by bola namieste aj pre ľubovoľný jednoznačný konečný automat.

zistujeme, že $d_{2,0} = -\frac{1}{4}$, $d_{2,1} = \frac{1}{8}$, $d_{0,0} = \frac{1}{4}$ a $d_{0,1} = \frac{1}{4}$. Dosadením do (5.7) a do (5.8) následne dostávame

$$S_{1,3}(t) = \frac{1}{4} \cdot 2^t - \frac{1}{4} \cdot \delta_{t,0} - \frac{1}{2} \cdot \delta_{t,1}$$

a

$$S_{1,4}(t) = \frac{1}{8} \cdot t \cdot 2^t - \frac{1}{4} \cdot 2^t + \frac{1}{4} \cdot \delta_{t,0} + \frac{1}{4} \cdot \delta_{t,1},$$

z čoho už hľadaný počet slov dĺžky t v jazyku $\|\mathcal{A}\|$ vieme vyjadriť ako

$$|\|\mathcal{A}\| \cap \Sigma^t| = S_{1,3}(t) + S_{1,4}(t) = \frac{1}{8} \cdot t \cdot 2^t - \frac{1}{4} \cdot \delta_{t,1}.$$

Ako skúška správnosti môže opäť poslúžiť konfrontácia tohto výsledku pre $t = 4, 5, \dots$ s hodnotami „nameranými“ prehľadávaním automatu.

5.4 Perronova-Frobeniova teória nezáporných matíc

Ľubovoľnú maticu nad množinou nezáporných reálnych čísel $\mathbb{R}_{\geq 0}$ nazveme *nezápornou*. Budeme sa teraz bližšie zaoberať – predovšetkým spektrálnymi – vlastnosťami nezáporných štvorcových matíc. Pozorovania, ku ktorým sa dopracujeme, budú tvoriť jadro takzvanej *Perronovej-Frobeniovej teórie*.

Teória nezáporných matíc patrí k tým oblastiam matematiky, v ktorých vzájomný súvis medzi maticami a grafmi zaujíma výsostné postavenie. Toto prepojenie bude pre nás užitočné „v oboch smeroch“: lepšie pochopenie štruktúry spektra nezáporných matíc využijeme pri riešení rozličných úloh súvisiacich s ohodnotenými grafmi nad $\mathbb{R}_{\geq 0}$ – a teda aj nad $\mathbb{N}$, či $\mathbb{B}$ – a grafová perspektíva nám naopak často podstatne zjednoduší a zintuitívni argumentáciu o nezáporných štvorcových maticiach.

Naše skúmanie začneme definíciou asi najvýznamnejšej triedy nezáporných štvorcových matíc – takzvaných *ireducibilných matíc*. Nepôjde však o žiadnu skutočne novú myšlienku: hneď vzápätí totiž ireducibilné matice charakterizujeme pomocou dôverne známeho pojmu silne súvislých grafov.

Najskôr si ale pripomeňme, že *permutačná matica* je štvorcová matica pozostávajúca z jednotiek a núl, ktorá má v každom riadku aj stĺpci práve jednu jednotku. Pre permutačnú maticu P typu $n \times n$ teda existuje permutácia $\pi: [n] \rightarrow [n]$ taká, že $P = (p_{i,j})_{n \times n}$, kde pre $k = 1, \dots, n$ je $p_{k,\pi(k)} = 1$ a $p_{k,j} = 0$ pre $j = 1, \dots, \pi(k) - 1, \pi(k) + 1, \dots, n$. Hovoríme potom, že P zodpovedá permutácii π . Každá permutačná matica je navyše nesingulárna a platí $P^{-1} = P^T$, pričom táto matica zodpovedá permutácii π^{-1} . Ľahko ďalej vidieť, že pre ľubovoľné nenulové prirodzené číslo n , ľubovoľnú maticu $A = (a_{i,j})_{n \times n}$ a ľubovoľnú permutačnú maticu P typu $n \times n$ zodpovedajúcu permutácii π je

$$PAP^{-1} = \begin{pmatrix} a_{\pi(1),\pi(1)} & a_{\pi(1),\pi(2)} & \cdots & a_{\pi(1),\pi(n)} \\ a_{\pi(2),\pi(1)} & a_{\pi(2),\pi(2)} & \cdots & a_{\pi(2),\pi(n)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\pi(n),\pi(1)} & a_{\pi(n),\pi(2)} & \cdots & a_{\pi(n),\pi(n)} \end{pmatrix}.$$

Ide teda o „povymieňanie riadkov aj stĺpcov matice A podľa permutácie π “.

Definícia 5.4.1. Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je matica. Maticu A nazveme *reducibilnou*, ak $n \geq 2$ a existuje permutačná matica P taká, že

$$PAP^{-1} = \begin{pmatrix} B_{1,1} & B_{1,2} \\ \mathbf{0} & B_{2,2} \end{pmatrix},$$

kde pre nejaké nenulové prirodzené čísla n_1, n_2 spĺňajúce $n_1 + n_2 = n$ sú $B_{1,1}, B_{1,2}, B_{2,2}$ bloky typov $n_1 \times n_1$, $n_1 \times n_2$, resp. $n_2 \times n_2$ a kde $\mathbf{0}$ označuje nulový blok typu $n_2 \times n_1$. Matica A je *ireducibilná*, ak nie je reducibilná.

Poznámka 5.4.2. Každú nezápornú maticu typu 0×0 alebo 1×1 teda považujeme za ireducibilnú.

Tvrdenie 5.4.3. *Nech $n \in \mathbb{N}$, $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je matica a $\mathcal{G} = ([n], \varphi)$ je ohodnotený graf nad $\mathbb{R}_{\geq 0}$ taký, že $A = \mathbf{A}(\mathcal{G})$ – alebo ekvivalentne: $\mathcal{G} = \bar{\text{Gr}}(A)$. Potom je matica A ireducibilná práve vtedy, keď je graf $\mathcal{G}$ silne súvislý.⁶*

Dôkaz. Pre $n = 0$ a $n = 1$ je tvrdenie zřejmé – každá matica typu 0×0 alebo 1×1 je ireducibilná a každý graf s najviac jedným vrcholom je silne súvislý. Pre $n \geq 2$ z definície 5.4.1 vyplýva, že matica A je *reducibilná* práve vtedy, keď existuje permutácia $\pi: [n] \rightarrow [n]$ a nenulové prirodzené čísla n_1, n_2 také, že $n_1 + n_2 = n$ a pre $i = n_1 + 1, \dots, n_1 + n_2$ a $j = 1, \dots, n_1$ je $a_{\pi(i), \pi(j)} = 0$. To znamená, že pre neprázdne podmnožiny $V_1 = \pi(\{1, \dots, n_1\})$ a $V_2 = \pi(\{n_1 + 1, \dots, n_1 + n_2\})$ množiny vrcholov grafu $\mathcal{G}$ nevedie z vrcholov množiny V_2 v grafe $\mathcal{G}$ žiadna hrana do vrcholu množiny V_1 . Takáto situácia ale očividne nastáva práve vtedy, keď graf $\mathcal{G}$ *nie* je silne súvislý. $\square$

Ako okamžitý dôsledok práve dokázaného tvrdenia dostávame, že nezáporná matica A typu $n \times n$ je ireducibilná práve vtedy, keď pre $u, v = 1, \dots, n$ existuje prirodzené číslo t také, že v matici A^t je na pozícii v u -tom riadku a v -tom stĺpci kladné reálne číslo.

Dôsledok 5.4.4. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je matica. Matica A je ireducibilná práve vtedy, keď pre všetky dvojice indexov $u, v \in [n]$ existuje $t \in \mathbb{N}$ také, že pre $A^t = (a_{i,j}[t])_{n \times n}$ je $a_{u,v}[t] > 0$.*

Dôkaz. Z predchádzajúceho tvrdenia vyplýva, že matica A je ireducibilná práve vtedy, keď pre všetky $u, v \in [n]$ existuje v grafe $\text{Gr}(A) = ([n], \varphi)$ aspoň jeden sled z u do v . Pre všetky $u, v \in [n]$ teda musí pre nejaké $t \in \mathbb{N}$ existovať v tomto grafe sled z u do v dĺžky t . V takom prípade musí pre *každý* sled $\gamma \in \mathcal{W}_{u,v}(\text{Gr}(A), t)$ byť $\varphi(\gamma) > 0$;⁷ podľa vety 1.3.10 teda $a_{u,v}[t] > 0$. Ak naopak žiaden sled dĺžky t vedúci z u do v neexistuje, z vety 1.3.10 vyplýva $a_{u,v}[t] = 0$. Uvedená podmienka je teda ekvivalentná existencii $t \in \mathbb{N}$ takého, že $a_{u,v}[t] > 0$. $\square$

Dôsledkom tvrdenia 5.4.3 je aj nasledujúca ekvivalentná charakterizácia ireducibilných matíc, ktorá sa nám neskôr zídne. Maticu nazveme *kladnou*, ak sú všetky jej prvky kladné reálne čísla.

Dôsledok 5.4.5. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je matica. Matica A je ireducibilná práve vtedy, keď je matica $(A + \mathbf{I})^{n-1}$ kladná. Táto situácia nastane práve vtedy, keď sú pre každý vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ kladné všetky zložky vektora $(A + \mathbf{I})^{n-1}\mathbf{x}$.*

Dôkaz. Matica A je podľa tvrdenia 5.4.3 ireducibilná práve vtedy, keď pre všetky $u, v \in [n]$ existuje v grafe $\text{Gr}(A)$ aspoň jeden sled z vrcholu u do vrcholu v – čo nastane práve vtedy, keď pre všetky $u, v \in [n]$ existuje aspoň jeden sled z u do v v grafe $\text{Gr}(A + \mathbf{I})$, ktorý oproti $\text{Gr}(A)$ obsahuje maximálne niekoľko slučiek navyše. Najkratší taký sled γ je ale dĺžky najviac $n - 1$, pričom s použitím slučiek ho možno predĺžiť na sled dĺžky práve $n - 1$. Matica A je teda ireducibilná práve vtedy, keď pre všetky dvojice vrcholov $u, v \in [n]$ existuje v grafe $\text{Gr}(A + \mathbf{I})$ sled z u do v dĺžky presne $n - 1$. To sa vďaka vete 1.3.10 a kladným ohodnoteniam jednotlivých hrán stane práve vtedy, keď sú na všetkých pozíciách matice $(A + \mathbf{I})^{n-1}$ kladné reálne čísla.

Je zřejmé, že pre ľubovoľnú kladnú maticu B typu $n \times n$ a ľubovoľný vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ musí byť vektor $B\mathbf{x}$ kladný. Ak naopak matica $B = (b_{i,j})_{n \times n} \in \mathbb{R}_{\geq 0}^{n \times n}$ nie je kladná, existujú $u, v \in [n]$ také, že $b_{u,v} = 0$. Pre ľubovoľný vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ s nenulovou v -tou zložkou a ostatnými zložkami nulovými je potom $B\mathbf{x}$ vektor, ktorého u -ta zložka je nulová. Nezáporná štvorcová matica B rádu n je teda kladná práve vtedy, keď je pre všetky $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ kladný vektor $B\mathbf{x}$; na dôkaz druhej časti dôsledku tak stačí vziať $B = (A + \mathbf{I})^{n-1}$. $\square$

⁶Graf $\mathcal{G} = (V, \varphi)$ je silne súvislý, ak pre všetky $u, v \in V(\mathcal{G})$ existuje v grafe $\mathcal{G}$ sled z vrcholu u do vrcholu v .

⁷Sled totiž musí pozostávať z hrán ohodnotených nenulovými prvkami množiny $\mathbb{R}_{\geq 0}$. Každá hrana v slede γ je teda ohodnotená kladným reálnym číslom, v dôsledku čoho je kladné aj ohodnotenie samotného sledu γ : pre prázdny sled je týmto ohodnotením 1 a pre neprázdne sledy ide o súčin nenulového počtu kladných reálnych čísel.

V nasledujúcom budeme pre nenulové prirodzené číslo n a dvojicu nezáporných reálnych vektorov

$$\mathbf{x} = (x_1, \dots, x_n)^T, \quad \mathbf{y} = (y_1, \dots, y_n)^T$$

písať $\mathbf{x} \leq \mathbf{y}$, ak pre $k = 1, \dots, n$ je $x_k \leq y_k$ a $\mathbf{x} \geq \mathbf{y}$, ak $\mathbf{y} \leq \mathbf{x}$.

Analýzu spektra nezápornej štvorcovej matice $A \in \mathbb{R}_{\geq 0}^{n \times n}$ do veľkej miery založíme na vlastnostiach takzvanej *Collatzovej-Wielandtovej funkcie* Γ_A prislúchajúcej k matici A , ktorú teraz zdefinujeme. Pôjde o funkciu, ktorá pre každý vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ udáva *najväčšie* nezáporné reálne číslo α také, že $A\mathbf{x} \geq \alpha\mathbf{x}$; túto jej vlastnosť explicitne sformulujeme v tvrdení 5.4.7. Prezentácia nasledujúcich partií tohto oddielu je do veľkej miery ovplyvnená knihou [86].

Definícia 5.4.6. Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je matica. *Collatzovou-Wielandtovou funkciou* prislúchajúcou k A nazveme zobrazenie $\Gamma_A: \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}_{\geq 0}$ dané pre všetky $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ a pre $A\mathbf{x} = (x'_1, \dots, x'_n)^T$ predpisom

$$\Gamma_A(\mathbf{x}) = \min_{\substack{k=1, \dots, n \\ x_k \neq 0}} \frac{x'_k}{x_k}.$$

Tvrdenie 5.4.7. Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$. Pre všetky $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ potom

$$\Gamma_A(\mathbf{x}) = \max \{ \alpha \in \mathbb{R}_{\geq 0} \mid A\mathbf{x} \geq \alpha\mathbf{x} \}.$$

Dôkaz. Nech $\mathbf{x} = (x_1, \dots, x_n)^T$ a $A\mathbf{x} = (x'_1, \dots, x'_n)^T$. Z definície Collatzovej-Wielandtovej funkcie potom vyplýva, že pre $k = 1, \dots, n$ je $x'_k \geq \Gamma_A(\mathbf{x})x_k$, z čoho $A\mathbf{x} \geq \Gamma_A(\mathbf{x})\mathbf{x}$. Musí však zároveň existovať $k \in [n]$ také, že $x_k \neq 0$ a $x'_k = \Gamma_A(\mathbf{x})x_k$. Pre všetky $\varepsilon > 0$ potom $x'_k < (\Gamma_A(\mathbf{x}) + \varepsilon)x_k$, a teda nemôže platiť nerovnosť $A\mathbf{x} \geq (\Gamma_A(\mathbf{x}) + \varepsilon)\mathbf{x}$. $\square$

Dokážeme teraz jednoduchú lemu, podľa ktorej pre násobenie vektora $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ kladným reálnym číslom – čiže skalárom z $\mathbb{R}_{> 0}$ – nemení hodnotu Collatzovej-Wielandtovej funkcie. Toto pozorovanie nám neskôr v prípade potreby umožní vektory normalizovať – napríklad tak, aby manhattanská norma

$$\|\mathbf{x}\|_1 = |x_1| + \dots + |x_n| = x_1 + \dots + x_n$$

nezáporného reálneho vektora $\mathbf{x} = (x_1, \dots, x_n)^T$ rôzneho od $\mathbf{0}$ bola rovná jednej.

Lema 5.4.8. Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$. Pre všetky $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ a $q \in \mathbb{R}_{> 0}$ potom

$$\Gamma_A(q\mathbf{x}) = \Gamma_A(\mathbf{x}).$$

Dôkaz. Z tvrdenia 5.4.7 priamo dostávame

$$\Gamma_A(q\mathbf{x}) = \max \{ \alpha \in \mathbb{R}_{\geq 0} \mid Aq\mathbf{x} \geq \alpha q\mathbf{x} \} = \max \{ \alpha \in \mathbb{R}_{\geq 0} \mid A\mathbf{x} \geq \alpha\mathbf{x} \} = \Gamma_A(\mathbf{x}). \quad \square$$

Pre každé nenulové prirodzené číslo n nazveme *štandardným $(n-1)$ -simplexom* podmnožinu Δ_{n-1} množiny $\mathbb{R}_{\geq 0}^n$ danú ako

$$\Delta_{n-1} = \left\{ (x_1, \dots, x_n)^T \in \mathbb{R}_{\geq 0}^n \mid \sum_{k=1}^n x_k = 1 \right\} = \{ \mathbf{x} \in \mathbb{R}_{\geq 0}^n \mid \|\mathbf{x}\|_1 = 1 \}.$$

Dokážeme, že Collatzova-Wielandtova funkcia Γ_A prislúchajúca k *ireducibilnej*⁸ matici $A \in \mathbb{R}_{\geq 0}^{n \times n}$ nadobúda na štandardnom simplexe Δ_{n-1} maximum. Štandardný simplex Δ_{n-1} je zrejme kompaktnou

⁸Rovnakú vlastnosť majú aj reducibilné nezáporné matice. Toto nie príliš dôležité tvrdenie nebudeme dokazovať, pôjde však o bezprostredný dôsledok našich neskorších úvah.

– čiže ohraničenou a uzavretou⁹ – podmnožinou $\mathbb{R}^n$. Ľubovoľná funkcia $f: \mathbb{R}^n \rightarrow \mathbb{R}$ spojitá na Δ_{n-1} teda na tejto množine nadobúda maximum. Ľahko možno dokázať, že Collatzova-Wielandtova funkcia je vždy spojitá v bodoch množiny $\mathbb{R}_{\geq 0}^n$ – to je náplňou cvičenia 3 na konci tejto kapitoly. Existuje však ireducibilná matica A taká, že k nej prislúchajúca Collatzova-Wielandtova funkcia nie je spojitá na $\Delta_{n-1} \setminus \mathbb{R}_{> 0}^n$, čiže v hraničných bodoch simplexu Δ_{n-1} ; nájdenie príkladu takejto matice je zas náplňou cvičenia 4 na konci kapitoly. Dôkaz nasledujúcej lemy tak bude o čosi menej priamočiary.

Lema 5.4.9. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná nezáporná matica. Collatzova-Wielandtova funkcia $\Gamma_A: \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}_{\geq 0}$ prislúchajúca k A potom na Δ_{n-1} nadobúda maximum.*

Dôkaz. Označme symbolom Δ_A množinu

$$\Delta_A = \{ (A + \mathbf{I})^{n-1} \mathbf{x} \mid \mathbf{x} \in \Delta_{n-1} \}.$$

Tá je očividne kompaktná – čiže ohraničená a uzavretá – a z dôsledku 5.4.5 vyplýva, že $\Delta_A \subseteq \mathbb{R}_{\geq 0}^n$. Collatzova-Wielandtova funkcia Γ_A je preto spojitá vo všetkých bodoch množiny Δ_A a v nejakom bode $\mathbf{y}_{max} \in \Delta_A$ teda nadobúda svoje maximum na Δ_A – čiže

$$\Gamma_A(\mathbf{y}_{max}) \geq \Gamma_A(\mathbf{y}) \quad (5.9)$$

pre všetky $\mathbf{y} \in \Delta_A$. Položme

$$\mathbf{x}_{max} := \frac{1}{\|\mathbf{y}_{max}\|_1} \mathbf{y}_{max},$$

pričom zjavne $\mathbf{x}_{max} \in \Delta_{n-1}$. Ukážeme, že pre všetky $\mathbf{x} \in \Delta_{n-1}$ je

$$\Gamma_A(\mathbf{x}_{max}) \geq \Gamma_A(\mathbf{x}), \quad (5.10)$$

čím bude lema dokázaná. Nech je teda vektor $\mathbf{x} \in \Delta_{n-1}$ daný pevne. Z tvrdenia 5.4.7 potom

$$A\mathbf{x} \geq \Gamma_A(\mathbf{x})\mathbf{x}$$

a pre násobením oboch strán zľava maticou $(A + \mathbf{I})^{n-1}$ dostávame

$$(A + \mathbf{I})^{n-1} A\mathbf{x} \geq \Gamma_A(\mathbf{x})(A + \mathbf{I})^{n-1} \mathbf{x},$$

z čoho – pretože podľa lemy 4.5.6 je $(A + \mathbf{I})^{n-1} A = A(A + \mathbf{I})^{n-1}$ – vyplýva

$$A[(A + \mathbf{I})^{n-1} \mathbf{x}] \geq \Gamma_A(\mathbf{x})[(A + \mathbf{I})^{n-1} \mathbf{x}].$$

Podľa tvrdenia 5.4.7 potom

$$\Gamma_A((A + \mathbf{I})^{n-1} \mathbf{x}) \geq \Gamma_A(\mathbf{x}). \quad (5.11)$$

Teraz už môžeme pristúpiť k samotnému dôkazu nerovnosti (5.10):

$$\Gamma_A(\mathbf{x}) \stackrel{(5.11)}{\leq} \Gamma_A((A + \mathbf{I})^{n-1} \mathbf{x}) \stackrel{(5.9)}{\leq} \Gamma_A(\mathbf{y}_{max}) = \Gamma_A(\|\mathbf{y}_{max}\|_1 \mathbf{x}_{max}) \stackrel{\text{Lema 5.4.8}}{=} \Gamma_A(\mathbf{x}_{max}),$$

čo bolo treba dokázať. □

S pomocou lemy 5.4.8 následne zisťujeme, že ľubovoľné maximum Collatzovej-Wielandtovej funkcie na štandardnom simplexe Δ_{n-1} je aj jej globálnym maximom – čiže maximom na celom jej definičnom obore $\mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$. Prichádzame tak k nasledujúcemu dôsledku.

⁹Ide o jednu z možných definícií kompaktnosti pre podmnožiny $\mathbb{C}^n$, ktorá pre ne súhlasí s topologickou definíciou kompaktnosti vďaka tzv. Heineho-Borelovej vete. Čitateľa tu len odkážeme napríklad na [99].

Dôsledok 5.4.10. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica. Collatzova-Wielandtova funkcia $\Gamma_A: \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}_{\geq 0}$ potom nadobúda aspoň jedno globálne maximum.*

Dôkaz. Nech $\mathbf{x}_{\max} \in \Delta_{n-1}$ je niektorý z bodov, v ktorých funkcia Γ_A nadobúda svoje maximum na štandardnom $(n-1)$ -simplexe – jeho existencia vyplýva z lemy 5.4.9. Ukážeme, že pre všetky $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ je

$$\Gamma_A(\mathbf{x}) \leq \Gamma_A(\mathbf{x}_{\max}).$$

Zvoľme vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ ľubovoľne. Vektor

$$\frac{1}{\|\mathbf{x}\|_1} \mathbf{x}$$

potom patrí do Δ_{n-1} a z lemy 5.4.8 dostávame

$$\Gamma_A(\mathbf{x}) = \Gamma_A\left(\frac{1}{\|\mathbf{x}\|_1} \mathbf{x}\right) \leq \Gamma_A(\mathbf{x}_{\max}),$$

čím je dôkaz dokončený. □

Z lemy 5.4.8 tiež vidieť, že Collatzova-Wielandtova funkcia je vždy jednoznačne daná jej zúžením na Δ_{n-1} . Hodnota tejto funkcie na ľubovoľnom vstupnom vektore $\mathbf{x}$ je totiž rovnaká ako na $(1/\|\mathbf{x}\|_1) \mathbf{x}$, čo je bod priestoru $\mathbb{R}^n$, ktorý si možno predstaviť ako priesečník jeho nadroviny obsahujúcej Δ_{n-1} a priamky prechádzajúcej cez body $\mathbf{x}$ a $\mathbf{0}$.

Môžeme teraz prejsť k skúmaniu vlastných čísel a vlastných vektorov ireducibilných nezáporných matic. Začnime jednoduchou lemov, podľa ktorej je ľubovoľný nezáporný vlastný vektor takejto matice v skutočnosti kladný.

Lema 5.4.11. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica. Nech $\mathbf{v} \in \mathbb{R}_{\geq 0}^n$ je vlastný vektor matice A . Potom $\mathbf{v} \in \mathbb{R}_{> 0}^n$.*

Dôkaz. Vlastný vektor musí byť nenulový; z predpokladu $\mathbf{v} \in \mathbb{R}_{\geq 0}^n$ preto vyplýva $\mathbf{v} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$. Ak je ďalej $\lambda \in \mathbb{C}$ vlastné číslo prislúchajúce k $\mathbf{v}$, musí byť $A\mathbf{v} = \lambda\mathbf{v}$, a teda – podľa vety 4.8.3 – aj

$$(A + \mathbf{I})\mathbf{v} = (\lambda + 1)\mathbf{v}.$$

V dôsledku toho

$$(A + \mathbf{I})^{n-1}\mathbf{v} = (\lambda + 1)^{n-1}\mathbf{v}. \quad (5.12)$$

Za účelom sporu teraz predpokladajme, že je niektorá zložka vektora $\mathbf{v}$ nulová. Potom musí byť nulová aj niektorá zložka vektora $(\lambda + 1)^{n-1}\mathbf{v}$. Keďže ale $\mathbf{v} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$, musí byť podľa dôsledku 5.4.5 vektor $(A + \mathbf{I})^{n-1}\mathbf{v}$ kladný. Tieto dve skutočnosti však dohromady očividne odporujú rovnosti (5.12). □

Pre ľubovoľné nenulové prirodzené číslo n a ľubovoľný vektor $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{C}^n$ ďalej definujme vektor $|\mathbf{x}|$ ako

$$|\mathbf{x}| := (|x_1|, \dots, |x_n|)^T \in \mathbb{R}_{\geq 0}^n.$$

Podobne pre maticu $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ bude $|A|$ označovať maticu $|A| := (|a_{i,j}|)_{n \times n} \in \mathbb{R}_{\geq 0}^{n \times n}$. Dôsledkom nasledujúcej lemy okrem iného bude, že pre nezápornú maticu $A \in \mathbb{R}_{\geq 0}^{n \times n}$ a ľubovoľný jej vlastný vektor $\mathbf{v} \in \mathbb{C}^n$ prislúchajúci k vlastnému číslu λ platí $A|\mathbf{v}| \geq |\lambda| \cdot |\mathbf{v}|$.

Lema 5.4.12. *Nech $n \in \mathbb{N} \setminus \{0\}$, $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ a $\mathbf{v} = (v_1, \dots, v_n)^T \in \mathbb{C}^n$ je vlastný vektor matice A prislúchajúci k vlastnému číslu $\lambda \in \sigma(A)$. Potom*

$$|A| \cdot |\mathbf{v}| \geq |\lambda| \cdot |\mathbf{v}|$$

Dôkaz. Keďže je $\mathbf{v}$ vlastný vektor matice A prislúchajúci k vlastnému číslu λ , platí

$$A\mathbf{v} = \lambda\mathbf{v},$$

z čoho pre $k = 1, \dots, n$ dostávame

$$\sum_{j=1}^n a_{k,j} v_j = \lambda v_k$$

a v dôsledku toho aj

$$\sum_{j=1}^n |a_{k,j}| \cdot |v_j| \geq |\lambda| \cdot |v_k|.$$

Pre vektor $|\mathbf{v}| = (|v_1|, \dots, |v_n|)^T$ je preto

$$|A| \cdot |\mathbf{v}| \geq |\lambda| \cdot |\mathbf{v}|,$$

čo bolo treba dokázať. □

Spektrálnym polomerom matice – môže pritom ísť o ľubovoľnú štvorcovú maticu komplexných čísel typu $n \times n$ pre $n > 0$ – rozumieme najväčšiu spomedzi absolútnych hodnôt jej vlastných čísel.

Definícia 5.4.13. Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{C}^{n \times n}$. *Spektrálnym polomerom* matice A nazveme nezáporné reálne číslo $\varrho(A)$ dané ako

$$\varrho(A) = \max \{ |\lambda| \mid \lambda \in \sigma(A) \}.$$

Dokážeme teraz prvú z kľúčových vlastností spektier ireducibilných nezáporných matíc: spektrálny polomer každej takejto matice je jej vlastným číslom. Medzi vlastnými číslami s maximálnou absolútnou hodnotou je tak pre ireducibilnú nezápornú maticu vždy nezáporné reálne číslo.

Veta 5.4.14. Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica. Potom je $\varrho := \varrho(A)$ vlastným číslom matice A , ku ktorému prislúcha aspoň jeden kladný vlastný vektor. Hodnota ϱ je navyše globálnym maximom Collatzovej-Wielandtovej funkcie Γ_A a vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ je vlastným vektorom matice A prislúchajúcim k ϱ práve vtedy, keď $\Gamma_A(\mathbf{x}) = \varrho$; každý takýto vektor $\mathbf{x}$ pritom patrí do $\mathbb{R}_{> 0}^n$.

Dôkaz. Uvažujme ľubovoľný vektor $\mathbf{x}_{\max} \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$, pre ktorý funkcia Γ_A nadobúda svoje globálne maximum – existencia aspoň jedného takého vektora je zaručená vďaka dôsledku 5.4.10. Ukážeme najprv, že $\mathbf{x}_{\max}$ je vlastný vektor matice A s prislúchajúcim vlastným číslom $\Gamma_A(\mathbf{x}_{\max})$.

Z tvrdenia 5.4.7 vyplýva, že musí platiť nerovnosť

$$A\mathbf{x}_{\max} \geq \Gamma_A(\mathbf{x}_{\max})\mathbf{x}_{\max}.$$

Potrebuje ukázať, že táto nerovnosť je v skutočnosti rovnosťou. Za účelom sporu predpokladajme, že to tak nie je; vektor

$$A\mathbf{x}_{\max} - \Gamma_A(\mathbf{x}_{\max})\mathbf{x}_{\max}$$

je potom prvkom $\mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$. Podľa dôsledku 5.4.5 je teda vektor

$$\begin{aligned} (A + \mathbf{I})^{n-1} (A\mathbf{x}_{\max} - \Gamma_A(\mathbf{x}_{\max})\mathbf{x}_{\max}) &= (A + \mathbf{I})^{n-1} A\mathbf{x}_{\max} - \Gamma_A(\mathbf{x}_{\max})(A + \mathbf{I})^{n-1} \mathbf{x}_{\max} = \\ &= A[(A + \mathbf{I})^{n-1} \mathbf{x}_{\max}] - \Gamma_A(\mathbf{x}_{\max})[(A + \mathbf{I})^{n-1} \mathbf{x}_{\max}] \end{aligned}$$

kladný. Existuje teda $\varepsilon \in \mathbb{R}_{> 0}$ také, že

$$A[(A + \mathbf{I})^{n-1} \mathbf{x}_{\max}] - (\Gamma_A(\mathbf{x}_{\max}) + \varepsilon)[(A + \mathbf{I})^{n-1} \mathbf{x}_{\max}] \geq \mathbf{0}$$

a z tvrdenia 5.4.7 dostávame

$$\Gamma_A((A + \mathbf{I})^{n-1} \mathbf{x}_{max}) \geq \Gamma_A(\mathbf{x}_{max}) + \varepsilon,$$

čo odporuje predpokladu, že $\Gamma_A(\mathbf{x}_{max})$ je globálne maximum funkcie Γ_A . Nutne teda

$$A\mathbf{x}_{max} = \Gamma_A(\mathbf{x}_{max})\mathbf{x}_{max}$$

a $\mathbf{x}_{max}$ je tak skutočne vlastným vektorom matice A prislúchajúcim k vlastnému číslu $\Gamma_A(\mathbf{x}_{max})$.

Dokážeme, že $\Gamma_A(\mathbf{x}_{max}) = \varrho$. Za týmto účelom zjavne stačí ukázať, že pre každé vlastné číslo $\lambda \in \sigma(A)$ je $|\lambda| \leq \Gamma_A(\mathbf{x}_{max})$. Nech je ale vlastné číslo λ dané pevne a $\mathbf{v}$ je ľubovoľný k nemu prislúchajúci vlastný vektor. Z lemy 5.4.12 potom

$$A|\mathbf{v}| \geq |\lambda| \cdot |\mathbf{v}|,$$

kde $|\mathbf{v}| \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$. S použitím tvrdenia 5.4.7 teda zisťujeme, že

$$|\lambda| \leq \Gamma_A(|\mathbf{v}|) \leq \Gamma_A(\mathbf{x}_{max}).$$

Každý vlastný vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n$ matice A prislúchajúci k ϱ napokon podľa lemy 5.4.11 patrí do $\mathbb{R}_{>0}^n$, pričom z rovnosti

$$A\mathbf{x} = \varrho\mathbf{x}$$

a tvrdenia 5.4.7 vyplýva $\Gamma_A(\mathbf{x}) \geq \varrho$. Keďže je teda ϱ globálnym maximom Collatzovej-Wielandtovej funkcie Γ_A , musí byť $\Gamma_A(\mathbf{x}) = \varrho$ a dokázané sú aj zvyšné časti vety. $\square$

Spektrálny polomer $\varrho(A)$ ireducibilnej nezápornej matice A je teda vždy jeho vlastným číslom. Tento poznatok teraz ešte o niečo upresníme: dokážeme, že algebraická násobnosť vlastného čísla $\varrho(A)$ je pre ireducibilnú nezápornú maticu A vždy rovná jednej a spektrálny polomer $\varrho(A)$ tak je jednoduchým koreňom charakteristického polynómu ireducibilnej matice A .

Veta 5.4.15. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica. Pre vlastné číslo $\varrho := \varrho(A)$ matice A je potom $\alpha(A, \varrho) = 1$.*

Dôkaz. Spektrálny polomer $\varrho = \varrho(A)$ je vlastným číslom matice A podľa vety 5.4.14; znenie dokazovaného tvrdenia teda dáva zmysel. Na jeho dôkaz využijeme vetu 4.5.10, podľa ktorej stačí ukázať, že dimenzia zovšeobecneného vlastného priestoru $G(A, \varrho)$ prislúchajúceho k ϱ je rovná jednej.

Podľa vety 5.4.14 prislúcha k ϱ aspoň jeden *kladný* vlastný vektor $\mathbf{x}_\varrho = (x_1, \dots, x_n)^T \in \mathbb{R}_{>0}^n$. Treba dokázať, že všetky vlastné vektory matice A prislúchajúce k ϱ sú skalárnymi násobkami vektora $\mathbf{x}_\varrho$ a že k vlastnému číslu ϱ matice A neprislúcha žiaden zovšeobecnený vlastný vektor rádu 2.

Nech teda $\mathbf{v} = (v_1, \dots, v_n)^T$ je ľubovoľný vlastný vektor prislúchajúci k ϱ ; z lemy 5.4.12 potom

$$A|\mathbf{v}| \geq \varrho|\mathbf{v}|. \quad (5.13)$$

Evidentne však $|\mathbf{v}| \in \mathbb{R}_{\geq 0}^n \setminus \{\mathbf{0}\}$ a z (5.13) tak podľa tvrdenia 5.4.7 dostávame

$$\Gamma_A(|\mathbf{v}|) \geq \varrho.$$

Z vety 5.4.14 potom vyplýva, že táto nerovnosť je v skutočnosti rovnosťou a $|\mathbf{v}|$ je kladný vlastný vektor prislúchajúci k ϱ . Špeciálne teda $|v_1| > 0$, z čoho $v_1 \neq 0$. Prvá zložka vektora

$$v_1\mathbf{x}_\varrho - x_1\mathbf{v} \quad (5.14)$$

je navyše nulová, pretože je rovná $v_1x_1 - x_1v_1$. Vektor (5.14) preto nemôže byť vlastným vektorom prislúchajúcim k ϱ , lebo v opačnom prípade by podľa argumentácie vyššie bol vektor

$$|v_1\mathbf{x}_\varrho - x_1\mathbf{v}|$$

kladným vlastným vektorom prislúchajúcim k ϱ , hoci jeho prvá zložka je tiež evidentne nulová. Súčasne však vektor (5.14) patrí do $E(A, \varrho)$, v dôsledku čoho zisťujeme, že musí platiť

$$v_1 \mathbf{x}_\varrho - x_1 \mathbf{v} = \mathbf{0},$$

a teda aj

$$\mathbf{v} = \frac{v_1}{x_1} \mathbf{x}_\varrho.$$

Vektor $\mathbf{v}$ je teda skutočne skalárnym násobkom vektora $\mathbf{x}_\varrho$ a prvá časť dôkazu je dokončená.¹⁰

Dokážeme teraz, že k vlastnému číslu ϱ matice A neprislúcha žiaden zovšeobecnený vlastný vektor rádu 2; využijeme pritom menší trik s transponovanou maticou, ktorý je prebratý z [76]. Sporom: nech $\mathbf{v}$ je zovšeobecnený vlastný vektor matice A rádu 2 prislúchajúci k ϱ . Potom

$$\mathbf{v}' = (A - \varrho \mathbf{I})\mathbf{v}$$

je vlastný vektor matice A prislúchajúci k vlastnému číslu ϱ . Z dokázaného teda vyplýva, že existuje konštanta $q \in \mathbb{C} \setminus \{0\}$ taká, že $\mathbf{v}' = q\mathbf{x}_\varrho$.¹¹ Pre zovšeobecnený vlastný vektor $\mathbf{w} := \frac{1}{q}\mathbf{v}$ preto dostávame

$$(A - \varrho \mathbf{I})\mathbf{w} = \frac{1}{q}\mathbf{v}' = \mathbf{x}_\varrho; \quad (5.15)$$

pripomeňme si, že vektor $\mathbf{x}_\varrho$ je *kladný*. Ireducibilnou nezápornou maticou je však aj matica A^T , ktorej spektrálny polomer je vďaka vete 4.8.1 tiež ϱ – podľa vety 5.4.14 tak existuje *kladný* vektor $\mathbf{y}_\varrho$ taký, že

$$(A^T - \varrho \mathbf{I})\mathbf{y}_\varrho = \mathbf{0},$$

čo možno ekvivalentne prepísať ako

$$\mathbf{y}_\varrho^T (A - \varrho \mathbf{I}) = \mathbf{0}^T. \quad (5.16)$$

Z rovností (5.15) a (5.16) napokon dostávame

$$\mathbf{y}_\varrho^T \mathbf{x}_\varrho = \mathbf{y}_\varrho^T (A - \varrho \mathbf{I})\mathbf{w} = \mathbf{0}^T \mathbf{w} = 0,$$

čo odporuje skutočnosti, že obidva vektory $\mathbf{x}_\varrho$ a $\mathbf{y}_\varrho$ sú kladné: spor. □

Spektrálny polomer $\varrho(A)$ je teda pre každú ireducibilnú nezápornú maticu A jej vlastným číslom s algebraickou násobnosťou 1. V nasledujúcom postupne charakterizujeme aj zvyšné vlastné čísla λ matice A také, že $|\lambda| = \varrho(A)$ – získame tak úplnú predstavu o najvýznamnejších vlastných číslach ireducibilných matíc. Prvým krokom k analýze týchto vlastných čísel bude *Wielandtova veta*, ktorú teraz vyslovíme a dokážeme.

Nech n je nenulové prirodzené číslo, $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ a $B = (b_{i,j})_{n \times n} \in \mathbb{R}_{\geq 0}^{n \times n}$. Budeme hovoriť, že nezáporná matica B *dominuje* matici A a písať

$$|A| \leq B,$$

ak pre $i, j = 1, \dots, n$ je $|a_{i,j}| \leq b_{i,j}$. Wielandtova veta hovorí o vzťahu medzi spektrálnym polomerom komplexnej matice a spektrálnym polomerom *ireducibilnej* nezápornej matice, ktorá jej dominuje.

¹⁰Okrem iného sme v tejto prvej časti dôkazu ukázali, že *geometrická* násobnosť vlastného čísla ϱ je rovná jednej.

¹¹Nenulovosť konštanty q je dôsledkom toho, že $\mathbf{v}'$ je vlastný vektor – musí teda byť $\mathbf{v}' \neq \mathbf{0}$.

Veta 5.4.16 (Wielandt). *Nech $n \in \mathbb{N} \setminus \{0\}$, $A = (a_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ je matrica a $B = (b_{i,j})_{n \times n} \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matrica taká, že $|A| \leq B$. Pre každé $\lambda \in \sigma(A)$ je potom*

$$|\lambda| \leq \varrho(B).$$

Vlastné číslo matice A s absolútnou hodnotou $\varrho(B)$ navyše existuje práve vtedy, keď existuje $\omega \in \mathbb{C}$ také, že $|\omega| = 1$ a

$$A = \omega D B D^{-1} \quad \text{resp.} \quad D^{-1} A D = \omega B$$

pre nejakú diagonálnu maticu $D \in \mathbb{C}^{n \times n}$ s absolútnou hodnotou $|D| = \mathbf{I}$. V takom prípade množina takýchto vlastných čísel obsahuje práve všetky $\lambda = \omega \varrho(B)$, kde ω je číslo s uvedenou vlastnosťou.

Dôkaz. Nech λ je ľubovoľné vlastné číslo matice A a nech $\mathbf{v}$ je vlastný vektor matice prislúchajúci k λ . Z predpokladu $|A| \leq B$ a lemy 5.4.12 potom

$$B|\mathbf{v}| \geq |A| \cdot |\mathbf{v}| \geq |\lambda| \cdot |\mathbf{v}|, \quad (5.17)$$

z čoho podľa tvrdenia 5.4.7 a vety 5.4.14 dostávame

$$|\lambda| \leq \Gamma_B(|\mathbf{v}|) \leq \varrho(B), \quad (5.18)$$

čo dokazuje prvú časť vety.

Za účelom dôkazu druhej časti vety najprv predpokladajme, že $D^{-1}AD = \omega B$ pre nejaké $\omega \in \mathbb{C}$ s $|\omega| = 1$ a diagonálnu maticu $D \in \mathbb{C}^{n \times n}$ s $|D| = \mathbf{I}$. Matice A a ωB si sú v takom prípade podobné, pričom $\varrho(B)$ je podľa vety 5.4.14 vlastným číslom matice B . Podľa vety 4.8.2 tak $\lambda = \omega \varrho(B)$ musí byť vlastným číslom matice ωB , a teda aj s ňou podobnej matice A . Zjavne pritom $|\lambda| = \varrho(B)$.

Pre dôkaz opačnej implikácie predpokladajme, že λ je vlastné číslo matice A také, že $|\lambda| = \varrho(B)$. Ak $\varrho(B) = 0$, musí byť matica B evidentne nilpotentná¹² a z jej ireducibility tak nutne $B = (0)_{1 \times 1}$. Preto aj $A = (0)_{1 \times 1}$ a zvyšok dokazovaného tvrdenia je triviálne pravdivý.

Nech teda $\varrho(B) = |\lambda| > 0$ a $\mathbf{v}$ je vlastný vektor matice A prislúchajúci k λ . Podľa (5.18) potom

$$|\lambda| = \Gamma_B(|\mathbf{v}|) = \varrho(B)$$

a z vety 5.4.14 vyplýva, že $|\mathbf{v}|$ je *kladným* vlastným vektorom matice B prislúchajúcim k vlastnému číslu $\varrho(B)$. Z (5.17) ďalej dostávame

$$B|\mathbf{v}| = |A| \cdot |\mathbf{v}| = \varrho(B)|\mathbf{v}|,$$

z čoho

$$(B - |A|)|\mathbf{v}| = \mathbf{0}. \quad (5.19)$$

Keďže ale $|A| \leq B$, je matica $B - |A|$ nezáporná, v dôsledku čoho (5.19) implikuje

$$B = |A|. \quad (5.20)$$

Nech teraz $\omega := \frac{\lambda}{\varrho(B)}$ a pre $\mathbf{v} = (v_1, \dots, v_n)^T$ položme

$$D := \begin{pmatrix} \frac{v_1}{|v_1|} & 0 & \dots & 0 \\ 0 & \frac{v_2}{|v_2|} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{v_n}{|v_n|} \end{pmatrix}.$$

¹²Ide o náplň cvičenia 8 kapitoly 4. Ľahko to ale vidieť napríklad aj z viet 5.2.1 a 1.3.10.

Ľahko vidieť, že $|\omega| = 1$, $|D| = \mathbf{I}$ a $\lambda = \omega \varrho(B)$. Stačí teda dokázať, že

$$D^{-1}AD = \omega B.$$

Zrejme navyše

$$D|\mathbf{v}| = \mathbf{v}. \quad (5.21)$$

Keďže je $\mathbf{v}$ vlastným vektorom matice A prislúchajúcim k λ , platí $A\mathbf{v} = \lambda\mathbf{v}$, z čoho s použitím rovnosti (5.21) dostávame

$$AD|\mathbf{v}| = \lambda D|\mathbf{v}|,$$

a teda aj

$$D^{-1}AD|\mathbf{v}| = \lambda|\mathbf{v}| = \omega \varrho(B)|\mathbf{v}|.$$

Avšak $|\mathbf{v}|$ je vlastným vektorom matice B prislúchajúcim k $\varrho(B)$ – uvedenú rovnosť tak možno prepísať ako

$$D^{-1}AD|\mathbf{v}| = \omega B|\mathbf{v}|,$$

z čoho

$$(\omega^{-1}D^{-1}AD - B)|\mathbf{v}| = \mathbf{0}. \quad (5.22)$$

Z rovnosti (5.20) ale evidentne

$$B = |A| = |\omega^{-1}D^{-1}AD|.$$

Ak teda pre každú maticu $Z = (z_{i,j})_{n \times n} \in \mathbb{C}^{n \times n}$ zavedieme označenia

$$\operatorname{Re} Z = (\operatorname{Re} z_{i,j})_{n \times n} \quad \text{a} \quad \operatorname{Im} Z = (\operatorname{Im} z_{i,j})_{n \times n},$$

z trojuholníkovej nerovnosti nutne

$$\operatorname{Re}(\omega^{-1}D^{-1}AD) \leq B$$

a keďže je vektor $|\mathbf{v}|$ kladný, z rovnosti (5.22) dostávame

$$\operatorname{Re}(\omega^{-1}D^{-1}AD) = B.$$

Zo vzťahu $B = |\omega^{-1}D^{-1}AD|$ je potom zrejmé, že $\operatorname{Im}(\omega^{-1}D^{-1}AD) = \mathbf{0}$, z čoho $\omega^{-1}D^{-1}AD = B$, a teda naozaj

$$D^{-1}AD = \omega B.$$

Tým je dôkaz vety dokončený. $\square$

Wielandtovu vetu teraz využijeme pri dôkaze charakterizácie najvýznamnejších vlastných čísel ireducibilných nezáporných matíc – to jest tých vlastných čísel, ktorých absolútna hodnota je rovná spektrálnemu polomeru uvažovanej matice. Dokážeme, že pre každú ireducibilnú nezápornú maticu A existuje prirodzené číslo $p \geq 1$, pre ktoré sú tieto vlastné čísla dané ako $\omega_p^0 \varrho, \omega_p^1 \varrho, \dots, \omega_p^{p-1} \varrho$, kde ϱ je spektrálny polomer matice A a ω_p je p -ta komplexná odmocnina jednej s minimálnym nenulovým argumentom – $\omega_p^0, \dots, \omega_p^{p-1}$ sú tak práve všetky p -te komplexné odmocniny čísla 1. Ukážeme navyše, že algebraická násobnosť každého z týchto vlastných čísel je rovná jednej.

Veta 5.4.17. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica o spektrálnom polomere $\varrho = \varrho(A)$. Potom existuje nenulové prirodzené číslo p také, že matica A má práve p vlastných čísel s absolútnou hodnotou ϱ , ktoré sú dané ako*

$$e^{0(2\pi i/p)} \varrho, e^{1(2\pi i/p)} \varrho, \dots, e^{(p-1)(2\pi i/p)} \varrho,$$

pričom algebraická násobnosť každého z nich je 1. Matica A je navyše pre $k = 0, \dots, p-1$ podobná matici $e^{k(2\pi i/p)} A$; matice $e^{k(2\pi i/p)} A$ pre $k = 0, \dots, p-1$ teda majú rovnaké spektrum.

Dôkaz. Pre $\varrho = 0$ je tvrdenie triviálne, pretože v takom prípade nutne $A = (0)_{1 \times 1}$. Jediným vlastným číslom takejto matice je 0 – stačí tak položiť $p = 1$.

Nech teda $\varrho > 0$ a p je – nutne nenulový – počet (po dvoch rôznych) vlastných čísel matice A s absolútnou hodnotu ϱ . Predpokladajme, že sú tieto vlastné čísla dané ako

$$\lambda_0 = e^{\theta_0 i} \varrho, \quad \dots, \quad \lambda_{p-1} = e^{\theta_{p-1} i} \varrho,$$

kde $\theta_0, \dots, \theta_{p-1}$ sú po dvoch rôzne reálne čísla z intervalu $[0, 2\pi)$; z vety 5.4.14 vyplýva, že jedno z nich je určite rovné 0. Bez ujmy na všeobecnosti predpokladajme, že

$$0 = \theta_0 < \dots < \theta_{p-1}.$$

Dokážeme, že pre $k = 0, \dots, p-1$ je $\theta_k = k(2\pi/p)$.

Iste ale $|A| \leq A$ a $\lambda_k = e^{\theta_k i} \varrho$ je pre $k = 0, \dots, p-1$ vlastným číslom matice A takým, že $|\lambda_k| = \varrho$. Podľa vety 5.4.16 teda zisťujeme, že pre $k = 0, \dots, p-1$ existuje diagonálna matica D_k taká, že $|D_k| = \mathbf{I}$ a

$$A = e^{\theta_k i} D_k A D_k^{-1}.$$

Pre ľubovoľnú dvojicu $r, s \in \{0, \dots, p-1\}$ navyše

$$A = e^{\theta_r i} D_r A D_r^{-1} = e^{\theta_r i} D_r e^{\theta_s i} D_s A D_s^{-1} D_r^{-1} = e^{(\theta_s + \theta_r) i} (D_r D_s) A (D_r D_s)^{-1};$$

z vety 5.4.16 teda vyplýva, že $e^{(\theta_s + \theta_r) i} \varrho$ je tiež vlastným číslom matice A .

Množina

$$M = \{\theta_0, \theta_1, \dots, \theta_{p-1}\} \subseteq [0, 2\pi)$$

je teda p -prvková, uzavretá na súčet modulo 2π a zároveň obsahujúca číslo 0. Ľahko teda vidieť, že musí byť

$$M = \{0(2\pi/p), 1(2\pi/p), \dots, (p-1)(2\pi/p)\}$$

a pre $k = 0, \dots, p-1$ naozaj $\theta_k = k(2\pi/p)$.

Z vety 5.4.16 ďalej vyplýva, že pre $k = 0, \dots, p-1$ je

$$e^{k(2\pi i/p)} A = D_k^{-1} A D_k,$$

kde D_k je diagonálna matica taká, že $|D_k| = \mathbf{I}$. Matice A a $e^{k(2\pi i/p)} A$ sú teda podobné. V dôsledku toho majú rovnaké spektrum a algebraické násobnosti jednotlivých vlastných čísel. Algebraická násobnosť vlastného čísla $e^{k(2\pi i/p)} \varrho$ matice A je však pre $k = 1, \dots, p-1$ rovná, podľa vety 4.8.2, algebraickej násobnosti vlastného čísla ϱ matice $e^{(p-k)(2\pi i/p)} A$, a tá musí byť rovnako ako pri matici A rovná jednej. Zisťujeme teda, že vlastné čísla

$$e^{0(2\pi i/p)} \varrho, e^{1(2\pi i/p)} \varrho, \dots, e^{(p-1)(2\pi i/p)} \varrho$$

sú všetky algebraickej násobnosti 1 a veta je dokázaná. □

Nasledujúca veta hovorí, že každá ireducibilná nezáporná matica A má (až na skalárne násobky) *jediný* nezáporný vlastný vektor – ten prislúcha k vlastnému číslu $\varrho(A)$ a je kladný.

Veta 5.4.18. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica o spektrálnom polomere $\varrho = \varrho(A)$. K vlastnému číslu ϱ potom prislúcha kladný vlastný vektor $\mathbf{x}_\varrho \in \mathbb{R}_{>0}^n$, pričom každý ďalší nezáporný vlastný vektor $\mathbf{x} \in \mathbb{R}_{\geq 0}^n$ matice A je skalárnym násobkom $\mathbf{x}_\varrho$ a prislúcha (iba) k ϱ .*

Dôkaz. Existencia kladného vlastného vektora $\mathbf{x}_\varrho$ matice A prislúchajúceho k ϱ vyplýva z vety 5.4.14. Transponovaná matica A^T má navyše rovnaké spektrum ako matica A a s použitím vety 5.4.14 tak zisťujeme, že k vlastnému číslu ϱ matice A^T prislúcha nejaký kladný vlastný vektor $\mathbf{y}_\varrho$. Uvažujme teraz ľubovoľný nezáporný vlastný vektor $\mathbf{x}$ matice A prislúchajúci k nejakému vlastnému číslu λ ; ten je podľa lemy 5.4.11 istotne kladný. Potom

$$\lambda \mathbf{x}^T \mathbf{y}_\varrho = (A\mathbf{x})^T \mathbf{y}_\varrho = \mathbf{x}^T A^T \mathbf{y}_\varrho = \varrho \mathbf{x}^T \mathbf{y}_\varrho. \quad (5.23)$$

Keďže sú vektory $\mathbf{y}_\varrho$ a $\mathbf{x}$ kladné, nutne $\mathbf{x}^T \mathbf{y}_\varrho > 0$ a z rovnosti (5.23) dostávame $\lambda = \varrho$. Ľubovoľný nezáporný vlastný vektor $\mathbf{x}$ tak skutočne musí prislúchať k vlastnému číslu ϱ , pričom jeho jedinečnosť (až na skalárny násobok) napokon vyplýva z vety 5.4.15. $\square$

Môžeme tak zhrnúť naše doterajšie zistenia do tvrdenia, ktorého varianty bývajú v literatúre občas nazývané *Perronovou-Frobeniovou vetou*.

Veta 5.4.19 (Perron, Frobenius). *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica. Potom:*

- (i) *Spektrálny polomer $\varrho := \varrho(A)$ je vlastným číslom matice A .*
- (ii) *Algebraická násobnosť vlastného čísla ϱ matice A je rovná jednej.*
- (iii) *Existuje prirodzené číslo $p \geq 1$ také, že matica A má práve p vlastných čísel absolútnej hodnoty ϱ , ktoré sú dané ako*

$$e^{0(2\pi i/p)} \varrho, e^{1(2\pi i/p)} \varrho, \dots, e^{(p-1)(2\pi i/p)} \varrho$$

- (iv) *Algebraická násobnosť všetkých vlastných čísel λ matice A spĺňajúcich $|\lambda| = \varrho$ je rovná jednej.*
- (v) *Ak p je číslo z tvrdenia (iii), sú si všetky matice*

$$A, e^{1(2\pi i/p)} A, \dots, e^{(p-1)(2\pi i/p)} A$$

navzájom podobné.

- (vi) *K vlastnému číslu ϱ matice A prislúcha kladný vlastný vektor $\mathbf{x}_\varrho$.*
- (vii) *Každý nezáporný vlastný vektor matice A je skalárnym násobkom $\mathbf{x}_\varrho$ a prislúcha tak k ϱ .*

Dôkaz. Vyplýva bezprostredne z viet 5.4.14, 5.4.15, 5.4.17 a 5.4.18. $\square$

Perronovu-Frobeniovu vetu možno využiť aj pri skúmaní spektrálnych vlastností *reducibilných* nezáporných matíc. Základom tu je takzvaný *normálny tvar reducibilnej matice*.

Definícia 5.4.20. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$. Matica A je v normálnom tvare reducibilnej matice, ak pre nejaké nenulové prirodzené číslo m nadobúda blokový rozklad*

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,m} \\ \mathbf{0} & A_{2,2} & \dots & A_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & A_{m,m} \end{pmatrix},$$

kde pre $i = 1, \dots, m$ existuje nenulové prirodzené číslo n_i také, že $A_{i,i} \in \mathbb{R}_{\geq 0}^{n_i \times n_i}$ je ireducibilná matica.

Táto terminológia môže byť trochu zavádzajúca, pretože špeciálne každá ireducibilná matica je v normálnom tvare reducibilnej matice. O *normálnom tvare* môžeme hovoriť vďaka nasledujúcej vete, ktorej dôkaz je jednoduchým cvičením na použitie matematickej indukcie spolu s definíciou 5.4.1.

Veta 5.4.21. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$. Potom existuje permutačná matica P typu $n \times n$ taká, že matica*

$$PAP^{-1}$$

je v normálnom tvare reducibilnej matice.

Dôkaz. Prenechaný čitateľovi ako cvičenie 5 na konci kapitoly. □

Z uvedenej vety okrem iného vyplýva, že spektrum každej nezápornej matice A možno vyjadriť ako zjednotenie spektier niekoľkých ireducibilných matíc – presnejšie diagonálnych blokov jej vyjadrenia v normálnom tvare reducibilnej matice; algebraická násobnosť každého vlastného čísla λ matice A je pritom daná súčtom algebraických násobností tohto vlastného čísla cez všetky diagonálne bloky, spektra ktorých je λ prvkom.

5.5 Primitívne matice a aperiodické grafy

Podľa Perronovej-Frobeniovej vety sú pre každú *ireducibilnú* nezápornú maticu A jej najvýznamnejšie vlastné čísla – teda vlastné čísla s absolútnou hodnotou rovnou spektrálnemu polomeru ϱ tejto matice – dané ako

$$e^{0(2\pi i/p)}\varrho, e^{1(2\pi i/p)}\varrho, \dots, e^{(p-1)(2\pi i/p)}\varrho,$$

kde $p \geq 1$ je prirodzené číslo; všetky tieto vlastné čísla sú navyše algebraickej násobnosti 1. Prirodzene vyvstáva otázka, ako súvisí počet týchto vlastných čísel – čiže číslo p – so štrukturálnymi vlastnosťami matice A a zodpovedajúceho ohodnoteného grafu.

V nasledujúcom sa tak budeme zaoberať počtom najvýznamnejších vlastných čísel ireducibilnej nezápornej matice, ktorý nazveme jej *indexom primitívnosti*. Ireducibilné matice s jediným najvýznamnejším vlastným číslom ϱ pritom nazveme *primitívnymi* – úlohy, v ktorých vystupuje takáto matica, sú často riešiteľné jednoduchšie, než v prípade všeobecných ireducibilných matíc. Index primitívnosti ireducibilnej matice navyše dáme do súvisu s periodickým správaním jej mocnín, čo vyjadríme v reči príslušných silne súvislých grafov charakterizáciou prostredníctvom pojmu *periódy* grafu, ktorý bude priamym „grafovým náprotivkom“ indexu primitívnosti.¹³

Definícia 5.5.1. Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica s $\varrho(A) = \varrho > 0$.¹⁴ *Indexom primitívnosti* matice A potom nazveme najväčšie prirodzené číslo $p(A) \geq 1$ také, že

$$e^{0(2\pi i/p(A))}\varrho, e^{1(2\pi i/p(A))}\varrho, \dots, e^{(p(A)-1)(2\pi i/p(A))}\varrho$$

sú vlastné čísla matice A . Ireducibilnú maticu A nazveme *primitívnou*, ak $p(A) = 1$.

Nasledujúca veta charakterizuje index primitívnosti ireducibilnej matice pomocou vlastností jej charakteristického polynómu. Toto prepojenie neskôr využijeme pri skúmaní súvislostí medzi indexom primitívnosti matice a vlastnosťami príslušného ohodnoteného grafu.

¹³Niektorí autori tak priamo hovoria aj o *perióde matice*. Podobne pojem *indexu primitívnosti* (v angl. *imprimitivity index*, čiže index „neprimitívnosti“) sa občas používa aj pre silne súvislé grafy.

¹⁴Predpokladom $\varrho > 0$ iba vylúčime triviálny prípad nulovej matice typu 1×1 . V prípade potreby by pre túto maticu bolo možné index primitívnosti konvenčne dodefinovať napríklad ako ∞ alebo 0.

Veta 5.5.2. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica s $\varrho(A) = \varrho > 0$. Nech*

$$\text{ch}_A(z) = z^n + a_1 z^{n_1} + a_2 z^{n_2} + \dots + a_k z^{n_k},$$

kde $k \in \mathbb{N}$, $a_1, \dots, a_k \in \mathbb{C} \setminus \{0\}$ a $n_1, \dots, n_k \in \mathbb{N}$ sú také, že $n > n_1 > n_2 > \dots > n_k \geq 0$.¹⁵ Potom $k \geq 1$ a

$$p(A) = \gcd\{n - n_1, n - n_2, \dots, n - n_k\} = \gcd\{n - n_1, n_1 - n_2, \dots, n_{k-1} - n_k\},$$

kde $\gcd$ označuje najväčší spoločný deliteľ čísel danej množiny.

Dôkaz. Keďže má matica A aspoň jedno nenulové vlastné číslo, nutne $k \geq 1$. Nech teraz $p \geq 1$ je prirodzené číslo také, že $e^{2\pi i/p} \varrho$ je vlastným číslom uvažovanej matice A . Označme $\omega := e^{2\pi i/p}$. Podľa Perronovej-Frobeniovej vety sú matice A a ωA podobné a na základe tvrdenia 4.1.14 tak majú rovnaký charakteristický polynóm. Charakteristický polynóm matice ωA ale možno podľa vety 4.8.2 vyjadriť ako

$$\text{ch}_{\omega A}(z) = \omega^n \text{ch}_A(z/\omega) = z^n + \omega^{n-n_1} a_1 z^{n_1} + \omega^{n-n_2} a_2 z^{n_2} + \dots + \omega^{n-n_k} a_k z^{n_k}$$

a rovnosť charakteristických polynómov matíc ωA a A teda možno zapísať ako

$$z^n + \omega^{n-n_1} a_1 z^{n_1} + \omega^{n-n_2} a_2 z^{n_2} + \dots + \omega^{n-n_k} a_k z^{n_k} = z^n + a_1 z^{n_1} + a_2 z^{n_2} + \dots + a_k z^{n_k}.$$

V dôsledku toho

$$\omega^{n-n_1} = \omega^{n-n_2} = \dots = \omega^{n-n_k} = 1$$

a keďže $\omega = e^{2\pi i/p}$, musí byť číslo p deliteľom každého z čísel $n - n_1, n - n_2, \dots, n - n_k$.

Ak je naopak p spoločným deliteľom čísel $n - n_1, n - n_2, \dots, n - n_k$, pre $\omega = e^{2\pi i/p}$ dostávame

$$\omega^{n-n_1} = \omega^{n-n_2} = \dots = \omega^{n-n_k} = 1,$$

z čoho vyplýva, že matice A a ωA majú rovnaký charakteristický polynóm, a teda aj rovnaké spektrum – $\omega \varrho$ je tak vlastným číslom matice A .

Index primitívnosti $p(A)$ je teraz *najväčším* číslom $p \in \mathbb{N} \setminus \{0\}$ takým, že $e^{2\pi i/p} \varrho$ je vlastným číslom matice A . Z dokázaného vyplýva, že $p(A)$ musí byť deliteľom čísel $n - n_1, n - n_2, \dots, n - n_k$. Keby ďalej bolo spoločným deliteľom týchto čísel aj nejaké prirodzené číslo $r > p(A)$, vyplývalo by z dokázaného, že $e^{2\pi i/r} \varrho$ je vlastným číslom matice A , čo by evidentne odporovalo definícii indexu primitívnosti $p(A)$ ako najväčšieho takého čísla. Index primitívnosti $p(A)$ je tak skutočne *najväčším* spoločným deliteľom čísel $n - n_1, n - n_2, \dots, n - n_k$, pričom ľahko vidieť, že v takom prípade musí byť aj najväčším spoločným deliteľom čísel $n - n_1, n_1 - n_2, \dots, n_{k-1} - n_k$. $\square$

Index primitívnosti ireducibilnej nezápornej matice teraz charakterizujeme prostredníctvom vlastností príslušného ohodnoteného grafu nad $\mathbb{R}_{\geq 0}$. Nech $\mathcal{G} = (V, \varphi)$ je takýto ohodnotený graf. *Uzavretým sledom* v grafe $\mathcal{G}$ nazveme ľubovoľný sled

$$\gamma = (u_0, u_1, u_2, \dots, u_t)$$

v grafe $\mathcal{G}$ taký, že $u_t = u_0$; dĺžka $|\gamma| = t$ uzavretého sledu γ môže byť aj nulová. *Orientovanou kružnicou*¹⁶ v grafe $\mathcal{G}$ nazveme uzavretý sled

$$\gamma = (u_0, u_1, u_2, \dots, u_t)$$

v grafe $\mathcal{G}$ taký, že $t \geq 1$ a vrcholy $u_0, u_1, \dots, u_{t-1}$ sú po dvoch rôzne. Definujeme teraz parameter silne súvislých grafov, ktorý je obdobou indexu primitívnosti ireducibilných matíc.

¹⁵Podľa tvrdenia 4.2.7 možno v tomto tvare vyjadriť charakteristický polynóm ľubovoľnej matice typu $n \times n$.

¹⁶Orientované kružnice možno opísať – podobne ako kružnice v neorientovaných grafoch – ako uzavreté sledy nenulovej dĺžky, v ktorých sa až na prvý a zároveň posledný vrchol sledu neopakujú vrcholy ani hrany. Na rozdiel od kružníc v neorientovaných grafoch ale napríklad môžu existovať aj orientované kružnice dĺžky 2. Orientované kružnice o jednej hrane nazývame *slučkami*.

Definícia 5.5.3. Nech $\mathcal{G} = (V, \varphi)$ je silne súvislý ohodnotený graf nad $\mathbb{R}_{\geq 0}$ obsahujúci aspoň jednu hranu.¹⁷ *Periódou* grafu $\mathcal{G}$ nazveme najväčší spoločný deliteľ $p(\mathcal{G})$ dĺžok orientovaných kružníc v $\mathcal{G}$:

$$p(\mathcal{G}) = \gcd\{|\gamma| \mid \gamma \text{ je orientovaná kružnica v } \mathcal{G}\}.$$

Graf $\mathcal{G}$ nazveme *aperiodický*, ak $p(\mathcal{G}) = 1$.

Dokážeme teraz vetu, podľa ktorej sú pojmy periódy silne súvislého grafu a indexu primitívnosti ireducibilnej matice len dvoma odlišnými pohľadmi na ten istý parameter – indexom primitívnosti matice susednosti silne súvislého grafu $\mathcal{G}$ je perióda grafu $\mathcal{G}$ a periódou grafu ireducibilnej matice A je index primitívnosti matice A .

Veta 5.5.4. Nech $n \in \mathbb{N} \setminus \{0\}$, $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica s $\varrho(A) = \varrho > 0$ a $\mathcal{G} = ([n], \varphi)$ je silne súvislý ohodnotený graf nad $\mathbb{R}_{\geq 0}$ taký, že $A = \mathbf{A}(\mathcal{G})$ resp. $\mathcal{G} = \text{Gr}(A)$. Potom

$$p(A) = p(\mathcal{G}).$$

Špeciálne je matica A primitívna práve vtedy, keď je graf $\mathcal{G}$ aperiodický.

Dôkaz. Ak $A = (a_{i,j})_{n \times n}$, z definície determinantu vyplýva

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \sum_{\pi \in S_n} \left(\text{sgn}(\pi) \prod_{k=1}^n (\delta_{k,\pi(k)}z - a_{k,\pi(k)}) \right), \quad (5.24)$$

kde S_n je symetrická grupa n -prvkovej množiny pozostávajúca z práve všetkých permutácií

$$\pi: [n] \rightarrow [n],$$

zápis $\text{sgn}(\pi)$ označuje znamienko¹⁸ permutácie π a Kroneckerova delta $\delta_{r,s}$ je pre všetky $r, s \in \mathbb{N}$ definovaná ako

$$\delta_{r,s} = \begin{cases} 1 & \text{ak } r = s, \\ 0 & \text{ak } r \neq s. \end{cases}$$

Pre každé $\pi \in S_n$ označme

$$\text{prod}_{\pi}(z) := \prod_{k=1}^n (\delta_{k,\pi(k)}z - a_{k,\pi(k)});$$

rovnosť (5.24) tak možno prepísať ako

$$\text{ch}_A(z) = \sum_{\pi \in S_n} \text{sgn}(\pi) \cdot \text{prod}_{\pi}(z).$$

Uvažujme teraz ľubovoľný rozklad niektorej permutácie $\pi \in S_n$ na disjunktné cykly. Ľahko vidieť, že ak polynóm $\text{prod}_{\pi}(z)$ nie je nulový, tak každý cyklus $(u_1 u_2 \dots u_k)$ v tomto rozklade, okrem pevných bodov, musí zodpovedať orientovanej kružnici

$$(u_1, u_2, \dots, u_k, u_1)$$

v grafe $\mathcal{G}$. Pevný bod môže alebo nemusí zodpovedať slučke v grafe $\mathcal{G}$.

Ukážeme najprv, že *perióda $p(\mathcal{G})$ grafu $\mathcal{G}$ musí byť deliteľná indexom primitívnosti $p(A)$ matice A* . Presnejšie dokážeme indukciou vzhľadom na $r = 1, \dots, n$, že číslo r musí byť deliteľné číslom $p(A)$ kedykoľvek graf $\mathcal{G}$ obsahuje orientovanú kružnicu dĺžky r .

¹⁷Týmto predpokladom iba vylúčime, podobne ako pri definícii indexu primitívnosti ireducibilnej matice, triviálny prípad silne súvislého grafu pozostávajúceho z jediného vrcholu bez slučky; tomuto grafu zodpovedá matica $A = (0)$, ktorú sme vylúčili v definícii 5.5.1. Podobne ako pri maticiach by však prostredníctvom vhodnej konvencie bolo v prípade potreby možné rozšíriť definíciu aj na tento triviálny prípad.

¹⁸Číže $\text{sgn}(\pi) = 1$, ak je permutácia π párna a $\text{sgn}(\pi) = -1$, ak je táto permutácia nepárna.

Ak graf $\mathcal{G}$ obsahuje orientovanú kružnicu dĺžky $r = 1 - t$. j. slučku – existuje aspoň jedno $k \in [n]$ také, že $a_{k,k} > 0$. Ľahko potom vidieť, že $\text{prod}_\pi(z)$ je polynóm s nenulovým koeficientom pri z^{n-1} práve vtedy, keď je π identická permutácia. V dôsledku toho je aj $\text{ch}_A(z)$ polynómom s nenulovým koeficientom pri z^{n-1} a z vety 5.5.2 vyplýva $p(A) = 1$; číslo $r = 1$ tak je deliteľné číslom $p(A)$.

Nech teraz tvrdenie platí pre $r = 1, \dots, q$, kde $1 \leq q < n$ a uvažujme $r = q + 1$. Nech teda graf $\mathcal{G}$ obsahuje aspoň jednu orientovanú kružnicu

$$\gamma = (u_1, u_2, \dots, u_{q+1}, u_1)$$

dĺžky $q + 1$. Pre permutáciu π_γ s rozkladom na cyklus $(u_1 u_2 \dots u_{q+1})$ a $n - (q + 1)$ pevných bodov je potom $\text{prod}_{\pi_\gamma}(z)$ očividne polynómom s nenulovým koeficientom pri $z^{n-(q+1)}$. Znamienko $\text{sgn}(\pi)$ každej permutácie π takto rozložiteľnej na cyklus dĺžky $q + 1$ a $n - (q + 1)$ pevných bodov je pritom dané ako $(-1)^q$. Ak sú teda jedinými permutáciami π s nenulovým koeficientom pri $z^{n-(q+1)}$ v $\text{prod}_\pi(z)$ permutácie tohto typu, musí byť nenulový aj koeficient pri $z^{n-(q+1)}$ v polynóme $\text{ch}_A(z)$. Z vety 5.5.2 potom vyplýva, že $p(A)$ je deliteľom čísla $r = q + 1$, čo bolo treba dokázať.

Ak naopak existujú aj iné permutácie π také, že koeficient pri $z^{n-(q+1)}$ v $\text{prod}_\pi(z)$ je nenulový, musí byť každá z takýchto permutácií rozložiteľná na $n - (q + 1)$ pevných bodov a $s \geq 2$ cyklov dĺžok $1 \leq q_1, \dots, q_s \leq q$ – niektoré z týchto cyklov teda môžu byť tiež pevnými bodmi – takých, že $q_1 + \dots + q_s = q + 1$ a graf $\mathcal{G}$ obsahuje orientovanú kružnicu každej z dĺžok $q_1, \dots, q_s$. Z indukčného predpokladu potom vyplýva, že čísla $q_1, \dots, q_s$ sú deliteľné číslom $p(A)$, v dôsledku čoho je číslom $p(A)$ deliteľný aj ich súčet $q + 1$. Indukčný krok je tým pádom dokázaný.

Zostáva teda dokázať, že *index primitívnosti* $p(A)$ matice A je deliteľný periódou $p(\mathcal{G})$ grafu $\mathcal{G}$. Tu potrebujeme prísť k pozorovaniu, že $r \in [n]$ musí byť deliteľné číslom $p(\mathcal{G})$ kedykoľvek je koeficient pri z^{n-r} v $\text{ch}_A(z)$ nenulový. Podobne ako vyššie ale zisťujeme, že v prípade nenulovosti tohto koeficientu musí existovať permutácia $\pi \in S_n$ rozložiteľná na $n - r$ pevných bodov a $s \geq 1$ cyklov o dĺžkach $1 \leq r_1, \dots, r_s \leq r$ takých, že $r_1 + \dots + r_s = r$ a graf $\mathcal{G}$ obsahuje orientovanú kružnicu každej z dĺžok $r_1, \dots, r_s$. Perióda $p(\mathcal{G})$ je tak spoločným deliteľom čísel $r_1, \dots, r_s$, v dôsledku čoho musí deliť aj ich súčet – čiže číslo r . $\square$

Primitívne matice, zodpovedajúce podľa práve dokázanej vety *aperiodickým grafom*, tvoria triedu ireducibilných matíc s obzvlášť jednoduchou – až triviálnou – štruktúrou najvýznamnejšej časti spektra. Z Perronovej-Frobeniovej vety totiž vyplýva, že spektrálny polomer ϱ primitívnej matice A je vždy jej vlastným číslom a pre každé ďalšie vlastné číslo λ matice A je $|\lambda| < \varrho$. Nasledujúce dva dôsledky vety 5.5.4 hovoria o postačujúcich podmienkach primitívnosti matice.

Dôsledok 5.5.5. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A = (a_{i,j})_{n \times n} \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica s $\varrho(A) > 0$ taká, že pre nejaké $k \in [n]$ je $a_{k,k} > 0$.¹⁹ Potom je matica A primitívna.*

Dôkaz. Graf $\text{Gr}(A)$ každej takejto matice A obsahuje slučku – čiže orientovanú kružnicu dĺžky 1. Primitívnosť matice A je tak bezprostredným dôsledkom vety 5.5.4. $\square$

Dôsledok 5.5.6. *Nech $n \in \mathbb{N} \setminus \{0\}$. Každá kladná matica $A \in \mathbb{R}_{>0}^{n \times n}$ je primitívna.*

Dôkaz. Vyplýva bezprostredne z dôsledku 5.5.5. $\square$

Po zvyšok tohto oddielu budeme dokazovať nasledujúcu ekvivalentnú charakterizáciu primitívnych matíc: ireducibilná matica A s nenulovým spektrálnym polomerom je primitívna práve vtedy, keď je niektorá jej mocnina kladnou maticou. Prvým krokom k tomuto výsledku bude nasledujúca lema poskytujúca dve alternatívne ekvivalentné vyjadrenia periódy grafu.

Hovoríme, že sled

$$\gamma = (u_0, u_1, u_2, \dots, u_t)$$

v grafe $\mathcal{G} = (V, \varphi)$ prechádza cez vrchol u , ak existuje index $k \in \{0, \dots, t\}$ taký, že $u_k = u$.

¹⁹Matica A teda obsahuje na svojej diagonále aspoň jeden nenulový – to v tomto prípade znamená kladný – prvok. Ekvivalentne možno túto podmienku sformulovať tak, že stopa $\text{tr}(A) = a_{1,1} + \dots + a_{n,n}$ matice A je kladná.

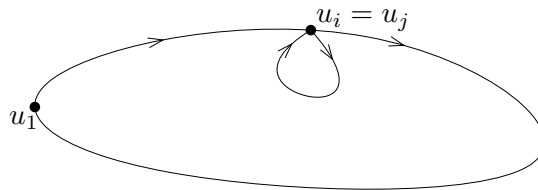
Lema 5.5.7. *Nech $\mathcal{G} = (V, \varphi)$ je silne súvislý ohodnotený graf nad množinou $\mathbb{R}_{\geq 0}$ obsahujúci aspoň jednu hranu, $p \in \mathbb{N} \setminus \{0\}$ a $u \in V$. Potom sú nasledujúce tvrdenia ekvivalentné:*

- (i) Číslo p je periódou grafu $\mathcal{G}$ – čiže $p(\mathcal{G}) = p$.
- (ii) Číslo p je najväčším spoločným deliteľom dĺžok uzavretých sledov v $\mathcal{G}$.
- (iii) Číslo p je najväčším spoločným deliteľom dĺžok uzavretých sledov v $\mathcal{G}$ prechádzajúcich cez vrchol u .

Dôkaz. Dokážeme najprv ekvivalenciu tvrdení (i) a (ii). Je zrejmé, že každý spoločný deliteľ dĺžok uzavretých sledov v $\mathcal{G}$ musí deliť aj dĺžky všetkých orientovaných kružníc v $\mathcal{G}$. Stačí teda dokázať, že pre každý uzavretý sled dĺžky t v grafe $\mathcal{G}$ musí byť číslo t deliteľné periódou $p(\mathcal{G})$ uvažovaného grafu $\mathcal{G}$. Za účelom sporu predpokladajme, že v $\mathcal{G}$ existuje uzavretý sled

$$\gamma = (u_1, u_2, \dots, u_t, u_1)$$

dĺžky $t \in \mathbb{N}$ takej, že $p(\mathcal{G})$ nedelí t . Bez ujmy na všeobecnosti navyše predpokladajme, že je sled γ minimálny s touto vlastnosťou. Nemôže ísť o orientovanú kružnicu, pretože v takom prípade by číslo t muselo byť deliteľné periódou $p(\mathcal{G})$; nutne tiež $t \geq 1$. Existujú teda indexy $i, j \in [t]$ také, že $i < j$ a $u_i = u_j$. Predpokladajme, že sú indexy $i < j$ s touto vlastnosťou zvolené tak, aby bol ich rozdiel $j - i$ najmenší možný – sled $(u_i, \dots, u_j)$ je teda orientovanou kružnicou. Táto situácia je schematicky znázornená na obrázku 5.3.



Obr. 5.3: Uzavretý sled γ musí obsahovať orientovanú kružnicu začínajúcu a končiacu vo vrchole $u_i = u_j$.

Nenulová dĺžka $j - i$ orientovanej kružnice $(u_i, \dots, u_j)$ musí byť deliteľná periódou $p(\mathcal{G})$. Uzavretý sled

$$\gamma' = (u_1, u_2, \dots, u_i, u_{j+1}, \dots, u_t, u_1)$$

ktorý vznikne zo sledu γ odstránením orientovanej kružnice $(u_i, \dots, u_j)$, však musí byť dĺžky $t - (j - i)$. Tá nemôže byť deliteľná periódou $p(\mathcal{G})$, pretože v opačnom prípade by bolo periódou $p(\mathcal{G})$ deliteľné aj číslo $t = (t - (j - i)) + (j - i)$. Dostávame teda spor s minimálnosťou sledu γ .

Pokračujeme ekvivalenciou tvrdení (ii) a (iii). Ľubovoľný spoločný deliteľ všetkých dĺžok uzavretých sledov v $\mathcal{G}$ musí súčasne deliť aj všetky dĺžky uzavretých sledov prechádzajúcich cez vrchol u . Zostáva dokázať, že každý spoločný deliteľ p dĺžok uzavretých sledov prechádzajúcich cez u delí aj dĺžku ľubovoľného iného uzavretého sledu v $\mathcal{G}$. Nech je takýto deliteľ p daný pevne a za účelom sporu predpokladajme, že

$$\gamma = (u_1, u_2, \dots, u_t, u_1)$$

je uzavretý sled dĺžky t , ktorá nie je deliteľná číslom p . Sled γ zrejme nemôže prechádzať cez vrchol u . Keďže je ale graf $\mathcal{G}$ silne súvislý, musí existovať uzavretý sled $\gamma' = (u_1, u_2, \dots, u_k, u_1)$ začínajúci a končiaci vo vrchole u_1 a súčasne prechádzajúci cez u . Jeho dĺžka k musí byť deliteľná číslom p . Spojením sledov γ a γ' však získavame uzavretý sled prechádzajúci cez vrchol u , ktorého dĺžka $t + k$ nie je deliteľná číslom p – spor. $\square$

Pripomeňme si *Bézoutovu rovnosť* – presnejšie jej všeobecnejší variant pre $n \geq 1$ celých čísel. Podľa nej pre ľubovoľné $x_1, \dots, x_n \in \mathbb{Z}$ existujú celočíselné koeficienty $a_1, \dots, a_n \in \mathbb{Z}$ také, že

$$a_1 x_1 + \dots + a_n x_n = \gcd\{x_1, \dots, x_n\}. \quad (5.25)$$

Na dôkaz Bézoutovej rovnosti možno využiť pozorovanie, že množina

$$I = \{a_1x_1 + \dots + a_nx_n \mid a_1, \dots, a_n \in \mathbb{Z}\} \quad (5.26)$$

evidentne tvorí ideál v okruhu $\mathbb{Z}$. Ak $x_1 = \dots = x_n = 0$, je $I = \{0\}$ a $\gcd\{x_1, \dots, x_n\} = 0$. V opačnom prípade ideál I obsahuje aspoň jedno kladné číslo a musí byť generovaný svojím najmenším kladným prvkom m – to jest

$$I = (m) = \{cm \mid c \in \mathbb{Z}\}. \quad (5.27)$$

Zrejme totiž $I \supseteq (m)$ a keby I obsahoval prvok x , ktorý nie je deliteľný číslom m , bolo by číslo

$$|x| - m \left\lfloor \frac{|x|}{m} \right\rfloor$$

kladným prvkom I menším, než m .²⁰ Keďže zrejme $x_1, \dots, x_n \in I$, musí byť m spoločným deliteľom týchto čísel. Pre každý spoločný deliteľ d čísel $x_1, \dots, x_n$ navyše z (5.26) vyplýva $I \subseteq (d)$, a preto m , pre ktoré je $I = (m)$, musí byť ich *najväčší* spoločný deliteľ. Teda $m = \gcd\{x_1, \dots, x_n\}$ a existencia koeficientov $a_1, \dots, a_n$ z Bézoutovej rovnosti (5.25) je dôsledkom rovností (5.26) a (5.27).

Pri dôkaze ekvivalentnej charakterizácie primitívnych matíc budeme používať jednoduchý dôsledok Bézoutovej rovnosti, ktorý je sformulovaný v nasledujúcej leme.

Lema 5.5.8. *Nech $x_1, \dots, x_n \in \mathbb{N} \setminus \{0\}$ sú také, že $\gcd\{x_1, \dots, x_n\} = 1$. Množina*

$$C(x_1, \dots, x_n) = \{c_1x_1 + \dots + c_nx_n \mid c_1, \dots, c_n \in \mathbb{N}\}$$

potom obsahuje skoro všetky prirodzené čísla (čiže jej komplement v $\mathbb{N}$ je konečný).

Dôkaz. S odvolaním sa na Bézoutovu rovnosť môžeme predpokladať existenciu čísel $a_1, \dots, a_n \in \mathbb{Z}$ takých, že $a_1x_1 + \dots + a_nx_n = 1$. Evidentne pritom existujú $b_1, \dots, b_n \in \mathbb{N}$ také, že pre $k = 1, \dots, n$ je $b_k + (x_1 - 1)a_k \geq 0$. Je zrejme, že pre takéto $b_1, \dots, b_n$ bude uvedená nerovnosť splnená, aj keď v nej číslo $x_1 - 1$ nahradíme ľubovoľným menším alebo rovným prirodzeným číslom j : pre $k = 1, \dots, n$ a $j = 0, \dots, x_1 - 1$ teda

$$b_k + ja_k \geq 0. \quad (5.28)$$

Dokážeme, že množina $C(x_1, \dots, x_n)$ obsahuje všetky prirodzené čísla väčšie alebo rovné číslu

$$m := b_1x_1 + \dots + b_nx_n.$$

Pre všetky prirodzené čísla $x \in C(x_1, \dots, x_n)$ musí byť aj $x + x_1 \in C(x_1, \dots, x_n)$: ak totiž

$$x = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

pre nejaké $c_1, \dots, c_n \in \mathbb{N}$, tak

$$x + x_1 = (c_1 + 1)x_1 + c_2x_2 + \dots + c_nx_n.$$

Na dôkaz lemy teda stačí ukázať, že množina $C(x_1, \dots, x_n)$ obsahuje čísla $m, m + 1, \dots, m + x_1 - 1$. Pre $j = 0, \dots, x_1 - 1$ ale

$$m + j = m + j1 = b_1x_1 + \dots + b_nx_n + j(a_1x_1 + \dots + a_nx_n) = (b_1 + ja_1)x_1 + \dots + (b_n + ja_n)x_n,$$

pričom z (5.28) vyplýva, že všetky koeficienty $(b_1 + ja_1), \dots, (b_n + ja_n)$ sú nezáporné. V dôsledku toho teda $m + j \in C(x_1, \dots, x_n)$, čo bolo treba dokázať. $\square$

So znalosťou práve dokázanej lemy môžeme prejsť k už spomenuť ekvivalentnej charakterizácii primitívnych matíc.

²⁰Mimochodom sme práve dokázali, že $\mathbb{Z}$ je okruh hlavných ideálov.

Veta 5.5.9. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A = (a_{i,j})_{n \times n} \in \mathbb{R}_{\geq 0}^{n \times n}$ je ireducibilná matica s $\varrho(A) > 0$. Matica A je primitívna práve vtedy, keď existuje prirodzené číslo m také, že $A^m \in \mathbb{R}_{> 0}^{n \times n}$.*

Dôkaz. Predpokladajme najprv, že je matica A primitívna. Graf $\mathcal{G} = \text{Gr}(A)$ je v takom prípade aperiodický a podľa lemy 5.5.7 tak je pre každý vrchol $u \in V(\mathcal{G})$ najväčším spoločným deliteľom dĺžok uzavretých sledov v $\mathcal{G}$ prechádzajúcich cez u číslo 1. Ku každému z týchto uzavretých sledov zrejme prislúcha uzavretý sled rovnakej dĺžky, ktorého počiatočným a zároveň koncovým vrcholom je vrchol u . Ak pritom budeme zakaždým vyberať sled, ktorého dĺžka nie je deliteľná najväčším spoločným deliteľom dĺžok už vybraných sledov, ľahko pre každý vrchol $u \in V(\mathcal{G})$ získame konečne veľa uzavretých sledov $\gamma_{u,1}, \dots, \gamma_{u,n_u}$ začínajúcich a končiacich v u takých, že $\gcd\{|\gamma_{u,1}|, \dots, |\gamma_{u,n_u}|\} = 1$.

Podľa lemy 5.5.8 tak pre všetky $u \in V(\mathcal{G})$ existuje prirodzené číslo m_u také, že pre všetky $t \geq m_u$ existuje v grafe $\mathcal{G}$ uzavretý sled dĺžky t začínajúci a končiaci vo vrchole u . Pre

$$m = (n - 1) + \max\{m_u \mid u \in V(\mathcal{G})\}$$

a všetky $t \geq m$ tak musí pre každú dvojicu vrcholov $u, v \in V(\mathcal{G})$ existovať sled γ dĺžky t vedúci v grafe $\mathcal{G}$ z vrcholu u do vrcholu v : ten pozostáva zo sledu γ_1 z u do v dĺžky najviac $n - 1$ a z uzavretého sledu γ_2 začínajúceho a končiaceho vo vrchole v o dĺžke $t - |\gamma_1| \geq m - (n - 1) \geq m_v$.

Keďže je $\mathcal{G}$ ohodnoteným grafom nad $\mathbb{R}_{\geq 0}$, musí byť ohodnotenie ľubovoľného sledu v tomto grafe kladné. Podľa vety 1.3.10 a vyššie učených zistení je teda matica A^t kladná pre všetky $t \geq m$; špeciálne je kladná aj matica A^m .

Nech je teraz pre nejaké $m \in \mathbb{N}$ matica A^m kladná. Matica A je ireducibilná a vďaka predpokladu $\varrho(A) > 0$ nejde o nulovú maticu typu 1×1 . Ľahko teda vidieť, že každý riadok aj stĺpec matice A musí obsahovať aspoň jeden nenulový prvok. V dôsledku toho a kladnosti matice A^m ľahko zisťujeme, že pre všetky $t \geq m$ musí byť kladná aj matica A^t . Špeciálne tak musí byť kladný napríklad aj prvok v prvom riadku a prvom stĺpci tejto matice. Pre vrchol 1 grafu $\mathcal{G} = \text{Gr}(A)$ tak z vety 1.3.10 a kladnosti ohodnotení všetkých sledov v $\mathcal{G}$ vyplýva, že pre všetky $t \geq m$ existuje v $\mathcal{G}$ uzavretý sled dĺžky t začínajúci a končiaci vo vrchole 1. To je evidentne možné len v prípade, že najväčší spoločný deliteľ dĺžok uzavretých sledov prechádzajúcich cez vrchol 1 je rovný jednej. Podľa lemy 5.5.7 teda $p(\mathcal{G}) = 1$, z čoho vďaka vete 5.5.4 vyplýva aj $p(A) = 1$. Matica A je teda primitívna. $\square$

Poznámka 5.5.10. Ľahko vidieť, že v prípade existencie $m \in \mathbb{N}$ takého, že pre maticu $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je $A^m \in \mathbb{R}_{> 0}^{n \times n}$, musí byť graf matice A silne súvislý a matica A teda musí byť ireducibilná.

Kritérium z predchádzajúcej vety tak v skutočnosti možno použiť pre ľubovoľnú nezápornú maticu $A \in \mathbb{R}_{\geq 0}^{n \times n}$ takú, že $n \in \mathbb{N} \setminus \{0\}$ a $\varrho(A) > 0$: ak $A^m \in \mathbb{R}_{> 0}^{n \times n}$ pre nejaké $m \in \mathbb{N}$, musí byť A primitívna ireducibilná matica.

5.6 Aplikácie Perronovej-Frobeniovej teórie

Naznačíme teraz niekoľko možných aplikácií teórie nezáporných matíc a Perronovej-Frobeniovej vety. Začneme pritom tým, že preskúmame dôsledky našich zistení o nezáporných maticiach pre enumeračnú metódu z oddielov 5.2 a 5.3. Každá matica nad $\mathbb{N}$ – a v dôsledku toho aj každá matica susednosti ohodnoteného grafu nad $\mathbb{N}$ – je totiž nezápornou maticou.

Príklad 5.6.1. Nech $\mathcal{G} = ([n], \varphi)$ je aperiodický silne súvislý ohodnotený graf nad $\mathbb{N}$ s aspoň jednou hranou a $i, j \in [n]$ jeho vrcholy. Predpokladajme, že potrebujeme nájsť asymptotický odhad pre počet $S_{i,j}(t)$ sledov dĺžky $t \rightarrow \infty$ vedúcich v multigrafe reprezentovanom grafom $\mathcal{G}$ z vrcholu i do vrcholu j .

Z viet 5.4.19 a 5.5.4 vyplýva, že spektrálny polomer $\varrho > 0$ matice $A := \mathbf{A}(\mathcal{G})$ je jej vlastným číslom algebraickej násobnosti 1 a pre každé ďalšie vlastné číslo λ matice A je $|\lambda| < \varrho$. K vlastnému číslu ϱ navyše prislúcha kladný vlastný vektor $\mathbf{x}_\varrho = (x_1, \dots, x_n)^T$.

S použitím vety 5.2.1 potom zisťujeme, že existuje konštanta $a \in \mathbb{C}$ taká, že

$$S_{i,j}(t) = a\varrho^t + o(\varrho^t). \quad (5.29)$$

Keďže $S_{i,j}(t)$ musí byť pre všetky $t \in \mathbb{N}$ prirodzené číslo, nutne $a \in \mathbb{R}_{\geq 0}$. Dokážeme, že $a > 0$. Skutočne: z vety 1.3.10 vyplýva

$$S_{i,j}(t) = \mathbf{e}_i^T A^t \mathbf{e}_j,$$

kde $\mathbf{e}_k$ sú pre $k = 1, \dots, n$ vektory štandardnej bázy priestoru $\mathbb{C}^n$. Podľa vety 5.5.9 navyše existuje prirodzené číslo m také, že $A^m \in \mathbb{R}_{>0}^{n \times n}$, a teda aj $A^m \mathbf{e}_j \in \mathbb{R}_{>0}^n$. Existuje potom konštanta $\alpha \in \mathbb{R}_{>0}$ taká, že

$$A^m \mathbf{e}_j \geq \alpha \mathbf{x}_\varrho,$$

z čoho pre všetky prirodzené čísla $t \geq m$ vyplýva

$$S_{i,j}(t) = \mathbf{e}_i^T A^t \mathbf{e}_j = \mathbf{e}_i^T A^{t-m} A^m \mathbf{e}_j \geq \alpha \mathbf{e}_i^T A^{t-m} \mathbf{x}_\varrho = \alpha \mathbf{e}_i^T \varrho^{t-m} \mathbf{x}_\varrho = \frac{\alpha x_i}{\varrho^m} \varrho^t,$$

z čoho je zrejmé, že v (5.29) musí skutočne byť $a > 0$. Môžeme teda uzavrieť, že

$$S_{i,j}(t) = \Theta(\varrho^t).$$

Pre aperiodické silne súvislé grafy tak spektrálny polomer ich matice susednosti obsahuje takmer celú informáciu o asymptotických vlastnostiach funkcií $S_{i,j}(t)$.

Príklad 5.6.2. Uvažujme teraz namiesto aperiodického grafu ľubovoľný silne súvislý ohodnotený graf $\mathcal{G} = ([n], \varphi)$ nad $\mathbb{N}$. Opäť skúmame pre ľubovoľnú dvojicu vrcholov $i, j \in [n]$ asymptotické vlastnosti funkcie $S_{i,j}(t)$ v prípade, že $t \rightarrow \infty$.

Pokiaľ graf $\mathcal{G}$ pozostáva z jediného vrcholu bez slučky, je zrejmé, že v ňom existuje jediný sled dĺžky 0 a žiaden sled nenulovej dĺžky. V opačnom prípade $p(\mathcal{G}) = p$ pre nejaké nenulové prirodzené číslo p a matica $A := \mathbf{A}(\mathcal{G})$ má spektrálny polomer $\varrho > 0$. Z viet 5.4.19 a 5.5.4 potom vyplýva, že

$$\varrho, e^{1(2\pi i/p)} \varrho, \dots, e^{(p-1)(2\pi i/p)} \varrho$$

sú vlastnými číslami matice A algebraickej násobnosti 1 a pre každé ďalšie vlastné číslo λ matice A je $|\lambda| < \varrho$; k vlastnému číslu ϱ opäť prislúcha kladný vlastný vektor $\mathbf{x}_\varrho = (x_1, \dots, x_n)^T$. Z vety 5.2.1 následne vyplýva existencia konštánt $a_0, \dots, a_{p-1} \in \mathbb{C}$ takých, že

$$S_{i,j}(t) = \left(a_0 + a_1 e^{t(2\pi i/p)} + a_2 e^{2t(2\pi i/p)} + \dots + a_{p-1} e^{(p-1)t(2\pi i/p)} \right) \varrho^t + o(\varrho^t). \quad (5.30)$$

Dokážeme, že pre $q = 0, \dots, p-1$ buď pre všetky $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ platí

$$S_{i,j}(t) = 0$$

alebo pre $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ je

$$S_{i,j}(t) = C_q \varrho^t + o(\varrho^t) = \Theta(\varrho^t),$$

kde C_q je kladná reálna konštanta. Ak teda množinu $\mathbb{N}$ rozložíme na p tried podľa zvyšku po delení číslom p , je funkcia $S_{i,j}(t)$ na každej z týchto tried buď nulová, alebo sa hodnoty

$$a_0, a_1 e^{t(2\pi i/p)}, a_2 e^{2t(2\pi i/p)}, \dots, a_{p-1} e^{(p-1)t(2\pi i/p)}$$

nasčítajú na kladnú konstantu C_q , v dôsledku čoho je pre t z tejto triedy $S_{i,j}(t) = \Theta(\varrho^t)$.

Keďže ale musí byť hodnota $S_{i,j}(t)$ pre všetky $t \in \mathbb{N}$ prirodzená, z reprezentácie (5.30) je zrejmé, že pre $q = 0, \dots, p-1$ a $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ je konštanta²¹

$$C_q := a_0 + a_1 e^{t(2\pi i/p)} + a_2 e^{2t(2\pi i/p)} + \dots + a_{p-1} e^{(p-1)t(2\pi i/p)}$$

reálna a nezáporná, pričom pre $q = 0, \dots, p-1$ a $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ z (5.30) dostávame

$$S_{i,j}(t) = C_q \varrho^t + o(\varrho^t). \quad (5.31)$$

Zostáva ukázať, že $S_{i,j}(t) = 0$ kedykoľvek $C_q = 0$.

²¹Naozaj ide pre každé q o konstantu, pretože hodnoty $e^{t(2\pi i/p)}, \dots, e^{(p-1)t(2\pi i/p)}$ závisia len na zvyšku čísla t po delení číslom p .

Pre všetky $t \in \mathbb{N}$ ale hodnota

$$S_{i,*}(t) := \mathbf{e}_i^T A^t \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

udáva počet sledov dĺžky t vedúcich v grafe $\mathcal{G}$ z vrcholu i do ľubovoľného cieľového vrcholu. Zjavne ďalej existuje konštanta $\alpha \in \mathbb{R}_{>0}$ taká, že $(1, \dots, 1)^T \geq \alpha \mathbf{x}_\varrho$. Potom

$$S_{i,*}(t) \geq \alpha \mathbf{e}_i^T A^t \mathbf{x}_\varrho = \alpha \mathbf{e}_i^T \varrho^t \mathbf{x}_\varrho = \alpha x_i \varrho^t. \quad (5.32)$$

Súčasne ale pre všetky $t \in \mathbb{N}$ je

$$S_{i,*}(t) = \sum_{k=1}^n S_{i,k}(t). \quad (5.33)$$

Preto²²

$$S_{i,*}(t) = \Theta(\varrho^t).$$

Uvažujme teraz pre pevne dané $q \in \{0, 1, \dots, p-1\}$ množinu K_q pozostávajúcu zo všetkých vrcholov $k \in V(\mathcal{G})$ takých, že pre nejaké $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ je $S_{i,k}(t) > 0$. Pre $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ teda

$$S_{i,*}(t) = \sum_{k \in K_q} S_{i,k}(t) = \Theta(\varrho^t). \quad (5.34)$$

Ak je množina $K_q = \{k\}$ jednoprvková, zjavne musí pre všetky takéto t byť $S_{i,k}(t) = S_{i,*}(t)$, v dôsledku čoho $S_{i,k}(t) = \Theta(\varrho^t)$ a $S_{i,k'}(t) = 0$ pre všetky $k' \in V(\mathcal{G}) \setminus \{k\}$. Predpokladajme teraz, že množina K_q obsahuje dva rôzne vrcholy k_1, k_2 a uvažujme ľubovoľný sled γ_1 z vrcholu i do vrcholu k_1 taký, že $|\gamma_1| \equiv q \pmod{p}$ a pevne daný sled γ_2 z vrcholu i do vrcholu k_2 taký, že $d_2 := |\gamma_2| \equiv q \pmod{p}$. Graf $\mathcal{G}$ je silne súvislý, v dôsledku čoho možno sled γ_1 predĺžiť na sled γ'_1 tak, aby sled γ'_1 končil vo vrchole i a aby bolo $|\gamma'_1| = |\gamma_1| + d_1$ pre nejaké fixné $d_1 \in \mathbb{N}$. Sled γ'_1 je uzavretý a jeho dĺžka je preto podľa lemy 5.5.7 deliteľná číslom p . Spojením sledov γ'_1 a γ_2 potom získavame sled z i do k_2 , ktorého dĺžka $|\gamma_1| + d_1 + d_2$ má po delení číslom p zvyšok q . Musí tak existovať $h \in \mathbb{N} \setminus \{0\}$ také, že $d_1 + d_2 = hp$.

V dôsledku toho zisťujeme, že ľubovoľný sled dĺžky $t \equiv q \pmod{p}$ vedúci z vrcholu i do vrcholu k_1 možno predĺžiť na sled dĺžky $t + hp$ z vrcholu i do vrcholu k_2 , a teda

$$S_{i,k_1}(t) \leq S_{i,k_2}(t + hp).$$

Keďže sú však k_1, k_2 ľubovoľné dva prvky množiny K_q , musí pre nejaké $h' \in \mathbb{N} \setminus \{0\}$ a všetky prípustné t platiť aj nerovnosť

$$S_{i,k_2}(t) \leq S_{i,k_1}(t + h'p).$$

Z (5.31) a (5.34) teda zisťujeme, že pre $j \in K_q$ musí skutočne pre všetky $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ byť

$$S_{i,j}(t) = C_q \varrho^t + o(\varrho^t),$$

kde C_q je kladná reálna konštanta. Keďže pre $j \notin K_q$ priamo z definície množiny K_q vyplýva $S_{i,j}(t) = 0$ pre všetky $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$, je dôkaz nášho tvrdenia týmto pozorovaním dokončený.

²²Z (5.32) dostávame odhad $S_{i,*}(t) = \Omega(\varrho^t)$. Odhad $S_{i,*}(t) = O(\varrho^t)$ zas vyplýva z (5.33) a (5.30).

Príklad 5.6.3. Analýzu z predošlého príkladu možno prostredníctvom normálneho tvaru reducibilnej matice rozšíriť aj na ľubovoľné – t. j. nie nutne silne súvislé – grafy. Takéto úvahy presahujú rámec tohto textu; poznamenajme však aspoň, že v tomto všeobecnom prípade možno množinu $\mathbb{N}$ rozložiť na p tried podľa zvyšku po delení nejakým číslom p tak, že pre počet sledov $S_{i,j}(t)$ vedúcich z vrcholu i do vrcholu j pre $q = 0, \dots, p-1$ a $t \in \mathbb{N}$ s $t \equiv q \pmod{p}$ je buď

$$S_{i,j}(t) = 0,$$

alebo

$$S_{i,j}(t) = \Theta(t^{k_q} \varrho_q^t),$$

kde ϱ_q je najväčší spomedzi spektrálnych polomerov matíc silne súvislých komponentov, cez ktoré prechádza nejaký sled z i do j dĺžky t takej, že $t \equiv q \pmod{p}$ a k_q je maximálne číslo také, že niektorý z takýchto sledov prechádza cez k_q komponentov, ktorých matica má spektrálny polomer ϱ_q .

Nasledujúci príklad súvisí s jednou z najvýznamnejších oblastí aplikácie teórie nezáporných matíc vôbec – s *Markovovými reťazcami*.

Príklad 5.6.4. *Konečným Markovovým reťazcom* môžeme nazvať – aj keď tým pádom nepôjde o úplne bežnú definíciu – ľubovoľný ohodnotený graf $\mathcal{M} = ([n], \varphi)$ nad $[0, 1]$ taký, že pre všetky $i \in [n]$ je

$$\sum_{j=1}^n \varphi(i, j) = 1.$$

Vrcholy Markovovho reťazca sa obyčajne nazývajú jeho *stavmi*. Ohodnotenia hrán reprezentujú *pravdepodobnosti prechodov* medzi stavmi, pričom súčet týchto pravdepodobností cez všetky hrany vychádzajúce z pevne daného stavu $i \in [n]$ je vždy rovný jednej. Ak teda $\mathbf{A}(\mathcal{M}) = A = (a_{i,j})_{n \times n}$, pre $i = 1, \dots, n$ je

$$\sum_{j=1}^n a_{i,j} = 1;$$

hovoríme, že matica A je *riadkovo stochastická*. Ľahko vidieť, že pre spektrálny polomer riadkovo stochastickej matice A je vždy $\varrho(A) = \varrho = 1$, pričom k ϱ prislúcha vlastný vektor $\mathbf{x}_\varrho = (1, \dots, 1)^T$.

Konečný Markovov reťazec $\mathcal{M}$ nazveme *nerozložiteľným*, ak je graf $\mathcal{M}$ silne súvislý – čiže ak je matica $A = \mathbf{A}(\mathcal{M})$ ireducibilná. V takom prípade je podľa Perronovej-Frobeniovej vety vlastné číslo $\varrho = 1$ matice A algebraickej násobnosti 1. Ak je navyše Markovov reťazec $\mathcal{M}$ *aperiodický* – to je práve vtedy, keď je $\mathcal{M}$ aperiodický ako graf – musí byť $\varrho = 1$ jediným vlastným číslom matice A s absolútnou hodnotou 1, pričom matica A je v tomto prípade primitívna.

Distribúcia pravdepodobností jednotlivých stavov je daná riadkovým vektorom

$$\mathbf{p}^T = (p_1, \dots, p_n)$$

takým, že $p_1 + \dots + p_n = 1$. *Počiatočná distribúcia* je daná nejakým takýmto vektorom $\mathbf{i}^T$ a distribúciu po t krokoch možno vyjadriť ako $\mathbf{i}^T A^t$. Distribúciu $\mathbf{p}^T$ nazveme *stacionárnou distribúciou reťazca* $\mathcal{M}$, ak $\mathbf{p}^T A = \mathbf{p}^T$ – pravdepodobnosti jednotlivých stavov sa už teda ďalej nemenia.

Uvažujme teraz ľubovoľný *nerozložiteľný* konečný Markovov reťazec $\mathcal{M}$ a distribúciu $\mathbf{p}^T$ takú, že

$$\mathbf{p}^T A = \mathbf{p}^T. \quad (5.35)$$

To možno prepísať ako

$$A^T \mathbf{p} = \mathbf{p},$$

z čoho vyplýva, že $\mathbf{p}$ je vlastným vektorom matice A^T prislúchajúcim k vlastnému číslu 1. Podobne ako matica A je ale aj jej transpozícia A^T ireducibilnou maticou so spektrálnym polomerom rovným jednej. Podľa Perronovej-Frobeniovej vety teda existuje *práve jedna* stacionárna distribúcia $\mathbf{p}^T$ reťazca $\mathcal{M}$.

Táto distribúcia však ešte nemusí byť *limitnou distribúciou* reťazca $\mathcal{M}$ – nemusí totiž platiť

$$\lim_{t \rightarrow \infty} \mathbf{i}^T A^t = \mathbf{p}^T.$$

Ľahko vidieť, že limitná distribúcia, ak existuje, musí byť stacionárnou distribúciou. Čitateľ však istotne ľahko príde s príkladom nerozložiteľného Markovovho reťazca takého, že jeho stacionárna distribúcia nie je limitnou distribúciou (a v dôsledku toho žiadna limitná distribúcia neexistuje).

Stacionárna distribúcia *aperiodického* nerozložiteľného reťazca $\mathcal{M}$ je ale vždy aj jeho limitnou distribúciou. Skutočne: keďže je reťazec $\mathcal{M}$ aperiodický, matica A musí byť primitívna. Jej spektrálny polomer $\varrho = 1$ je teda jej vlastným číslom algebraickej násobnosti 1 a pre každé ďalšie vlastné číslo λ je $|\lambda| < \varrho$. Podobne ako vo vete 5.2.1 tak zisťujeme, že pre $i, j = 1, \dots, n$ existuje nezáporná reálna konštanta $c_{i,j}$ taká, že pre maticu $A^t = (a_{i,j}[t])_{n \times n}$ pre všetky $t \in \mathbb{N}$ platí

$$a_{i,j}[t] = c_{i,j}1^t + o(1^t) = c_{i,j} + o(1).$$

Je potom zrejmé, že $\lim_{t \rightarrow \infty} A^t = (c_{i,j})_{n \times n}$, v dôsledku čoho existuje aj limita

$$\lim_{t \rightarrow \infty} \mathbf{i}^T A^t = \mathbf{r}^T.$$

Z predchádzajúcich úvah pritom vyplýva, že takéto $\mathbf{r}^T$ musí byť stacionárnou distribúciou reťazca $\mathcal{M}$.

5.7 Maticovo-grafový slovník

V tejto a aj v predchádzajúcich kapitolách sme sa z veľkej časti zaoberali súvisom medzi štvorcovými maticami a ohodnotenými orientovanými grafmi – prvky matíc resp. ohodnotenia hrán sme pritom vyberali z množín s nulou, z polokruhov, zo (spočítateľne) úplných polokruhov, z polokruhov formálnych mocninových radov (nad polokruhom, spočítateľne úplným polokruhom, či poľom), z množiny komplexných čísel, či z množiny nezáporných reálnych čísel, ktorých špeciálnym prípadom sú prirodzené čísla. V závislosti od štruktúry na uvažovanej množine prvkov matíc resp. ohodnotení grafov sme potom definovali rôzne operácie a vlastnosti, ktoré sa neraz ukázali byť vyjadriteľné dvojakým spôsobom: k maticovému pojmu často prislúchal prirodzený grafový náprotivok a naopak. Niektoré z takýchto dvojíc pojmov v nasledujúcom stručne zrekapitulujeme; pridáme však aj niekoľko ďalších dvojíc, ktorých vzájomnú korešpondenciu čitateľ istotne ľahko nahliadne.²³

Maticový pojem	Zodpovedajúci grafový pojem	Množina prvkov
Nenulový prvok matice	Ohodnotenie hrany	Množina s nulou
Nenulový diagonálny prvok	Ohodnotenie slučky	Množina s nulou
Prvok t -tej mocniny matice	Súčet ohodnotení sledov dĺžky t	Polokruh
Prvok iterácie matice	Súčet ohodnotení všetkých sledov	Spočítateľne úplný polokruh, $S \langle\langle \Sigma^+ \rangle\rangle$ pre polokruh S , ...
Transponovaná matica	Transponovaný graf (obrátenie hrán)	Množina s nulou
Ireducibilná matica	Silne súvislý graf	$\mathbb{R}_{\geq 0}$
Reducibilná matica	Graf s aspoň 2 komponentmi	$\mathbb{R}_{\geq 0}$
Nilpotentná matica	Acyklický orientovaný graf	$\mathbb{R}_{\geq 0}$
Kladná matica	Úplný orientovaný graf (so slučkami)	$\mathbb{R}_{\geq 0}$
Normálny tvar reducibilnej matice	Topologické usporiadanie komponentov silnej súvislosti	$\mathbb{R}_{\geq 0}$
Index primitívnosti	Periódna	$\mathbb{R}_{\geq 0}$
Primitívna matica	Aperiodický graf	$\mathbb{R}_{\geq 0}$

²³Viacere z pojmov definovaných pre množinu $\mathbb{R}_{\geq 0}$ by bolo možné rozšíriť aj na všeobecnejšie množiny prvkov; prestala by však potom pre ne platiť väčšina zaujímavých tvrdení.

Cvičenia

1. Preformulujte vetu 5.2.1 pre špeciálny prípad grafov s diagonalizovateľnými maticami susednosti.
2. Bolo spomenuté, že vo vyjadrení $S_{i,j}(t)$ podľa vety 5.2.1 vystupuje práve n neznámych konštánt $a_{\lambda,k}$ pre $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$. Identifikujte grafy, pre ktoré možno $S_{i,j}(t)$ vyjadriť s menším počtom takýchto konštánt a preformulujte vetu 5.2.1 tak, aby pre každý graf poskytovala vyjadrenie hodnôt $S_{i,j}(t)$ s minimálnym možným počtom komplexných konštánt.²⁴
3. Dokážte, že Collatzova-Wielandtova funkcia Γ_A prislúchajúca k nezápornej matici $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je spojitá vo všetkých bodoch množiny $\mathbb{R}_{> 0}^n$.
4. Dokážte, že Collatzova-Wielandtova funkcia Γ_A prislúchajúca k ireducibilnej matici $A \in \mathbb{R}_{\geq 0}^{n \times n}$, kde n je nenulové prirodzené číslo, nemusí byť spojitá v hraničných bodoch simplexu Δ_{n-1} .
5. Dokážte vetu 5.4.21.
6. Nech $\mathcal{G} = (V, \varphi)$ je silne súvislý ohodnotený graf nad množinou $\mathbb{R}_{\geq 0}$ obsahujúci aspoň jednu hranu a $u \in V$. Zistite, či perióda $p(\mathcal{G})$ grafu $\mathcal{G}$ musí byť najväčším spoločným deliteľom dĺžok orientovaných kružníc v $\mathcal{G}$ prechádzajúcich cez vrchol u .
7. Nech $\mathcal{G} = (V, \varphi)$ je silne súvislý ohodnotený graf nad množinou $\mathbb{R}_{\geq 0}$ obsahujúci aspoň jednu hranu. Dokážte, že perióda $p(\mathcal{G})$ grafu $\mathcal{G}$ je najväčším prirodzeným číslom p , pre ktoré existuje funkcia $\chi: V \rightarrow \mathbb{Z}_p$ taká, že pre všetky hrany $(u, v) \in E(\mathcal{G})$ je $\chi(v) = \chi(u) + 1$ (modulo p). Pre $p \geq 2$ je táto funkcia χ veľmi špeciálnym vrcholovým farbením grafu $\mathcal{G}$.
8. Nech n je pevne dané nenulové prirodzené číslo. Nájdite všetky $p \in \mathbb{N} \setminus \{0\}$, pre ktoré existuje ireducibilná matica $A \in \mathbb{R}_{\geq 0}^{n \times n}$ s indexom primitívnosti $p(A) = p$.
9. Nech $n \in \mathbb{N}$ a $\mathcal{I}_n$ je množina všetkých *ireducibilných* matíc $A \in \mathbb{R}_{\geq 0}^{n \times n}$. Zistite, či musí byť $\mathcal{I}_n$ uzavretá na nasledujúce maticové operácie:
 - a) súčet;
 - b) súčin;
 - c) umocňovanie na $t \in \mathbb{N}$;
 - d) umocňovanie na $t \in \mathbb{N} \setminus \{0\}$.
10. Nech $n \in \mathbb{N} \setminus \{0\}$ a $\mathcal{P}_n$ je množina všetkých *primitívnych* ireducibilných matíc $A \in \mathbb{R}_{\geq 0}^{n \times n}$. Zistite, či musí byť $\mathcal{P}_n$ uzavretá na nasledujúce maticové operácie:
 - a) súčet;
 - b) súčin;
 - c) umocňovanie na $t \in \mathbb{N}$;
 - d) umocňovanie na $t \in \mathbb{N} \setminus \{0\}$.
11. Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}_{\geq 0}^{n \times n}$ je nezáporná reálna matica taká, že pre nejaké $t \in \mathbb{N} \setminus \{0\}$ je matica A^t
 - a) ireducibilná;
 - b) primitívna.

Zistite, či musí mať rovnakú vlastnosť aj matica A .

²⁴Toto vyjadrenie by však aj naďalej malo byť rovnaké pre všetky dvojice vrcholov i, j .

Poznámky a odkazy na samoštúdium

Vandermondova matica sa v literatúre definuje viacerými jemne odlišnými spôsobmi. V najčastejšie sa vyskytujúcej variácii definície používanej v tomto texte sa pracuje s jej transpozíciou [57]; pre determinant takto definovanej Vandermondovej matice tak ostáva v platnosti rovnaký vzorec ako vo vete 5.1.2. Možno sa však stretnúť aj s ďalšími alternatívnymi definíciami – napríklad Knuth [66] definuje Vandermondovu maticu tak, že k -ty stĺpec matice $V(\alpha_1, \dots, \alpha_n)$ je pre $k = 1, \dots, n$ daný ako $(\alpha_k^1, \dots, \alpha_k^n)^T$; do vzorca pre determinant takto definovanej Vandermondovej matice tak ešte treba zakomponovať faktor $\alpha_1 \dots \alpha_n$.

Pod *zovšeobecnenou Vandermondovou maticou* sa v literatúre chápu rôzne, často aj vzájomne nesúvisiace, objekty. Pravdepodobne prvá zmienka o zovšeobecnenej Vandermondovej matici v zmysle tohto textu pochádza z Kalmanovho populárneho článku [60], v ktorom sú aj naznačené viaceré možnosti dôkazu vety 5.1.11 hovoriacej o jej determinante (pracuje sa tam však len s konkrétnymi hodnotami s a $n_1, \dots, n_s$). Dôkaz vety 5.1.11 uvedený v tomto texte je rozvedením jedného z týchto náznakov dôkazu pre všeobecný prípad. Rovnaký objekt sa pod zovšeobecnenou Vandermondovou maticou chápe napríklad aj v [34]. O *Hasseho deriváciách*, ktoré sme používali pri dôkaze vety 5.1.11, sa možno dočítať viac napríklad v [54].

Klasickými monografiami o nezáporných maticiach sú napríklad [9] a [86]. Kapitoly venované tejto problematike možno ďalej nájsť napríklad v [76, 48, 57, 84, 101]. Perronova-Frobeniova teória je pomenovaná po O. Perronovi, ktorý sa zaoberal spektrálnymi vlastnosťami *kladných* matíc [90] a po G. Frobeniovi, ktorý jeho zistenia zovšeobecnil na prípad *ireducibilných nezáporných* matíc [46]. Vetu 5.4.16 dokázal H. Wielandt v článku [103].

Pre *index primitívnosti* matice sa v anglickej literatúre zvyčajne používa termín *imprimitivity index*, čiže niečo ako „index neprimitívnosti“ [86, 84]. Občas sa používa aj termín *perióda matice* [78]; naopak pre *periódu grafu* sa možno stretnúť aj s pojmom *index primitívnosti grafu* (angl. *imprimitivity index of a graph*) [86, 17].

Asymptotické odhady z príkladov 5.6.1, 5.6.2 a 5.6.3 je možné, snáď o čosi elegantnejšie, odvodiť aj s využitím vytvárajúcich funkcií a analytickej kombinatoriky [44]. Táto metóda je však založená na skúmaní vlastností Laurentových rozvojov racionálnych funkcií v ich singularitách (póloch), čo je v porovnaní so stále relatívne elementárnou Perronovou-Frobeniovou teóriou o poznanie ťažšia mašinéria. Konečným Markovovým reťazcom je venovaná napríklad kniha [63].

Kapitola 6

Diferenčné rovnice

V krátkosti si teraz predstavíme ďalšiu významnú oblasť aplikácie teórie okolo vlastných čísel matic: riešenie určitých typov rekurencií, ktoré budeme nazývať *diferenčnými rovnicami*.

Pojem diferenčnej rovnice pre nás bude synonymom pre pojem rekurencie konečného rádu – čiže vzťahu, v ktorom vystupuje neznáma funkcia $x: \mathbb{N} \rightarrow \mathbb{C}$, pričom hodnota $x(t)$ tejto funkcie pre všetky $t \in \mathbb{N}$ závisí iba od konečného počtu predchádzajúcich hodnôt $x(t-1), \dots, x(t-k)$, kde k je fixná konštanta. Pomenovanie „diferenčná rovnica“ – prevládajúce napríklad v oblastiach, kde sa tieto vzťahy uvažujú v súvislosti s dynamickými systémami alebo s numerickým riešením diferenciálnych rovníc – vychádza z pozorovania, že ľubovoľný rekurentný vzťah uvedeného typu možno chápať aj ako rovnicu, v ktorej okrem neznámej funkcie x vystupuje konečný počet funkcií získaných z x opakovaným aplikovaním operátora *konečnej diferencie*, ktorý je diskretnou obdobou operátora derivácie. Diferenčné rovnice sa tak v určitom presnom zmysle stávajú diskretnou obdobou diferenciálnych rovníc. Podobne aj metódy, ktoré budeme na riešenie diferenčných rovníc používať, bude možné chápať ako diskkrétne analógie metód na riešenie niektorých základných typov diferenciálnych rovníc.

V skutočnosti možno na konečných diferenciách založiť kompletnú teóriu paralelnú s diferenciálnym a integrálnym počtom, v ktorej sa namiesto funkcií spojitaj reálnej alebo komplexnej premennej narába s funkciami diskretnaj premennej – čiže s postupnosťami. Táto teória sa nazýva *diferenčným počtom* alebo *konečným kalkuľom* a okrajovo do nej nazrieme v prvom oddiele tejto kapitoly. Podstatné však pre nás budú najmä metódy na riešenie diferenčných rovníc založené na lineárnej algebre, pri ktorých si vystačíme aj s bežným pohľadom na diferenčné rovnice ako na rekurencie. O diferenčnom počte sa tu zmieňujeme pre lepšie pochopenie genézy pojmu diferenčnej rovnice a s cieľom zdôraznenia hĺbky, ktorá sa skrýva za zjavnou podobnosťou metód riešenia diferenčných a diferenciálnych rovníc.

6.1 Základy diferenčného počtu

Analogicky k operátoru derivácie, definovaného na vektorovom priestore (v ľubovoľnom slova zmysle) diferencovateľných funkcií predpisom

$$\left(\frac{d}{dt}f\right)(t) = \frac{d}{dt}f(t) := \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h},$$

možno na vektorovom priestore $\mathcal{F} := \mathbb{C}^{\mathbb{N}}$ všetkých funkcií¹ z $\mathbb{N}$ do $\mathbb{C}$ uvažovať lineárny operátor² *konečnej diferencie* Δ , ktorý je pre všetky $f: \mathbb{N} \rightarrow \mathbb{C}$ a $t \in \mathbb{N}$ daný predpisom

$$(\Delta f)(t) = \Delta f(t) := f(t+1) - f(t).$$

¹Ide o vektorový priestor nad poľom $\mathbb{C}$, v ktorom funkcie z $\mathbb{N}$ do $\mathbb{C}$ uvažujeme spolu so štandardnými operáciami sčítania funkcií a násobenia funkcie skalárom z $\mathbb{C}$. Rovnako dobre by bolo možné pracovať aj nad priestorom $\mathbb{R}^{\mathbb{N}}$.

²Ide tu o skryté tvrdenie: aby sme dokázali, že je operátor konečnej diferencie skutočne lineárny, je potrebné argumentovať, že pre všetky $f, g \in \mathcal{F}$ a $a \in \mathbb{C}$ je $\Delta(f+g) = \Delta f + \Delta g$ a $\Delta(af) = a\Delta f$. Toto jednoduché mechanické cvičenie prenechávame čitateľovi.

Konečné diferencie sú teda diskretnou obdobou derivácií, pričom ich úloha v diferenčnom počte je porovnateľná s úlohou derivácií v diferenciálnom počte. V situáciách, keď je nutné zdôrazniť, že ide o konečnú diferenciu funkcie premennej t , píšeme namiesto Δ aj Δ_t .

Ľahko vidieť, že lineárne operátory na $\mathcal{F}$ tvoria spoločne s bežne definovanými operáciami súčtu a zloženia operátorov okruh, ktorý budeme označovať $\mathcal{L}(\mathcal{F})$. V tomto okruhu budeme interpretovať aj mocniny operátorov – napríklad Δ^k teda označuje konečnú diferenciu aplikovanú k -krát, pričom v takom prípade hovoríme o *konečnej diferencii k -teho rádu*. Priamym výpočtom získavame nasledujúce vzťahy pre konečné diferencie druhého a tretieho rádu:

$$\begin{aligned}\Delta^2 f(t) &= \Delta(\Delta f(t)) = \Delta f(t+1) - \Delta f(t) = \\ &= f(t+2) - f(t+1) - (f(t+1) - f(t)) = \\ &= f(t+2) - 2f(t+1) + f(t), \\ \Delta^3 f(t) &= \Delta^2(\Delta f(t)) = \Delta f(t+2) - \Delta^2 f(t+1) + \Delta f(t) = \\ &= f(t+3) - f(t+2) - (2f(t+2) - 2f(t+1)) + f(t+1) - f(t) = \\ &= f(t+3) - 3f(t+2) + 3f(t+1) - f(t).\end{aligned}$$

Indukciou možno ľahko dokázať, že pre všetky $k \in \mathbb{N}$ je

$$\begin{aligned}\Delta^k f(t) &= f(t+k) - \binom{k}{1}f(t+k-1) + \dots + \binom{k}{k-1}(-1)^{k-1}f(t+1) + (-1)^k f(t) = \\ &= \sum_{j=0}^k \binom{k}{j} (-1)^j f(t+k-j).\end{aligned}$$

V skutočnosti je tento vzorec dokonca bezprostredným dôsledkom binomickej vety

$$(x+y)^k = \sum_{j=0}^k \binom{k}{j} x^j y^{k-j},$$

ktorá evidentne platí pre každú dvojicu vzájomne komutujúcich prvkov x, y ľubovoľného polokruhu. Na priestore $\mathcal{F}$ môžeme uvažovať lineárny operátor posunu E daný pre všetky $f: \mathbb{N} \rightarrow \mathbb{C}$ a $t \in \mathbb{N}$ predpisom

$$(Ef)(t) = Ef(t) := f(t+1)$$

a takisto aj *identický operátor* Id , ktorý je pre všetky $f: \mathbb{N} \rightarrow \mathbb{C}$ a $t \in \mathbb{N}$ daný ako

$$(\text{Id}f)(t) = \text{Id}f(t) := f(t).$$

Ľahko vidieť, že operátory E a $-\text{Id}$ spolu komutujú a zrejme tiež

$$\Delta = E - \text{Id}.$$

Z binomickej vety pre komutujúce prvky okruhu lineárnych operátorov $\mathcal{L}(\mathcal{F})$ tak pre všetky $k \in \mathbb{N}$ dostávame

$$\begin{aligned}\Delta^k f(t) &= (E - \text{Id})^k f(t) = \sum_{j=0}^k \binom{k}{j} (-\text{Id})^j E^{k-j} f(t) = \\ &= \sum_{j=0}^k \binom{k}{j} (-1)^j f(t+k-j).\end{aligned}$$

Nasledujúce vzťahy pre konečné diferencie súčtu a podielu funkcií sú analógiami známych vzorcov pre derivácie.

Tvrdenie 6.1.1. *Nech $f, g: \mathbb{N} \rightarrow \mathbb{C}$ sú funkcie. Potom*

$$a) \Delta(fg) = (\Delta f)g + (Ef)(\Delta g) = (\Delta f)(Eg) + f(\Delta g).$$

Ak navyše $g(t) \neq 0$ pre všetky $t \in \mathbb{N}$, tak

$$b) \Delta\left(\frac{1}{g}\right) = -\frac{\Delta g}{g(Eg)},$$

$$c) \Delta\left(\frac{f}{g}\right) = \frac{(\Delta f)g - f(\Delta g)}{g(Eg)}.$$

Dôkaz.

a) Pre diferenciu súčinu pre všetky $t \in \mathbb{N}$ dostávame

$$\begin{aligned} \Delta(fg)(t) &= f(t+1)g(t+1) - f(t)g(t) = \\ &= (f(t+1)g(t+1) - f(t)g(t+1)) + (f(t)g(t+1) - f(t)g(t)) = \\ &= (\Delta f(t))g(t+1) + f(t)(\Delta g(t)) = ((\Delta f)(Eg) + f(\Delta g))(t). \end{aligned}$$

Po výmene funkcií f, g potom zisťujeme, že aj

$$\Delta(fg)(t) = ((\Delta f)g + (Ef)(\Delta g))(t).$$

b) Priamo z definície konečnej diferencie je

$$\Delta\left(\frac{1}{g}\right)(t) = \frac{1}{g(t+1)} - \frac{1}{g(t)} = \frac{g(t) - g(t+1)}{g(t)g(t+1)} = \left(-\frac{\Delta g}{g(Eg)}\right)(t).$$

c) S použitím predchádzajúcich dvoch dokázaných vzťahov zisťujeme, že

$$\Delta\left(\frac{f}{g}\right) = \Delta\left(\frac{1}{g}f\right) = \left(\Delta\left(\frac{1}{g}\right)\right)f + \left(E\left(\frac{1}{g}\right)\right)(\Delta f) = -\frac{f(\Delta g)}{g(Eg)} + \frac{\Delta f}{Eg} = \frac{(\Delta f)g - f(\Delta g)}{g(Eg)}. \quad \square$$

Evidentne $\Delta_t c = 0$ pre všetky $c \in \mathbb{C}$ a $\Delta_t t = 1$. Konečné diferencie mocninových funkcií t^k pre $k \geq 2$ sú však o niečo komplikovanejšieho charakteru, než je tomu v prípade derivácií. Za analógiu známeho vzorca pre derivácie polynomických funkcií možno považovať skôr nasledujúcu vlastnosť k -teho klesajúceho faktoriálu, definovaného pre $k \in \mathbb{N}$ a všetky $z \in \mathbb{C}$ ako

$$z^{\underline{k}} := z(z-1)\dots(z-k+1).$$

Tvrdenie 6.1.2. *Nech $k \in \mathbb{N} \setminus \{0\}$. Potom $\Delta_t t^{\underline{k}} = kt^{\underline{k-1}}$.*

Dôkaz. Z definície konečnej diferencie je

$$\begin{aligned} \Delta_t t^{\underline{k}} &= (t+1)^{\underline{k}} - t^{\underline{k}} = (t+1)t(t-1)\dots(t-k+2) - t(t-1)\dots(t-k+2)(t-k+1) = \\ &= ((t+1) - (t-k+1))t(t-1)\dots(t-k+2) = kt^{\underline{k-1}}, \end{aligned}$$

čo bolo treba dokázať. □

V skutočnosti možno definíciu k -tych klesajúcich faktoriálov elementárne rozšíriť na všetky $k \in \mathbb{Z}$ a s použitím funkcie gamma aj na všetky $k \in \mathbb{C}$. Dá sa dokázať, že vzorec z predchádzajúceho tvrdenia zostane v platnosti aj pri takomto zovšeobecnení – detaily možno nájsť napríklad v [62].

Obdobou neurčitých integrálov alebo primitívnych funkcií sú v diferenčnom počte *neurčité sumy*. Pre ľubovoľnú funkciu $f: \mathbb{N} \rightarrow \mathbb{C}$ nazveme jej *neurčitou sumou*

$$(\Delta^{-1}f)(t) = \Delta^{-1}f(t) = \sum_t f(t)$$

ľubovoľnú funkciu $F: \mathbb{N} \rightarrow \mathbb{C}$ takú, že $\Delta F = f$. Základným výsledkom o neurčitých sumách je nasledujúca veta, ktorá hovorí o existencii neurčitých súm a o ich jednoznačnosti až na konštantný člen.

Veta 6.1.3. *Nech $f: \mathbb{N} \rightarrow \mathbb{C}$ je funkcia. Neurčitými sumami funkcie f sú potom práve všetky funkcie*

$$\sum_t f(t) = F(t) + c,$$

kde $c \in \mathbb{C}$ je ľubovoľná konštanta a $F: \mathbb{N} \rightarrow \mathbb{C}$ je funkcia daná pre všetky $t \in \mathbb{N}$ predpisom

$$F(t) = \sum_{j=0}^{t-1} f(j).$$

Dôkaz. Zvoľme ľubovoľné $c \in \mathbb{C}$ a nájdime všetky neurčité sumy $F_c: \mathbb{N} \rightarrow \mathbb{C}$ funkcie f také, že

$$F_c(0) = c. \quad (6.1)$$

Aby bola funkcia F_c neurčitou sumou funkcie f , pre všetky $t \in \mathbb{N}$ musí byť

$$\Delta F_c(t) = F_c(t+1) - F_c(t) = f(t),$$

čiže

$$F_c(t+1) = F_c(t) + f(t). \quad (6.2)$$

Indukciou teraz možno ľahko dokázať, že z rovností (6.1) a (6.2) pre všetky $t \in \mathbb{N}$ vyplýva

$$F_c(t) = c + \sum_{j=0}^{t-1} f(j) = F(t) + c,$$

pričom takto definovaná funkcia je neurčitou sumou funkcie f . □

Priamo z definície neurčitých súm vyplýva platnosť vzťahu $\Delta\Delta^{-1}f = f$ pre všetky $f: \mathbb{N} \rightarrow \mathbb{C}$. Dokážeme teraz tvrdenie, podľa ktorého je tiež $\Delta^{-1}\Delta f = f$; všetky funkcie $\Delta^{-1}\Delta f$ pritom vzniknú z funkcie f pripočítaním nejakej komplexnej konštanty d .

Tvrdenie 6.1.4. *Nech $f: \mathbb{N} \rightarrow \mathbb{C}$ je funkcia. Neurčitými sumami funkcie Δf sú potom práve všetky funkcie*

$$\sum_t \Delta f(t) = f(t) + d,$$

kde $d \in \mathbb{C}$ je konštanta.

Dôkaz. Z vety 6.1.3 vyplýva, že neurčitými sumami funkcie Δf sú práve všetky funkcie

$$c + \sum_{j=0}^{t-1} \Delta f(j) = c + \sum_{j=0}^{t-1} (f(j+1) - f(j)) = f(t) - f(0) + c,$$

kde $c \in \mathbb{C}$. Stačí teda vziať $d := c - f(0)$ a tvrdenie je dokázané. □

V tomto texte nie je naším cieľom detailne preskúmať množstvo rôznorodých techník neurčitej sumácie, ktoré sú často obdobami známych techník na výpočet neurčitých integrálov. Naopak sa tu obmedzíme len na demonštráciu asi najzákladnejšej techniky, ktorá spočíva vo využití *linearity neurčitej sumácie* spoločne so znalosťou neurčitých súm niektorých základných funkcií. O linearite neurčitých súm hovorí nasledujúce tvrdenie.

Tvrdenie 6.1.5. *Nech $f, g: \mathbb{N} \rightarrow \mathbb{C}$ sú funkcie a $a, b \in \mathbb{C}$ sú konštanty. Potom³*

$$\sum_t (af(t) + bg(t)) = a \sum_t f(t) + b \sum_t g(t).$$

Dôkaz. S použitím vety 6.1.3 zisťujeme, že jednou z neurčitých súm funkcie $af(t) + bg(t)$ je funkcia

$$\sum_t (af(t) + bg(t)) = \sum_{j=0}^{t-1} (af(j) + bg(j)) = a \sum_{j=0}^{t-1} f(j) + b \sum_{j=0}^{t-1} g(j) = a \sum_t f(t) + b \sum_t g(t),$$

čo bolo treba dokázať. □

V rámci príkladov nižšie si vystačíme so znalosťou neurčitých súm pre k -te klesajúce faktoriály – vzorce pre ne sú priamym dôsledkom vzorcov pre konečné diferencie týchto funkcií.

Tvrdenie 6.1.6. *Nech $k \in \mathbb{N}$. Neurčitými sumami funkcie t^k sú potom práve všetky funkcie*

$$\sum_t t^k = \frac{t^{k+1}}{k+1} + c,$$

kde $c \in \mathbb{C}$ je konštanta.

Dôkaz. Z tvrdenia 6.1.2 a z linearity konečnej diferencie dostávame

$$\Delta \left(\frac{t^{k+1}}{k+1} \right) = \frac{\Delta t^{k+1}}{k+1} = \frac{(k+1)t^k}{k+1} = t^k$$

a funkcia $t^{k+1}/(k+1)$ je tak naozaj neurčitou sumou funkcie t^k . Z vety 6.1.3 potom vyplýva, že práve všetkými neurčitými sumami funkcie t^k sú funkcie typu

$$\frac{t^{k+1}}{k+1} + c$$

pre nejakú komplexnú konštantu c . Tým je tvrdenie dokázané. □

Použitie neurčitých súm je založené predovšetkým na ich nasledujúcej vlastnosti, ktorá ich dáva do súvisu s bežnými – alebo *určitými* – sumami. Nasledujúca veta je diskretnou obdobou Newtonovej-Leibnizovej vety o vzťahu určitých a neurčitých integrálov.

Veta 6.1.7. *Nech $f: \mathbb{N} \rightarrow \mathbb{C}$ je funkcia a $F = \Delta^{-1}f$ je neurčitá suma funkcie f . Pre všetky $r, s \in \mathbb{N}$ potom je*

$$\sum_{t=r}^s f(t) = F(s+1) - F(r).$$

³Nasledujúcu rovnosť treba chápať tak, že nastane pre *nejaký* výber neurčitých súm na jednotlivých jej stranách. Nejde tu o podstatný nedostatok, pretože ľubovoľné dve neurčité sumy tej istej funkcie sa líšia nanajvýš o aditívnu konštantu. Ak je ale aspoň jedna z konštánt a, b nenulová, ľahko dokážeme, že pre *ľubovoľný* výber neurčitých súm na niektorej strane rovnosti možno na jej opačnej strane vybrať neurčité sumy tak, aby si výsledné funkcie boli rovné.

Dôkaz. Z vety 6.1.3 vyplýva, že funkcia F musí byť pre všetky $t \in \mathbb{N}$ a nejakú konštantu $c \in \mathbb{C}$ daná ako

$$F(t) = c + \sum_{j=0}^{t-1} f(j).$$

Preto

$$F(s+1) - F(r) = \left(c + \sum_{j=0}^s f(j) \right) - \left(c + \sum_{j=0}^{r-1} f(j) \right) = \sum_{j=0}^s f(j) - \sum_{j=0}^{r-1} f(j) = \sum_{t=r}^s f(t),$$

čo bolo treba dokázať. $\square$

Príklad 6.1.8. Metódami diferenčného počtu vypočítame pre všetky $n \in \mathbb{N}$ sumu

$$\sum_{t=0}^n t^2.$$

Podľa tvrdenia 6.1.6 je neurčitou sumou funkcie t^2 napríklad funkcia $t^3/3$. Z vety 6.1.7 preto

$$\sum_{t=0}^n t^2 = \frac{(n+1)^3}{3} - \frac{0^3}{3} = \frac{(n+1)n(n-1)}{3}.$$

Príklad 6.1.9. Podobne teraz pre všetky $n \in \mathbb{N}$ vypočítame sumu

$$\sum_{t=0}^n t^2.$$

Tu využijeme vyjadrenie

$$t^2 = t(t-1) + t = t^2 + t^1.$$

S použitím tvrdení 6.1.5 a 6.1.6 tak dostávame

$$\sum_t t^2 = \sum_t (t^2 + t^1) = \sum_t t^2 + \sum_t t^1 = \frac{t^3}{3} + \frac{t^2}{2}.$$

Podľa vety 6.1.7 preto

$$\begin{aligned} \sum_{t=0}^n t^2 &= \left(\frac{(n+1)^3}{3} + \frac{(n+1)^2}{2} \right) - \left(\frac{0^3}{3} + \frac{0^2}{2} \right) = \frac{(n+1)n(n-1)}{3} + \frac{(n+1)n}{2} = \\ &= \frac{(2n+1)(n+1)n}{6}. \end{aligned}$$

Podobne možno postupovať pri výpočte sumy ľubovoľnej polynomickej funkcie.

Týmto jednoduchými príkladmi problematiku sumácie metódami diferenčného počtu zanecháme. O pokročilejších sumačných metódach založených na diferenčnom počte – napríklad o metóde *sumácie per partes*, ktorá je diskretnou obdobou integrovania *per partes* – sa možno dočítať napríklad v [51, 62].

6.2 Diferenčné rovnice a súvisiace pojmy

Prejdime teraz k hlavnému objektu skúmania tejto kapitoly – to jest k *diferenčným rovniciam*. Ak by sme hľadali čo možno najbližšiu analógiu s pojmom (obyčajnej) diferenciálnej rovnice, mohli by sme diferenčnou rovnicou nazvať vzťah

$$F(t, x(t), \Delta x(t), \dots, \Delta^k x(t)) = 0 \quad (6.3)$$

o neznámej funkcii $x: \mathbb{N} \rightarrow \mathbb{C}$ platný pre všetky $t \in \mathbb{N}$, kde $k \in \mathbb{N}$ a $F: \mathbb{N} \times \mathbb{C}^{k+1} \rightarrow \mathbb{C}$ je ľubovoľné zobrazenie.

Naša definícia diferenčnej rovnice bude skutočne tejto formulácii ekvivalentná; bude sa v nej však tiež odzrkadľovať fakt, že diferenčné rovnice sú súčasne aj rekurenciami určitého špeciálneho typu. Každý algebraický výraz utvorený z $t, x(t), \Delta x(t), \dots, \Delta^k x(t)$ je totiž zároveň aj výrazom utvoreným z $t, x(t), x(t+1), \dots, x(t+k)$ – napríklad

$$\begin{aligned} 2\Delta^2 x(t) - \Delta x(t) &= 2(x(t+2) - 2x(t+1) + x(t)) - (x(t+1) - x(t)) = \\ &= 2x(t+2) - 5x(t+1) + 3x(t). \end{aligned}$$

Diferenčnú rovnicu teda definujeme ako rovnicu, v ktorej okrem premennej $t \in \mathbb{N}$ a hodnoty neznámej funkcie $x(t)$ vystupujú aj hodnoty $x(t+1), \dots, x(t+k)$ – alebo ekvivalentne: $Ex(t), \dots, E^k x(t)$.

Definícia 6.2.1. *Diferenčnou rovnicou nazveme pre ľubovoľné zobrazenie $F: \mathbb{N} \times \mathbb{C}^{k+1} \rightarrow \mathbb{C}$ rovnicu*

$$F(t, x(t), x(t+1), \dots, x(t+k)) = 0 \quad (6.4)$$

resp.

$$F\left(t, x(t), Ex(t), \dots, E^k x(t)\right) = 0,$$

ktorej platnosť sa vyžaduje pre všetky $t \in \mathbb{N}$, pričom $k \in \mathbb{N}$ a x označuje neznámu funkciu z $\mathbb{N}$ do $\mathbb{C}$.

Bez ujmy na všeobecnosti budeme predpokladať, že zobrazenie $F: \mathbb{N} \times \mathbb{C}^{k+1} \rightarrow \mathbb{C}$ netriviálnym spôsobom závisí od svojho posledného argumentu, čiže je vzhľadom naň nekonštantné. To znamená, že existujú $m \in \mathbb{N}$, $c_0, \dots, c_{k-1} \in \mathbb{C}$ a $z_1, z_2 \in \mathbb{C}$ také, že

$$F(m, c_0, \dots, c_{k-1}, z_1) \neq F(m, c_0, \dots, c_{k-1}, z_2).$$

Rovnica (6.4) tak „naozaj“ závisí od $x(t+k)$.

Za tohto predpokladu má zmysel definovať *rád* diferenčnej rovnice (6.4) ako prirodzené číslo k . Takáto definícia rádu diferenčnej rovnice *nie je* úplne v súlade s definíciou rádu často používanou v literatúre [96], pri ktorej sa rád rovnice (6.4) chápe ako rozdiel $k - \ell$, kde $\ell \in \{0, \dots, k\}$ je najmenší index taký, že zobrazenie F netriviálne závisí od argumentu zodpovedajúceho $x(t + \ell)$. Pre nás bude napríklad aj diferenčná rovnica

$$x(t+4) + x(t+3) = 0$$

rovnice štvrtého rádu. To má svoje opodstatnenie – čoskoro napríklad uvidíme, že aj pri uvedenej diferenčnej rovnici je na nájdenie jednoznačného riešenia pre všetky $t \in \mathbb{N}$ potrebné zadať ako začiatočné podmienky prvé štyri hodnoty funkcie x , čo je charakteristickým znakom diferenčných rovníc štvrtého rádu. Neskôr sa budeme detailnejšie zaoberať teóriou lineárnych diferenčných rovníc, pričom prideme k zisteniu, že rovnice uvedeného typu sú skutočne akýmsi „degenerovanými“ rovnicami štvrtého rádu, ktorých charakteristický polynóm má aj nulové korene.

Výhodou našej definície rádu diferenčnej rovnice oproti jej spomínanej alternatíve je tiež jej konzistentnosť s chápaním diferenčných rovníc ako rovníc typu (6.3) – v prípade, že tu zobrazenie F netriviálne závisí od svojho posledného argumentu, ide o najvyšší rád k operátora Δ v rovnici (6.3). Naš pojem rádu diferenčnej rovnice je tak aj dobrou diskretnou analógiou rádu diferenciálnej rovnice.

Diferenčnú rovnicu v tvare (6.4) môžeme nazvať aj *implicitnou*, pretože z takéhoto zápisu nemusí byť jasné vyjadrenie hodnoty $x(t+k)$ pomocou hodnôt $t, x(t), x(t+1), \dots, x(t+k-1)$. Diferenčnú rovnicu k -teho rádu v tvare

$$x(t+k) = F(t, x(t), x(t+1), \dots, x(t+k-1)),$$

kde $k \in \mathbb{N}$, x je neznáma funkcia z $\mathbb{N}$ do $\mathbb{C}$ a $F: \mathbb{N} \times \mathbb{C}^k \rightarrow \mathbb{C}$ je zobrazenie, naopak nazveme *explicitnou diferenčnou rovnicou*.

Definícia 6.2.2. *Začiatočnou úlohou danou explicitnou diferenčnou rovnicou*

$$x(t+k) = F(t, x(t), x(t+1), \dots, x(t+k-1)) \quad (6.5)$$

k-teho rádu pre $k \in \mathbb{N}$ a začiatočnými podmienkami

$$x(0) = q_0, \quad x(1) = q_1, \quad \dots, \quad x(k-1) = q_{k-1} \quad (6.6)$$

pre $q_0, \dots, q_{k-1} \in \mathbb{C}$ nazveme problém nájdenia funkcie $f: \mathbb{N} \rightarrow \mathbb{C}$, ktorá po dosadení za neznámu funkciu x vyhovuje rovniciam (6.6) a pre všetky $t \in \mathbb{N}$ aj rovnici (6.5). Takúto funkciu f nazveme *riešením* začiatočnej úlohy danej diferenčnou rovnicou (6.5) a začiatočnými podmienkami (6.6).

Nasledujúce tvrdenie o existencii a jednoznačnosti riešení začiatočných úloh pre diferenčné rovnice je na rozdiel od podobných tvrdení v teórii diferenciálnych rovníc v podstate triviálne.

Veta 6.2.3 (O existencii a jednoznačnosti). *Uvažujme pre $k \in \mathbb{N}$ ľubovoľnú explicitnú diferenčnú rovnicu*

$$x(t+k) = F(t, x(t), x(t+1), \dots, x(t+k-1)) \quad (6.7)$$

k-teho rádu a konštanty $q_0, \dots, q_{k-1} \in \mathbb{C}$. Potom existuje práve jedno riešenie začiatočnej úlohy danej diferenčnou rovnicou (6.7) a začiatočnými podmienkami

$$x(0) = q_0, \quad x(1) = q_1, \quad \dots, \quad x(k-1) = q_{k-1}.$$

Dôkaz. Vezmime funkciu $f: \mathbb{N} \rightarrow \mathbb{C}$, ktorá je pre $t = 0, \dots, k-1$ daná ako $f(t) = q_t$ a pre všetky $t = k, k+1, \dots$ ako

$$f(t) = F(t-k, f(t-k), f(t-k+1), \dots, f(t-1)).$$

Táto funkcia je evidentne riešením uvažovanej začiatočnej úlohy. Ak naopak $g: \mathbb{N} \rightarrow \mathbb{C}$ je funkcia taká, že $g \neq f$, musí existovať $s \in \mathbb{N}$, pre ktoré je $g(s) \neq f(s)$. Bez ujmy na všeobecnosti predpokladajme, že s je najmenšie s touto vlastnosťou. Ak $s \leq k-1$, tak $g(s) \neq f(s) = q_s$ a funkcia g nevyhovuje začiatočným podmienkam; ak $s \geq k$, tak

$$\begin{aligned} g((s-k)+k) &= g(s) \neq f(s) = F(s-k, f(s-k), f(s-k+1), \dots, f(s-1)) = \\ &= F(s-k, g(s-k), g(s-k+1), \dots, g(s-1)) \end{aligned}$$

a funkcia g nevyhovuje diferenčnej rovnici (6.7). Funkcia f je teda naozaj jediným riešením uvažovanej začiatočnej úlohy. $\square$

Z uvedeného je zrejmé, že kedykoľvek pre diferenčnú rovnicu k -teho rádu zadáme menej ako k začiatočných podmienok, bude mať príslušná úloha vždy viac ako jedno riešenie. Ak naopak zadáme viac ako k začiatočných podmienok, nemusí mať výsledná úloha žiadne riešenie – iba v prípade, že „nadbytočné“ začiatočné podmienky vyhovujú diferenčnej rovnici, bude mať úloha jediné riešenie.

Popri jednotlivých diferenčných rovniciach je často užitočné študovať aj ich systémy – tie zohrávajú dôležitú úlohu nielen v rozličných aplikáciách, ale aj pri samotných základoch teórie diferenčných rovníc. Explicitným *systémom diferenčných rovníc* k -teho rádu pre $k \in \mathbb{N}$ nazveme vzťah

$$\mathbf{x}(t+k) = F(t, \mathbf{x}(t), \mathbf{x}(t+1), \dots, \mathbf{x}(t+k-1)) \quad (6.8)$$

s neznámu vektorovou funkciou $\mathbf{x}: \mathbb{N} \rightarrow \mathbb{C}^n$ pre nejaké $n \in \mathbb{N} \setminus \{0\}$, ktorého platnosť sa očakáva pre všetky $t \in \mathbb{N}$ a v ktorom vystupuje ľubovoľné zobrazenie $F: \mathbb{N} \times (\mathbb{C}^n)^k \rightarrow \mathbb{C}^n$. Často je užitočné pracovať s jednotlivými zložkami $x_1(t), \dots, x_n(t)$ vektora

$$\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T,$$

pričom $x_1, \dots, x_n$ môžeme interpretovať ako neznáme funkcie z $\mathbb{N}$ do $\mathbb{C}$. V takom prípade hovoríme, že (6.8) je systémom n diferenčných rovníc o n neznámych funkciách $x_1, \dots, x_n$.

Začiatočnou úlohou danou systémom n diferenčných rovníc

$$\mathbf{x}(t+k) = F(t, \mathbf{x}(t), \mathbf{x}(t+1), \dots, \mathbf{x}(t+k-1)) \quad (6.9)$$

k -teho rádu o n neznámych funkciách a začiatočnými podmienkami

$$\mathbf{x}(0) = \mathbf{q}_0, \quad \mathbf{x}(1) = \mathbf{q}_1, \quad \dots, \quad \mathbf{x}(k-1) = \mathbf{q}_{k-1} \quad (6.10)$$

pre $\mathbf{q}_0, \dots, \mathbf{q}_{k-1} \in \mathbb{C}^n$ nazveme problém nájdenia funkcie $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$, ktorá po dosadení za $\mathbf{x}$ vyhovuje rovniciam (6.10) a pre všetky $t \in \mathbb{N}$ aj rovnici (6.9). Takúto vektorovú funkciu $\mathbf{f}$ nazveme *riešením* príslušnej začiatočnej úlohy.

Veta 6.2.4 (O existencii a jednoznačnosti pre systémy). *Uvažujme pre $k \in \mathbb{N}$ a $n \in \mathbb{N} \setminus \{0\}$ ľubovoľný explicitný systém n diferenčných rovníc*

$$\mathbf{x}(t+k) = F(t, \mathbf{x}(t), \mathbf{x}(t+1), \dots, \mathbf{x}(t+k-1)) \quad (6.11)$$

k -teho rádu o n neznámych funkciách a konštantné vektory $\mathbf{q}_0, \dots, \mathbf{q}_{k-1} \in \mathbb{C}^n$. Potom existuje práve jedno riešenie začiatočnej úlohy danej systémom (6.11) a začiatočnými podmienkami

$$\mathbf{x}(0) = \mathbf{q}_0, \quad \mathbf{x}(1) = \mathbf{q}_1, \quad \dots, \quad \mathbf{x}(k-1) = \mathbf{q}_{k-1}.$$

Dôkaz. Podobne ako pri vete 6.2.3 môžeme uvažovať funkciu $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ danú pre $t = 0, \dots, k-1$ ako $\mathbf{f}(t) = \mathbf{q}_t$ a pre všetky $t = k, k+1, \dots$ ako

$$\mathbf{f}(t) = F(t-k, \mathbf{f}(t-k), \mathbf{f}(t-k+1), \dots, \mathbf{f}(t-1)).$$

Vektorová funkcia $\mathbf{f}$ je evidentne riešením uvažovanej začiatočnej úlohy. Ak je naopak $\mathbf{g}: \mathbb{N} \rightarrow \mathbb{C}^n$ funkcia taká, že $\mathbf{g} \neq \mathbf{f}$, musí existovať $s \in \mathbb{N}$, pre ktoré je $\mathbf{g}(s) \neq \mathbf{f}(s)$; predpokladajme, že s je najmenšie možné. Ak $s \leq k-1$, tak $\mathbf{g}(s) \neq \mathbf{f}(s) = \mathbf{q}_s$ a $\mathbf{g}$ nevyhovuje začiatočným podmienkam; ak $s \geq k$, tak

$$\begin{aligned} \mathbf{g}((s-k)+k) &= \mathbf{g}(s) \neq \mathbf{f}(s) = F(s-k, \mathbf{f}(s-k), \mathbf{f}(s-k+1), \dots, \mathbf{f}(s-1)) = \\ &= F(s-k, \mathbf{g}(s-k), \mathbf{g}(s-k+1), \dots, \mathbf{g}(s-1)) \end{aligned}$$

a funkcia $\mathbf{g}$ nevyhovuje systému diferenčných rovníc (6.11). Vektorová funkcia $\mathbf{f}$ je teda naozaj jediným riešením uvažovanej začiatočnej úlohy. $\square$

Dôležitou špeciálnou triedou diferenčných rovníc a ich systémov, na ktorú sa zameriame vo zvyšku tejto kapitoly, je trieda tvorená *lineárnymi* diferenčnými rovnicami a ich systémami. Pri lineárnej diferenčnej rovnici k -teho rádu je závislosť hodnoty $x(t+k)$ neznámej funkcie x od $x(t), \dots, x(t+k-1)$ lineárna – každú takúto rovnicu teda možno zapísať ako

$$x(t+k) = c_{k-1}(t)x(t+k-1) + \dots + c_1(t)x(t+1) + c_0(t)x(t) + g(t),$$

kde $c_0, \dots, c_{k-1}, g: \mathbb{N} \rightarrow \mathbb{C}$ sú funkcie. Podobne pri systéme lineárnych diferenčných rovníc k -teho rádu s vektorom neznámych funkcií

$$\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$$

závisí každá zložka vektora $\mathbf{x}(t+k)$ lineárne od zložiek vektorov $\mathbf{x}(t), \dots, \mathbf{x}(t+k-1)$. Takýto systém tak možno vyjadriť v maticovom tvare

$$\mathbf{x}(t+k) = A_{k-1}(t)\mathbf{x}(t+k-1) + \dots + A_1(t)\mathbf{x}(t+1) + A_0(t)\mathbf{x}(t) + \mathbf{g}(t),$$

kde $A_0, \dots, A_{k-1}: \mathbb{N} \rightarrow \mathbb{C}^{n \times n}$ a $\mathbf{g}: \mathbb{N} \rightarrow \mathbb{C}^n$ sú funkcie.

6.3 Systémy lineárnych diferenčných rovníc prvého rádu

Zameriame sa teraz na veľmi špeciálnu triedu systémov diferenčných rovníc – na systémy *lineárnych* diferenčných rovníc *prvého rádu*. Každý takýto systém možno vyjadriť v tvare

$$\begin{aligned}x_1(t+1) &= a_{1,1}(t)x_1(t) + a_{1,2}(t)x_2(t) + \dots + a_{1,n}(t)x_n(t) + g_1(t), \\x_2(t+1) &= a_{2,1}(t)x_1(t) + a_{2,2}(t)x_2(t) + \dots + a_{2,n}(t)x_n(t) + g_2(t), \\&\vdots \\x_n(t+1) &= a_{n,1}(t)x_1(t) + a_{n,2}(t)x_2(t) + \dots + a_{n,n}(t)x_n(t) + g_n(t),\end{aligned}$$

kde $n \in \mathbb{N} \setminus \{0\}$, pre $i, j = 1, \dots, n$ sú $a_{i,j}: \mathbb{N} \rightarrow \mathbb{C}$ a $g_i: \mathbb{N} \rightarrow \mathbb{C}$ ľubovoľné funkcie a $x_1, \dots, x_n: \mathbb{N} \rightarrow \mathbb{C}$ sú neznáme funkcie. Ak pre všetky $t \in \mathbb{N}$ položíme $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T$, $A(t) = (a_{i,j}(t))_{n \times n}$ a $\mathbf{g}(t) = (g_1(t), \dots, g_n(t))^T$, môžeme systém prepísať do maticovo-vektorového tvaru ako

$$\mathbf{x}(t+1) = A(t)\mathbf{x}(t) + \mathbf{g}(t). \quad (6.12)$$

Pokiaľ je vektor $\mathbf{g}(t)$ nulový pre všetky $t \in \mathbb{N}$, nazýva sa systém (6.12) *homogénnym*; v opačnom prípade hovoríme o *nehomogénnom* systéme. Systém (6.12) navyše nazveme *autonómnym*, ak $\mathbf{x}(t+1)$ závisí len od $\mathbf{x}(t)$ a nie od t samotného – matica $A(t)$ ani vektor $\mathbf{g}(t)$ teda nezávisia od t a ľubovoľný autonómny systém lineárnych diferenčných rovníc prvého rádu tak možno písať v tvare

$$\mathbf{x}(t+1) = A\mathbf{x}(t) + \mathbf{g},$$

kde $\mathbf{x}(t)$ označuje pre všetky $t \in \mathbb{N}$ vektor neznámych funkcií, $A \in \mathbb{C}^{n \times n}$ a $\mathbf{g} \in \mathbb{C}^n$. Autonómnym systémom sa budeme detailnejšie venovať neskôr; nateraz ostane pri všeobecných systémoch lineárnych diferenčných rovníc prvého rádu daných ako v (6.12).

Začnime jednoduchými pozorovaniami o tvaroch riešení začiatočných úloh ako pre homogénne, tak aj pre nehomogénne systémy lineárnych diferenčných rovníc prvého rádu.

Veta 6.3.1. *Nech $n \in \mathbb{N} \setminus \{0\}$ a pre všetky $t \in \mathbb{N}$ je $A(t) \in \mathbb{C}^{n \times n}$. Jediným riešením začiatočnej úlohy danej homogénnym systémom n lineárnych diferenčných rovníc prvého rádu*

$$\mathbf{x}(t+1) = A(t)\mathbf{x}(t)$$

a začiatočnými podmienkami

$$\mathbf{x}(0) = \mathbf{q}_0$$

pre nejaký vektor $\mathbf{q}_0 \in \mathbb{C}^n$ je potom vektorová funkcia $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ daná pre všetky $t \in \mathbb{N}$ ako

$$\mathbf{f}(t) = \left(\prod_{j=0}^{t-1} A(j) \right) \mathbf{q}_0.$$

Dôkaz. Matematickou indukciou vzhľadom na t možno ľahko overiť, že $\mathbf{f}$ je skutočne riešením uvažovanej začiatočnej úlohy. Jeho jedinečnosť je pritom dôsledkom vety 6.2.4. $\square$

Veta 6.3.2. *Nech $n \in \mathbb{N} \setminus \{0\}$ a pre všetky $t \in \mathbb{N}$ je $A(t) \in \mathbb{C}^{n \times n}$ a $\mathbf{g}(t) \in \mathbb{C}^n$. Jediným riešením začiatočnej úlohy danej nehomogénnym systémom n lineárnych diferenčných rovníc prvého rádu*

$$\mathbf{x}(t+1) = A(t)\mathbf{x}(t) + \mathbf{g}(t)$$

a začiatočnými podmienkami

$$\mathbf{x}(0) = \mathbf{q}_0$$

pre nejaký vektor $\mathbf{q}_0 \in \mathbb{C}^n$ je potom vektorová funkcia $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ daná pre všetky $t \in \mathbb{N}$ ako

$$\mathbf{f}(t) = \left(\prod_{j=0}^{t-1} A(j) \right) \mathbf{q}_0 + \sum_{s=0}^{t-1} \left(\prod_{j=s+1}^{t-1} A(j) \right) \mathbf{g}(s).$$

Dôkaz. Opäť jednoduché cvičenie na matematickú indukciu. $\square$

Uvažujme teraz ľubovoľný *homogénny* systém $n \in \mathbb{N} \setminus \{0\}$ lineárnych diferenčných rovníc prvého rádu daný ako

$$\mathbf{x}(t+1) = A(t)\mathbf{x}(t). \quad (6.13)$$

Partikulárnym riešením systému (6.13) nazveme vektorovú funkciu $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$, ktorá po dosadení za $\mathbf{x}$ vzťahu (6.13) vyhovuje. Každé takéto partikulárne riešenie $\mathbf{f}$ je pritom tiež riešením začiatkovej úlohy danej systémom (6.13) a začiatkovými podmienkami $\mathbf{x}(0) = \mathbf{f}(0)$. Všetky partikulárne riešenia uvažovaného systému sú teda dané rovnako ako vo vete 6.3.1, pričom vektor $\mathbf{q}_0$ môže byť ľubovoľný.

Lema 6.3.3. *Nech $\mathbf{f}_1, \mathbf{f}_2: \mathbb{N} \rightarrow \mathbb{C}^n$ sú partikulárne riešenia systému (6.13) a $a, b \in \mathbb{C}$. Potom je partikulárnym riešením systému (6.13) aj vektorová funkcia $a\mathbf{f}_1 + b\mathbf{f}_2$.*

Dôkaz. Pre ľubovoľnú dvojicu partikulárných riešení $\mathbf{f}_1, \mathbf{f}_2: \mathbb{N} \rightarrow \mathbb{C}^n$ a všetky $t \in \mathbb{N}$ je

$$\begin{aligned} (a\mathbf{f}_1 + b\mathbf{f}_2)(t+1) &= a\mathbf{f}_1(t+1) + b\mathbf{f}_2(t+1) = aA(t)\mathbf{f}_1(t) + bA(t)\mathbf{f}_2(t) = A(t)(a\mathbf{f}_1(t) + b\mathbf{f}_2(t)) = \\ &= A(t)(a\mathbf{f}_1 + b\mathbf{f}_2)(t) \end{aligned}$$

a vektorová funkcia $a\mathbf{f}_1 + b\mathbf{f}_2$ tak vyhovuje systému (6.13). $\square$

V dôsledku práve dokázanej lemy tvoria partikulárne riešenia homogénneho systému (6.13) vektorový priestor. Každé z týchto riešení $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ je pritom podľa vety 6.3.1 pre nejaký vektor $\mathbf{q}_0 \in \mathbb{C}^n$ a pre všetky $t \in \mathbb{N}$ dané ako

$$\mathbf{f}(t) = \left(\prod_{j=0}^{t-1} A(j) \right) \mathbf{q}_0 =: \mathbf{f}[\mathbf{q}_0](t).$$

Uvažujme teraz niekoľko takýchto partikulárných riešení $\mathbf{f}[\mathbf{v}_1], \dots, \mathbf{f}[\mathbf{v}_m]$. Tie sú lineárne nezávislé práve vtedy, keď sú lineárne nezávislé vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$, pretože pre ľubovoľné $c_1, \dots, c_m \in \mathbb{C}$ a $t \in \mathbb{N}$ je

$$c_1\mathbf{f}[\mathbf{v}_1](t) + \dots + c_m\mathbf{f}[\mathbf{v}_m](t) = \left(\prod_{j=0}^{t-1} A(j) \right) (c_1\mathbf{v}_1 + \dots + c_m\mathbf{v}_m). \quad (6.14)$$

Ak sú teda vektory $\mathbf{v}_1, \dots, \mathbf{v}_m$ lineárne závislé, možno nájsť konštanty $c_1, \dots, c_m \in \mathbb{C}$ také, že aspoň jedna z nich je nenulová a výraz (6.14) je nulový pre všetky $t \in \mathbb{N}$. Ak sú naopak lineárne nezávislé, bude v prípade nenulovosti aspoň jednej z konštánt $c_1, \dots, c_m$ výraz (6.14) nenulový pre $t = 0$.

Zisťujeme teda, že *dimenzia vektorového priestoru partikulárných riešení homogénneho systému n diferenčných rovníc (6.13) je vždy rovná n .*

Ľubovoľnú množinu n lineárne nezávislých riešení $\mathbf{f}_1, \dots, \mathbf{f}_n: \mathbb{N} \rightarrow \mathbb{C}^n$ systému (6.13) nazveme jeho *fundamentálnou množinou riešení*. Z úvah učinенých vyššie bezprostredne vyplýva, že množina po dvoch rôznych riešení $\mathbf{f}_1, \dots, \mathbf{f}_n$ systému (6.13) je fundamentálna práve vtedy, keď sú lineárne nezávislé vektory $\mathbf{f}_1(0), \dots, \mathbf{f}_n(0)$. Pre ľubovoľnú fundamentálnu množinu riešení $\mathbf{f}_1, \dots, \mathbf{f}_n$ systému (6.13) navyše nazveme *všeobecným riešením* systému (6.13) lineárnu kombináciu

$$c_1\mathbf{f}_1(t) + \dots + c_n\mathbf{f}_n(t)$$

pre bližšie nešpecifikované koeficienty $c_1, \dots, c_n \in \mathbb{C}$. Vhodnou voľbou týchto konštánt možno získať práve všetky partikulárne riešenia systému (6.13).

Uvažujme napokon ľubovoľný *nehomogénny* systém $n \in \mathbb{N} \setminus \{0\}$ lineárnych diferenčných rovníc prvého rádu daný vzťahom

$$\mathbf{x}(t+1) = A(t)\mathbf{x}(t) + \mathbf{g}(t). \quad (6.15)$$

Podobne ako pre homogénne systémy nazveme *partikulárnym riešením* systému (6.15) ľubovoľnú vektorovú funkciu $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$, ktorá po dosadení za $\mathbf{x}$ systému (6.15) vyhovuje.

Veta 6.3.4. *Nech $\mathbf{f}_{\text{par}}: \mathbb{N} \rightarrow \mathbb{C}^n$ je pevne dané partikulárne riešenie nehomogénneho systému (6.15). Vektorová funkcia $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ je potom partikulárnym riešením systému (6.15) práve vtedy, keď existuje partikulárne riešenie $\mathbf{f}_{\text{hom}}: \mathbb{N} \rightarrow \mathbb{C}^n$ homogénneho systému (6.13) také, že pre všetky $t \in \mathbb{N}$ je*

$$\mathbf{f}(t) = \mathbf{f}_{\text{hom}}(t) + \mathbf{f}_{\text{par}}(t).$$

Dôkaz. Predpokladajme najprv, že vektorová funkcia $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ je partikulárnym riešením nehomogénneho systému (6.15) a pre všetky $t \in \mathbb{N}$ položme

$$\mathbf{f}_{\text{hom}}(t) := \mathbf{f}(t) - \mathbf{f}_{\text{par}}(t).$$

Takto definovaná funkcia $\mathbf{f}_{\text{hom}}: \mathbb{N} \rightarrow \mathbb{C}^n$ je partikulárnym riešením homogénneho systému (6.13), keďže pre všetky $t \in \mathbb{N}$ je

$$\begin{aligned} \mathbf{f}_{\text{hom}}(t+1) &= \mathbf{f}(t+1) - \mathbf{f}_{\text{par}}(t+1) = (A(t)\mathbf{f}(t) + \mathbf{g}(t)) - (A(t)\mathbf{f}_{\text{par}}(t) + \mathbf{g}(t)) = \\ &= A(t) \cdot (\mathbf{f}(t) - \mathbf{f}_{\text{par}}(t)) = A(t)\mathbf{f}_{\text{hom}}(t) \end{aligned}$$

vďaka tomu, že $\mathbf{f}$ a $\mathbf{f}_{\text{par}}$ sú partikulárnymi riešeniami nehomogénneho systému (6.15).

Predpokladajme naopak, že $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ je pre všetky $t \in \mathbb{N}$ daná ako

$$\mathbf{f}(t) = \mathbf{f}_{\text{hom}}(t) + \mathbf{f}_{\text{par}}(t),$$

kde $\mathbf{f}_{\text{hom}}: \mathbb{N} \rightarrow \mathbb{C}^n$ je nejaké partikulárne riešenie homogénneho systému (6.13). Pre všetky $t \in \mathbb{N}$ potom

$$\begin{aligned} \mathbf{f}(t+1) &= \mathbf{f}_{\text{hom}}(t+1) + \mathbf{f}_{\text{par}}(t+1) = A(t)\mathbf{f}_{\text{hom}}(t) + (A(t)\mathbf{f}_{\text{par}}(t) + \mathbf{g}(t)) = \\ &= A(t) \cdot (\mathbf{f}_{\text{hom}}(t) + \mathbf{f}_{\text{par}}(t)) + \mathbf{g}(t) = A(t)\mathbf{f}(t) + \mathbf{g}(t) \end{aligned}$$

a $\mathbf{f}$ je tak riešením nehomogénneho systému (6.15). □

Aj pre nehomogénny systém (6.15) tak môžeme definovať jeho *všeobecné riešenie* ako ľubovoľný výraz typu

$$c_1\mathbf{f}_1(t) + \dots + c_n\mathbf{f}_n(t) + \mathbf{f}_{\text{par}}(t),$$

kde $\mathbf{f}_1, \dots, \mathbf{f}_n$ tvoria fundamentálnu množinu riešení príslušného *homogénneho* systému (6.13), $\mathbf{f}_{\text{par}}$ je pevne dané partikulárne riešenie nehomogénneho systému (6.15) a $c_1, \dots, c_n \in \mathbb{C}$ sú bližšie nešpecifikované konštanty. Vhodnou voľbou koeficientov $c_1, \dots, c_n$ v jeho všeobecnom riešení môžeme získať práve všetky partikulárne riešenia systému (6.15).

6.4 Autonómne systémy

Ako špeciálny prípad systémov z predchádzajúceho oddielu teraz uvažujme *autonómne* systémy lineárnych diferenčných rovníc prvého rádu. Budeme sa zaoberať iba homogénnymi systémami – metódy opísané nižšie však bude z veľkej časti možné rozšíriť aj na nehomogénne systémy, pričom východiskom sú v takom prípade pozorovania z konca predchádzajúceho oddielu. Ľubovoľný systém uvažovaného typu tak možno zapísať ako

$$\mathbf{x}(t+1) = A\mathbf{x}(t), \tag{6.16}$$

kde $A \in \mathbb{C}^{n \times n}$ pre nejaké $n \in \mathbb{N} \setminus \{0\}$. Je pritom zrejmé, že všetky jeho partikulárne riešenia sú dané ako

$$\mathbf{f}[\mathbf{q}_0](t) = A^t\mathbf{q}_0$$

pre nejaký vektor $\mathbf{q}_0 \in \mathbb{C}^n$.

Zo zistení predchádzajúceho oddielu tak vyplýva, že všeobecné riešenie systému (6.16) môže byť pre ľubovoľných n lineárne nezávislých vektorov $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{C}^n$ dané ako

$$c_1 \mathbf{f}[\mathbf{v}_1](t) + \dots + c_n \mathbf{f}[\mathbf{v}_n](t) = c_1 A^t \mathbf{v}_1 + \dots + c_n A^t \mathbf{v}_n.$$

Ak je pritom matica systému A diagonalizovateľná, možno za vektory $\mathbf{v}_1, \dots, \mathbf{v}_n$ vziať ľubovoľnú bázu priestoru $\mathbb{C}^n$ zloženú z vlastných vektorov matice A a zisťujeme, že všeobecné riešenie systému (6.16) možno vyjadriť v tvare

$$c_1 \lambda_1^t \mathbf{v}_1 + \dots + c_n \lambda_n^t \mathbf{v}_n, \quad (6.17)$$

kde pre $k = 1, \dots, n$ je λ_k vlastným číslom matice A prislúchajúcim k $\mathbf{v}_k$. Pre ľubovoľný vektor $\mathbf{q}_0 \in \mathbb{C}^n$ navyše možno jediné riešenie začiatočnej úlohy danej systémom (6.16) a počiatočnými podmienkami $\mathbf{x}(0) = \mathbf{q}_0$ získať zo všeobecného riešenia (6.17) dopočítaním koeficientov $c_1, \dots, c_n$. Keďže pre $t = 0$ musí byť

$$c_1 \lambda_1^0 \mathbf{v}_1 + \dots + c_n \lambda_n^0 \mathbf{v}_n = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n = \mathbf{q}_0,$$

je hľadaný vektor koeficientov $(c_1, \dots, c_n)^T$ jedným riešením sústavy lineárnych rovníc

$$\begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \uparrow \\ \mathbf{q}_0 \\ \downarrow \end{pmatrix}$$

o neznámych $y_1, \dots, y_n$.

Príklad 6.4.1. Uvažujme homogénny autonómny systém dvoch lineárnych diferenčných rovníc prvého rádu daný pre všetky $t \in \mathbb{N}$ ako

$$\begin{aligned} x_1(t+1) &= x_1(t) - 2x_2(t), \\ x_2(t+1) &= -2x_1(t) + x_2(t), \end{aligned}$$

čiže

$$\mathbf{x}(t+1) = A\mathbf{x}(t) \quad (6.18)$$

pre maticu

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}.$$

Charakteristickým polynómom matice A je

$$\text{ch}_A(z) = \det(z\mathbf{I} - A) = \det \begin{pmatrix} z-1 & 2 \\ 2 & z-1 \end{pmatrix} = z^2 - 2z - 3 = (z-3)(z+1),$$

z čoho zisťujeme, že vlastnými číslami matice A sú $\lambda_1 = 3$ a $\lambda_2 = -1$. Keďže sú tieto vlastné čísla rôzne, je matica A diagonalizovateľná. Riešením príslušných sústav lineárnych rovníc navyše možno nájsť vlastné vektory prislúchajúce k týmto vlastným číslam – k vlastnému číslu λ_1 prislúcha napríklad vlastný vektor $\mathbf{v}_1 = (-1, 1)^T$ a k vlastnému číslu λ_2 napríklad vlastný vektor $\mathbf{v}_2 = (1, 1)^T$.

Všeobecné riešenie autonómneho systému (6.18) tak môže byť dané ako

$$c_1 \lambda_1^t \mathbf{v}_1 + c_2 \lambda_2^t \mathbf{v}_2 = c_1 3^t \begin{pmatrix} -1 \\ 1 \end{pmatrix} + c_2 (-1)^t \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Napríklad pre začiatočnú úlohu danú systémom (6.18) a začiatočnými podmienkami $\mathbf{x}(0) = (2, 0)^T$ teda môžeme vypočítať jej jediné riešenie tak, že v nájdenom všeobecnom riešení systému (6.18) dopočítame koeficienty c_1 a c_2 . Tie musia byť riešením sústavy lineárnych rovníc

$$\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

o neznámych y_1, y_2 . Zisťujeme teda, že $c_1 = -1$ a $c_2 = 1$. Jediným riešením začiatkovej úlohy danej systémom (6.18) a začiatkovými podmienkami $\mathbf{x}(0) = (2, 0)^T$ je preto

$$\mathbf{f}(t) = -3^t \begin{pmatrix} -1 \\ 1 \end{pmatrix} + (-1)^t \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

čiže $f_1(t) = 3^t + (-1)^t$ a $f_2(t) = -3^t + (-1)^t$.

Pokiaľ matica A systému (6.16) nie je diagonalizovateľná, možno za $\mathbf{v}_1, \dots, \mathbf{v}_n$ vziať vhodnú bázu zovšeobecnených vlastných vektorov matice A . Možno tak dospieť k vyjadreniu všeobecného riešenia systému (6.16), ktoré je podobné ako v (6.17), akurát v ňom namiesto vlastných vektorov vystupujú zovšeobecnené vlastné vektory matice A a z exponenciálnych faktorov λ^t pre vlastné čísla λ matice A sa stanú polynomiálno-exponenciálne funkcie.

Začiatočnú úlohu danú systémom (6.16) a vektorom začiatkových podmienok $\mathbf{q}_0 \in \mathbb{C}^n$ však možno vyriešiť aj bez prácného výpočtu zovšeobecnených vlastných vektorov matice A , a to s použitím podobných myšlienok ako pri enumerácii sledov v oddiele 5.2. Základom tejto metódy je nasledujúca veta, ktorá je v skutočnosti len veľmi jemným zovšeobením vety 5.2.1.

Veta 6.4.2. *Nech $\mathbf{f}: \mathbb{N} \rightarrow \mathbb{C}^n$ je jediné riešenie začiatkovej úlohy danej systémom (6.16) a vektorom začiatkových podmienok $\mathbf{q}_0 \in \mathbb{C}^n$. Pre všetky $t \in \mathbb{N}$ označme $\mathbf{f}(t) = (f_1(t), \dots, f_n(t))^T$. Pre $j = 1, \dots, n$ potom ku každému $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$ existuje konštanta $c_{j,\lambda,k} \in \mathbb{C}$ taká, že pre všetky $t \in \mathbb{N}$ je*

$$f_j(t) = \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} c_{j,\lambda,k} t^k \lambda^t + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} c_{j,\lambda,k} \delta_{t,k},$$

kde Kroneckerova delta je pre všetky $r, s \in \mathbb{N}$ daná ako

$$\delta_{r,s} = \begin{cases} 1 & \text{ak } r = s, \\ 0 & \text{ak } r \neq s. \end{cases}$$

Dôkaz. Podobne ako pri vete 5.2.1 ide o jednoduchý dôsledok existencie matice J v Jordanovom kanonickom tvare takej, že pre nejakú nesingulárnu maticu P je

$$A = PJP^{-1}$$

a pre všetky $t \in \mathbb{N}$ teda aj

$$\mathbf{f}(t) = A^t \mathbf{q}_0 = PJ^t P^{-1} \mathbf{q}_0.$$

Z charakteru mocnín matíc v Jordanovom kanonickom tvare totiž vyplýva, že každý prvok $q_{i,j}[t]$ matice $J^t = (q_{i,j}[t])_{n \times n}$ je buď nulový pre všetky $t \in \mathbb{N}$, alebo tvaru

$$q_{i,j}[t] = \binom{t}{k} \lambda^{t-k}$$

pre nejaké $\lambda \in \sigma(A)$ a $k \in \mathbb{N}$ spĺňajúce $0 \leq k \leq \alpha(A, \lambda) - 1$. Ľubovoľná zložka vektora $\mathbf{f}(t)$ je pritom lineárnou kombináciou prvkov takéhoto typu, z čoho už bezprostredne vyplýva možnosť vyjadriť hodnoty funkcií $f_1(t), \dots, f_n(t)$ ako v znení vety. $\square$

Riešenie začiatkovej úlohy danej systémom (6.16) a vektorom začiatkových podmienok $\mathbf{q}_0 \in \mathbb{C}^n$ teda možno ľahko vypočítať s použitím predchádzajúcej vety, ktorá poskytuje spoločné všeobecné vyjadrenie všetkých zložiek $f_1(t), \dots, f_n(t)$ jediného riešenia $\mathbf{f}(t)$. Pre každú z týchto zložiek $f_j(t)$ možno n neznámych koeficientov $c_{j,\lambda,k} \in \mathbb{C}$ pre $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$ dopočítať ako riešenie sústavy lineárnych rovníc. Priamym výpočtom vektorov $A^0 \mathbf{q}_0, A^1 \mathbf{q}_0, \dots, A^{n-1} \mathbf{q}_0$ totiž nie je

problém nájsť hodnoty $f_j(0), f_j(1), \dots, f_j(n-1)$. Hľadané koeficienty $c_{j,\lambda,k}$ sú potom riešením sústavy lineárnych rovníc

$$\begin{aligned} \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} y_{\lambda,k} 0^k \lambda^0 + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} y_{\lambda,k} \delta_{0,k} &= f_j(0), \\ \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} y_{\lambda,k} 1^k \lambda^1 + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} y_{\lambda,k} \delta_{1,k} &= f_j(1), \\ &\vdots \\ \sum_{\lambda \in \sigma(A) \setminus \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} y_{\lambda,k} (n-1)^k \lambda^{n-1} + \sum_{\lambda \in \sigma(A) \cap \{0\}} \sum_{k=0}^{\alpha(A,\lambda)-1} y_{\lambda,k} \delta_{n-1,k} &= f_j(n-1), \end{aligned}$$

kde neznámymi sú $y_{\lambda,k}$ pre $\lambda \in \sigma(A)$ a $k = 0, \dots, \alpha(A, \lambda) - 1$. Matica tejto sústavy je neregulárna podľa dôsledku 5.1.12 – hľadané koeficienty sú teda jej jediným riešením.

Táto metóda môže byť pomerne prácna v prípade, keď potrebujeme identifikovať všetky zložky jej riešenia $\mathbf{f}(t)$ – pre $j = 1, \dots, n$ je totiž potrebné hľadať koeficienty pre $f_j(t)$ zvlášť. Často nás však zaujíma iba jediná zložka riešenia – pre tieto účely je opísaná metóda naopak pomerne efektívna.

Príklad 6.4.3. Uvažujme začiatočnú úlohu danú homogénnym autonómnym systémom troch lineárnych diferenčných rovníc prvého rádu

$$\begin{aligned} x_1(t+1) &= 3x_1(t) + 2x_2(t) + x_1(t), \\ x_2(t+1) &= 7x_2(t) + 2x_3(t), \\ x_3(t+1) &= -x_1(t) - 12x_2(t) - 3x_3(t) \end{aligned}$$

a začiatočnými podmienkami $x_1(0) = 1$, $x_2(0) = 0$ a $x_3(0) = 2$. Ak teda pre všetky $t \in \mathbb{N}$ položíme $\mathbf{x}(t) = (x_1(t), x_2(t), x_3(t))^T$, možno uvažovaný systém zapísať ako

$$\mathbf{x}(t+1) = A\mathbf{x}(t)$$

pre maticu

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 7 & 2 \\ -1 & -12 & -3 \end{pmatrix}$$

a začiatočné podmienky možno zapísať ako

$$\mathbf{x}(0) = \mathbf{q}_0 = (1, 0, 2)^T.$$

Nájdeme prvú zložku $f_1(t)$ jediného riešenia $\mathbf{f}(t) = (f_1(t), f_2(t), f_3(t))^T$ tejto začiatočnej úlohy.

Lahko sa možno presvedčiť o tom, že charakteristickým polynómom matice A je

$$\text{ch}_A(z) = z^3 - 7z^2 + 16z - 12 = (z-3)(z-2)^2.$$

Vlastnými číslami matice A preto sú $\lambda_1 = 3$ s $\alpha(A, 3) = 1$ a $\lambda_2 = 2$ s $\alpha(A, 2) = 2$. Z vety 6.4.2 teda vyplýva, že pre nejaké komplexné konštanty $c_{1,3,0}$, $c_{1,2,0}$ a $c_{1,2,1}$ a všetky $t \in \mathbb{N}$ je

$$f_1(t) = c_{1,3,0}3^t + c_{1,2,1}t2^t + c_{1,2,0}2^t.$$

Zostáva dopočítať tieto neznáme koeficienty. Priamym výpočtom ľahko zistujeme, že $A^0\mathbf{q}_0 = (1, 0, 2)^T$, $A^1\mathbf{q}_0 = (5, 4, -7)^T$ a $A^2\mathbf{q}_0 = (16, 14, -32)^T$. Z toho $f_1(0) = 1$, $f_1(1) = 5$ a $f_1(2) = 16$. Koeficienty

$c_{1,3,0}$, $c_{1,2,0}$ a $c_{1,2,1}$ sú preto jediným riešením sústavy lineárnych rovníc

$$\begin{aligned} y_{1,3,0} + y_{1,2,0} &= 1, \\ 3y_{1,3,0} + 2y_{1,2,1} + 2y_{1,2,0} &= 5, \\ 9y_{1,3,0} + 8y_{1,2,1} + 4y_{1,2,0} &= 16 \end{aligned}$$

o neznámych $y_{1,3,0}$, $y_{1,2,0}$ a $y_{1,2,1}$. Riešením tejto sústavy zisťujeme, že

$$c_{1,3,0} = 0, \quad c_{1,2,1} = \frac{3}{2}, \quad c_{1,2,0} = 1$$

a pre všetky $t \in \mathbb{N}$ teda

$$f_1(t) = \frac{3}{2}t^2 + 2^t.$$

6.5 Lineárne diferenčné rovnice k -teho rádu

Budeme sa teraz zaoberať začiatočnými úlohami pre lineárne homogénne diferenčné rovnice k -teho rádu s konštantnými koeficientmi, čiže pre diferenčné rovnice dané ako

$$x(t+k) = c_{k-1}x(t+k-1) + \dots + c_0x(t)$$

pre všetky $t \in \mathbb{N}$, kde $k \in \mathbb{N}$ a $c_0, \dots, c_{k-1} \in \mathbb{C}$. Ukážeme si, ako možno riešenie začiatočných úloh pre takéto diferenčné rovnice previesť na riešenie začiatočných úloh pre vhodné homogénne autonómne systémy lineárnych diferenčných rovníc prvého rádu.

Každú diferenčnú rovnicu uvedeného typu možno písať aj ako

$$x(t+k) + d_{k-1}x(t+k-1) + \dots + d_0x(t) = 0 \quad (6.19)$$

pre všetky $t \in \mathbb{N}$, kde $k \in \mathbb{N}$ a $d_0, \dots, d_{k-1} \in \mathbb{C}$. *Charakteristickým polynómom* rovnice (6.19) nazveme polynóm

$$p(z) = z^k + d_{k-1}z^{k-1} + \dots + d_0z^0.$$

K ľubovoľnému polynómu takéhoto typu môžeme ďalej uvažovať jeho *sprievodnú maticu*, ktorou je štvorcová matica A_p typu $k \times k$ daná ako

$$A_p = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -d_0 & -d_1 & -d_2 & \dots & -d_{k-1} \end{pmatrix};$$

sprievodná matica polynómu $p(z)$ sa zvykne nazývať aj jeho *pridruženou maticou*.

Ľahko teraz vidieť, že v prípade platnosti vzťahu (6.19) musí pre všetky $t \in \mathbb{N}$ byť

$$\begin{pmatrix} x(t+1) \\ x(t+2) \\ \vdots \\ x(t+k) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -d_0 & -d_1 & -d_2 & \dots & -d_{k-1} \end{pmatrix} \begin{pmatrix} x(t) \\ x(t+1) \\ \vdots \\ x(t+k-1) \end{pmatrix},$$

čiže pre $\mathbf{x}(t) = (x(t), \dots, x(t+k-1))^T$ dostávame

$$\mathbf{x}(t+1) = A_p \mathbf{x}(t). \quad (6.20)$$

Ľubovoľnú začiatočnú úlohu danú diferenčnou rovnicou (6.19) spolu so začiatočnými podmienkami $x(0) = q_0, \dots, x(k-1) = q_{k-1}$ tak možno previesť na začiatočnú úlohu danú systémom (6.20) a začiatočnými podmienkami $\mathbf{x}(0) = (q_0, \dots, q_{k-1})^T$. Hľadaným riešením začiatočnej úlohy pre (6.19) je potom prvá zložka riešenia začiatočnej úlohy pre (6.20).

Už toto pozorovanie samo o sebe je základom metódy riešenia začiatočných úloh pre rovnicu (6.19). V skutočnosti je však situácia ešte oveľa jednoduchšia – charakteristický polynóm sprievodnej matice A_p polynómu $p(z)$ totiž spĺňa

$$\begin{aligned} \text{ch}_{A_p}(z) &= \det \begin{pmatrix} z & -1 & 0 & \dots & 0 & 0 \\ 0 & z & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 0 \\ 0 & 0 & 0 & \dots & z & -1 \\ d_0 & d_1 & d_2 & \dots & d_{k-2} & z + d_{k-1} \end{pmatrix} = \\ &= z \det \begin{pmatrix} z & -1 & \dots & 0 & 0 \\ 0 & z & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & z & -1 \\ d_1 & d_2 & \dots & d_{k-2} & z + d_{k-1} \end{pmatrix} + (-1)^{k+1} d_0 \det \begin{pmatrix} -1 & 0 & \dots & 0 & 0 \\ z & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & 0 \\ 0 & 0 & \dots & z & -1 \end{pmatrix}, \end{aligned}$$

kde pri výpočte determinantu sme použili Laplaceov rozvoj podľa prvého stĺpca. Evidentne pritom

$$\det \begin{pmatrix} -1 & 0 & \dots & 0 & 0 \\ z & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & 0 \\ 0 & 0 & \dots & z & -1 \end{pmatrix} = (-1)^{k-1},$$

takže pozorovaný vzťah možno prepísať ako

$$\text{ch}_{A_p}(z) = d_0 + z \det \begin{pmatrix} z & -1 & \dots & 0 & 0 \\ 0 & z & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & z & -1 \\ d_1 & d_2 & \dots & d_{k-2} & d_{k-1} + z \end{pmatrix}.$$

Indukciou tak ľahko dokážeme, že

$$\text{ch}_{A_p}(z) = z^k + d_{k-1}z^{k-1} + \dots + d_0z^0 = p(z).$$

Sprievodnú maticu A_p polynómu $p(z)$ tak možno chápať ako maticu „skonštruovanú na mieru“ polynómu $p(z)$ tak, aby bol polynóm $p(z)$ jej charakteristickým polynómom. Vlastnými číslami matice A_p sú teda práve korene polynómu $p(z)$, pričom algebraická násobnosť každého vlastného čísla λ je daná násobnosťou čísla λ ako koreňa polynómu $p(z)$.

Z vety 6.4.2 tak vyplýva priamočiary postup nájdenia jediného riešenia $f: \mathbb{N} \rightarrow \mathbb{C}$ začiatočnej úlohy danej diferenčnou rovnicou (6.19) a začiatočnými podmienkami $x(0) = q_0, \dots, x(k-1) = q_{k-1}$. Nech

$$p(z) = \prod_{\lambda \in \sigma(p)} (z - \lambda)^{\alpha(p, \lambda)}$$

je charakteristický polynóm diferenčnej rovnice (6.19), kde $\sigma(p) \subseteq \mathbb{C}$ je konečná množina a pre každé $\lambda \in \sigma(p)$ je $\alpha(p, \lambda) \in \mathbb{N} \setminus \{0\}$. Pre všetky $t \in \mathbb{N}$ potom

$$f(t) = \sum_{\lambda \in \sigma(p) \setminus \{0\}} \sum_{r=0}^{\alpha(p, \lambda)-1} c_{\lambda, r} t^r \lambda^t + \sum_{\lambda \in \sigma(p) \cap \{0\}} \sum_{r=0}^{\alpha(p, \lambda)-1} c_{\lambda, r} \delta_{t, r},$$

kde $c_{\lambda,r}$ pre $\lambda \in \sigma(p)$ a $r = 0, \dots, \alpha(p, \lambda) - 1$ sú komplexné konštanty nezávislé od t , ktoré možno dopočítať ako jediné riešenie sústavy lineárnych rovníc

$$\begin{aligned} \sum_{\lambda \in \sigma(p) \setminus \{0\}} \sum_{r=0}^{\alpha(p,\lambda)-1} y_{\lambda,r} 0^r \lambda^0 + \sum_{\lambda \in \sigma(p) \cap \{0\}} \sum_{r=0}^{\alpha(p,\lambda)-1} y_{\lambda,r} \delta_{0,r} &= q_0, \\ \sum_{\lambda \in \sigma(p) \setminus \{0\}} \sum_{r=0}^{\alpha(p,\lambda)-1} y_{\lambda,r} 1^r \lambda^1 + \sum_{\lambda \in \sigma(p) \cap \{0\}} \sum_{r=0}^{\alpha(p,\lambda)-1} y_{\lambda,r} \delta_{1,r} &= q_1, \\ &\vdots \\ \sum_{\lambda \in \sigma(p) \setminus \{0\}} \sum_{r=0}^{\alpha(p,\lambda)-1} y_{\lambda,r} (k-1)^r \lambda^{k-1} + \sum_{\lambda \in \sigma(p) \cap \{0\}} \sum_{r=0}^{\alpha(p,\lambda)-1} y_{\lambda,r} \delta_{k-1,r} &= q_{k-1} \end{aligned}$$

o neznámych $y_{\lambda,r}$ pre $\lambda \in \sigma(p)$ a $r = 0, \dots, \alpha(p, \lambda) - 1$.

Poznámka 6.5.1. Jedinečnosť riešenia sústavy lineárnych rovníc tu vyplýva nielen z dôsledku 5.1.12, ale aj z vety 6.2.3 o existencii a jednoznačnosti riešení začiatočných úloh pre diferenčné rovnice. Podľa nej totiž musí mať uvedená sústava riešenie *pre všetky* vektory začiatočných podmienok $(q_0, \dots, q_{k-1})^T$, z čoho vyplýva, že matica sústavy – ktorou je práve matica z dôsledku 5.1.12 – musí byť neregulárna. Takto teda prichádzame k dôsledku 5.1.12 aj bez toho, aby sme museli využívať vzorec pre determinant zovšeobecnenej Vandermondovej matice. Keby sme navyše pri vyjadrení riešenia začiatočnej úlohy pracovali s funkciami tvaru $\binom{t}{k} \lambda^{t-k}$, mohli by sme rovnakým spôsobom dokázať aj neregulárnosť zovšeobecnenej Vandermondovej matice nad poľom komplexných čísel⁴ – opäť bez toho, aby sme čokoľvek vedeli o jej determinante. Špeciálnym prípadom oboch pozorovaní je, samozrejme, aj neregulárnosť Vandermondovej matice nad $\mathbb{C}$.

Príklad 6.5.2. V rámci príkladu 2.2.1 sme odvodili uzavretý tvar pre t -te Fibonacciho číslo pomocou vytvárajúcich funkcií. Teraz tento výsledok odvodíme metódou opísanou vyššie.

Postupnosť Fibonacciho čísel $(F_t)_{t \in \mathbb{N}}$ možno charakterizovať ako jediné riešenie začiatočnej úlohy danej diferenčnou rovnicou

$$x(t+2) = x(t+1) + x(t)$$

pre všetky $t \in \mathbb{N}$ a začiatočnými podmienkami $x(0) = 0$ a $x(1) = 1$. Uvedenú diferenčnú rovnicu možno prepísať aj ako

$$x(t+2) - x(t+1) - x(t) = 0$$

a jej charakteristickým polynómom tak je

$$p(z) = z^2 - z - 1 = (z - \varphi)(z - \psi),$$

kde

$$\varphi = \frac{1 + \sqrt{5}}{2} \quad \text{a} \quad \psi = \frac{1 - \sqrt{5}}{2}.$$

Existujú preto konštanty $c_{\varphi,0}, c_{\psi,0} \in \mathbb{C}$ také, že pre všetky $t \in \mathbb{N}$ je

$$F_t = c_{\varphi,0} \varphi^t + c_{\psi,0} \psi^t.$$

Koeficienty $c_{\varphi,0}, c_{\psi,0}$ sú pritom jediným riešením sústavy lineárnych rovníc

$$\begin{aligned} y_{\varphi,0} + y_{\psi,0} &= 0, \\ \varphi y_{\varphi,0} + \psi y_{\psi,0} &= 1 \end{aligned}$$

⁴V skutočnosti je pri troche opatrnosti možné teóriu systémov lineárnych diferenčných rovníc vybudovať nad ľubovoľným poľom a uvedeným spôsobom tak dokázať neregulárnosť zovšeobecnených Vandermondových matíc aj v tomto všeobecnejšom rámci.

a po jej vyriešení zistujeme, že

$$c_{\varphi,0} = \frac{1}{\sqrt{5}} \quad \text{a} \quad c_{\psi,0} = -\frac{1}{\sqrt{5}}.$$

Pre všetky $t \in \mathbb{N}$ preto

$$F_t = \frac{1}{\sqrt{5}}\varphi^t - \frac{1}{\sqrt{5}}\psi^t = \frac{1}{\sqrt{5}}(\varphi^t - \psi^t),$$

čo sa zhoduje s výsledkom získaným metódou vytvárajúcich funkcií v rámci príkladu 2.2.1.

Príklad 6.5.3. Nájdeme jediné riešenie začiatočnej úlohy danej diferenčnou rovnicou

$$x(t+3) - 5x(t+2) + 8x(t+1) - 4x(t) = 0$$

pre všetky $t \in \mathbb{N}$ a počiatočnými podmienkami $x(0) = 0$, $x(1) = 1$ a $x(2) = 2$. Charakteristickým polynómom uvažovanej diferenčnej rovnice je

$$p(z) = z^3 - 5z^2 + 8z - 4 = (z-2)^2(z-1).$$

Existujú preto konštanty $c_{2,0}, c_{2,1}, c_{1,0} \in \mathbb{C}$ také, že jediné riešenie $f: \mathbb{N} \rightarrow \mathbb{C}$ uvažovanej začiatočnej úlohy je pre všetky $t \in \mathbb{N}$ dané ako

$$f(t) = c_{2,1}t2^t + c_{2,0}2^t + c_{1,0}.$$

Koeficienty $c_{2,1}, c_{2,0}, c_{1,0}$ nájdeme ako jediné riešenie sústavy lineárnych rovníc

$$\begin{aligned} y_{2,0} + y_{1,0} &= 0, \\ 2y_{2,1} + 2y_{2,0} + y_{1,0} &= 1, \\ 8y_{2,1} + 4y_{2,0} + y_{1,0} &= 2. \end{aligned}$$

Zisťujeme tak, že $c_{2,1} = -1/2$, $c_{2,0} = 2$ a $c_{1,0} = -2$. Pre všetky $t \in \mathbb{N}$ preto

$$f(t) = -t2^{t-1} + 2^{t+1} - 2.$$

Cvičenia

1. Dokážte nasledujúci variant vzorca pre konečnú diferenciu podielu: pre ľubovoľnú dvojicu funkcií $f: \mathbb{N} \rightarrow \mathbb{C}$ a $g: \mathbb{N} \rightarrow \mathbb{C} \setminus \{0\}$ je

$$\Delta \left(\frac{f}{g} \right) = \frac{(\Delta f)(Eg) - (Ef)(\Delta g)}{g(Eg)}.$$

2. Nech $a \in \mathbb{C} \setminus \{0\}$. Nájdite vzorce pre konečné diferencie nasledujúcich funkcií prirodzenej premennej t : 2^t , a^t , $\sin at$, $\cos at$.
3. Podrobne opíšte metódu nájdania všeobecného riešenia homogénneho autonómneho systému lineárnych diferenčných rovníc prvého rádu

$$\mathbf{x}(t+1) = A\mathbf{x}(t)$$

v prípade, že matica A nie je diagonalizovateľná.

4. Nájdite riešenie začiatočnej úlohy danej systémom lineárnych diferenčných rovníc

$$\mathbf{x}(t+1) = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \mathbf{x}(t) \quad \text{pre všetky } t \in \mathbb{N}$$

a vektorom začiatočných podmienok $\mathbf{x}(0) = (1, -1, 1)^T$.

5. Nájdite riešenia nasledujúcich začiatočných úloh pre lineárne diferenčné rovnice:

a) $x(t+3) = 3x(t+1) + 2x(t)$ pre všetky $t \in \mathbb{N}$, pričom $x(0) = 0$, $x(1) = 1$ a $x(2) = 1$.

b) $x(t+3) = x(t+2) + 2x(t+1)$ pre všetky $t \in \mathbb{N}$, pričom $x(0) = 2$, $x(1) = 1$ a $x(2) = 1$.

6. Uvažujme polynóm $p(z) = z^k + d_{k-1}z^{k-1} + \dots + d_0z^0$ s komplexnými koeficientmi $d_0, \dots, d_{k-1} \in \mathbb{C}$ a príslušnú sprievodnú maticu A_p .

a) Určte geometrické násobnosti všetkých vlastných čísel matice A_p .

b) Charakterizujte polynómy $p(z)$, pre ktoré je matica A_p diagonalizovateľná.

c) Charakterizujte polynómy $p(z)$, pre ktoré je matica A_p nesingulárna.

Poznámky a odkazy na samoštúdium

Spomedzi učebníc zameraných na oblasť diferenčných rovníc možno odporúčať predovšetkým [62, 34]; ďalšími relevantnými úvodnými textmi sú [58, 85]. O pokročilejších aspektoch diferenčných rovníc sa možno dočítať v monografii [1]; výbornou algebraickejšie ladenou monografiou o rekurentných postupnostiach je tiež [41].

Dejiny diferenčného počtu siahajú ešte do doby pred Newtonom a Leibnizom [96] a pozornosť sa tejto oblasti venovala aj neskôr – autorom klasickej monografie o diferenčnom počte z roku 1860 je napríklad aj George Boole [15]. O diferenčnom počte sa možno dočítať vo väčšine učebníc diferenčných rovníc alebo napríklad v [51], kde sa preň používa názov *konečný kalkul* resp. *konečný počet* (angl. *finite calculus*). Ďalšou knihou o diferenčnom počte je [59].

Očividná podobnosť mnohých aspektov diferenčného a diferenciálneho počtu dala vzniknúť aj teórii, ktorá obidva tieto počty zjednocuje – tá je známa ako *počet časových škál* (angl. *time-scale calculus*). Takzvané *dynamické rovnice* (angl. *dynamic equations*) na časových škálach umožňujú študovať diferenčné a diferenciálne rovnice v spoločnom rámci. Dočítať sa o nich možno v [14].

Kapitola 7

Vlastné čísla a grafy II

Stretli sme sa už s niekoľkými aplikáciami vlastných čísel matíc susednosti orientovaných grafov. V rámci tejto záverečnej kapitoly naznačíme niektoré možnosti využitia vlastných čísel pri analýze štrukturálnych vlastností *neorientovaných* grafov. Tie sa ukazujú ako podstatne bohatšie, než tomu bolo pri orientovaných grafoch: do veľkej miery ide o dôsledok skutočnosti, že matice prislúchajúce neorientovaným grafom sú typicky symetrické – špeciálne to tak je pre *matice susednosti* neorientovaných grafov, ako aj pre takzvané *Laplaceove matice* neorientovaných grafov, ktoré si v rámci tejto kapitoly takisto predstavíme. Oblasť výskumu, na pôde ktorej sa v spojitosti s takýmito úvahami ocitneme, je známa ako *spektrálna teória grafov*. Po ceste dokážeme aj viacero ďalších výsledkov z klasickej lineárnej algebry a maticového počtu, súvisiacich so symetrickými maticami.

Obmedzíme sa len na vyslovené základy vybraných partií spektrálnej teórie grafov a na ochutnávku niektorých jej typických metód. Nie je úmyslom tejto kapitoly priblížiť sa k ucelenému výkladu najvýznamnejších výsledkov spektrálnej teórie grafov a už tobôž nie podať vyčerpávajúci prehľad množstva rôznorodých smerov výskumu v tejto oblasti. Čitateľovi so záujmom o hlbšie alebo širšie štúdium spektrálnej teórie grafov a príbuzných oblastí je k dispozícii relatívne veľké množstvo špecializovanej literatúry; odkazy na vybrané zdroje možno nájsť v poznámkach za touto kapitolou.

7.1 Neorientované grafy

Podobne ako pri orientovaných grafoch zvolíme jemne atypickú formalizáciu tradičnej definície neorientovaných grafov, pri ktorej budeme množinu hrán *de facto* reprezentovať jej charakteristickou funkciou. To by opäť bolo možné využiť na zavedenie pojmu *ohodnotených* neorientovaných grafov. Keďže sa ale v rámci spektrálnej teórie grafov vlastnosťami ohodnotených grafov zaoberať nebudeme, obmedzíme sa len na definíciu pojmu neorientovaných *multigrafov*, ktoré môžeme chápať aj ako ohodnotené grafy nad $\mathbb{N}$; multigrafy sa nám zídu v oddiele 7.8 ako pomocný nástroj.

Definícia 7.1.1. *Neorientovaný graf* je dvojica $\mathcal{G} = (V, \varphi)$, kde V je ľubovoľná konečná množina a $\varphi: \binom{V}{2} \cup \binom{V}{1} \rightarrow \mathbb{B}$ je zobrazenie.

Pre každý takýto graf podobne ako pri orientovaných grafoch píšeme $V(\mathcal{G}) := V$ a $E(\mathcal{G}) := \text{supp}(\varphi)$. Prvky množiny $V(\mathcal{G})$ pritom nazývame *vrcholmi* grafu $\mathcal{G}$ a prvky množiny $E(\mathcal{G})$, ktorými sú vybrané jednoprvkové alebo dvojprvkové množiny vrcholov $e = \{u, v\}$, nazývame *hranami* grafu $\mathcal{G}$. Hranu $e = \{u, v\} \in E(\mathcal{G})$ také, že $u = v$, nazývame *slučkami*. *Stupňom* $\deg_{\mathcal{G}}(v)$ vrcholu $v \in V(\mathcal{G})$ nazveme počet hrán $e \in E(\mathcal{G})$ takých, že $v \in e$, pričom prípadnú slučku započítame dvakrát:

$$\deg_{\mathcal{G}}(v) = |\{e \in E(\mathcal{G}) \mid v \in e\}| + |\{e \in E(\mathcal{G}) \mid e = \{v\}\}|.$$

Keďže sa v nasledujúcom budeme zaoberať maticami prislúchajúcimi ku grafom, budeme väčšinou bez ujmy na všeobecnosti predpokladať, že $V(\mathcal{G}) = [n]$ pre nejaké prirodzené číslo n . Matice grafov líšiacich sa len očíslovaním vrcholov budú vždy podobné – pri ich spektrách teda pôjde o invarianty nezávislé od voľby tohto očíslovania.

Pri definícii matice susednosti neorientovaného grafu, ktorá je inak priamočiaraou obdobou definície pre orientované grafy, existuje v literatúre určitá nekonzistencia ohľadom diagonálnych prvkov – tie sa v prípade existencie slučky niekedy definujú ako 1 a inokedy ako 2. S ohľadom na definíciu stupňa vrcholu, pri ktorej sú slučky započítané dvakrát, a ktorá je motivovaná predovšetkým platnosťou vzťahu

$$\sum_{v \in V(\mathcal{G})} \deg_{\mathcal{G}}(v) = 2 |E(\mathcal{G})|,$$

sa druhá alternatíva javí ako o niečo vhodnejšia. Na druhej strane by sme však pri takejto definícii prišli o možnosť chápať maticu susednosti neorientovaného grafu so slučkami ako booleovskú maticu, čo sa v kapitole 1 pri orientovaných grafoch ukázalo ako relatívne užitočné. Zvolíme teda nasledujúci prístup: maticu susednosti neorientovaného grafu definujeme tak, aby vždy išlo o maticu núl a jednotiek – aj v prípade existencie slučky v niektorom vrchole tak bude príslušný diagonálny prvok matice susednosti rovný jednej. Vzápätí ale definujeme pojem neorientovaného multigrafu a každý neorientovaný graf bude možné považovať aj za špeciálny neorientovaný multigraf. Maticou susednosti multigrafu bude matica prirodzených čísel – tú pritom definujeme tak, aby sa všetky prípadné slučky do príslušných diagonálnych prvkov matice započítali dvakrát. Namiesto úvah o tom, ktorá definícia matice susednosti je vhodnejšia, tak budeme mať na výber: ku každému neorientovanému grafu budeme môcť uvažovať ako booleovskú maticu susednosti daného grafu, tak aj maticu susednosti príslušného multigrafu pozostávajúcu z prirodzených čísel.

Tieto detaily sú však pre samotnú spektrálnu teóriu grafov okrajového významu – *pováčšine sa totiž budeme zaoberať neorientovanými grafmi bez slučiek.*

Definícia 7.1.2. Nech $n \in \mathbb{N}$ a $\mathcal{G} = ([n], \varphi)$ je neorientovaný graf. Maticou susednosti grafu $\mathcal{G}$ nazveme štvorcovú maticu $\mathbf{A}(\mathcal{G}) \in \mathbb{B}^{n \times n}$ danú ako $\mathbf{A}(\mathcal{G}) = (\varphi(\{i, j\}))_{n \times n}$.

Ak je teda $\mathcal{G} = ([n], \varphi)$ neorientovaný graf a $\mathbf{A}(\mathcal{G}) = (a_{i,j})_{n \times n}$, tak $a_{i,j} = 1$ práve vtedy, keď $\{i, j\} \in E(\mathcal{G})$. Matica susednosti $\mathbf{A}(\mathcal{G})$ je pre každý neorientovaný graf $\mathcal{G}$ evidentne symetrická a v spektrálnej teórii grafov ju obvykle chápeme ako maticu prirodzených čísel 0 a 1.

Hovoríme, že neorientovaný graf $\mathcal{G} = (V, \varphi)$ je *neorientovaný graf bez slučiek*, ak $E(\mathcal{G}) \subseteq \binom{V}{2}$. V takom prípade chápeme φ aj ako zobrazenie typu $\varphi: \binom{V}{2} \rightarrow \mathbb{B}$. Pod grafom (bez ďalšieho prívlastku) budeme po zvyšok tejto kapitoly vždy rozumieť neorientovaný graf bez slučiek. Pre takéto grafy $\mathcal{G}$ sú všetky diagonálne prvky matice $\mathbf{A}(\mathcal{G})$ vždy nulové.

Zavedme ešte už spomínaný pojem *neorientovaného multigrafu*, pod ktorým budeme rozumieť neorientovaný graf s možnosťou existencie násobných hrán medzi jednotlivými vrcholmi, ako aj násobných slučiek. V podstate tak pôjde o ohodnotené neorientované grafy nad $\mathbb{N}$, kde ohodnotenie každej z hrán udáva jej násobnosť; niektoré z definícií súvisiacich s neorientovanými multigrafmi sa ale budú od definícií pre ohodnotené grafy trochu odlišovať.

Definícia 7.1.3. Neorientovaný multigraf je dvojica $\mathcal{G} = (V, \varphi)$, kde V je ľubovoľná konečná množina a $\varphi: \binom{V}{2} \cup \binom{V}{1} \rightarrow \mathbb{N}$ je zobrazenie.

Podobne ako vyššie kladieme $V(\mathcal{G}) := V$ a prvky tejto množiny nazývame *vrcholmi* multigrafu $\mathcal{G}$. Na rozdiel od ohodnotených orientovaných grafov z kapitoly 1 však teraz nebudeme pracovať s hranami ohodnotenými prirodzenými číslami, ale medzi ľubovoľnou dvojicou vrcholov $u, v \in V(\mathcal{G})$ budeme uvažovať práve $\varphi(\{u, v\})$ rôznych hrán. Množinu hrán $E(\mathcal{G})$ multigrafu $\mathcal{G}$ tak definujeme ako

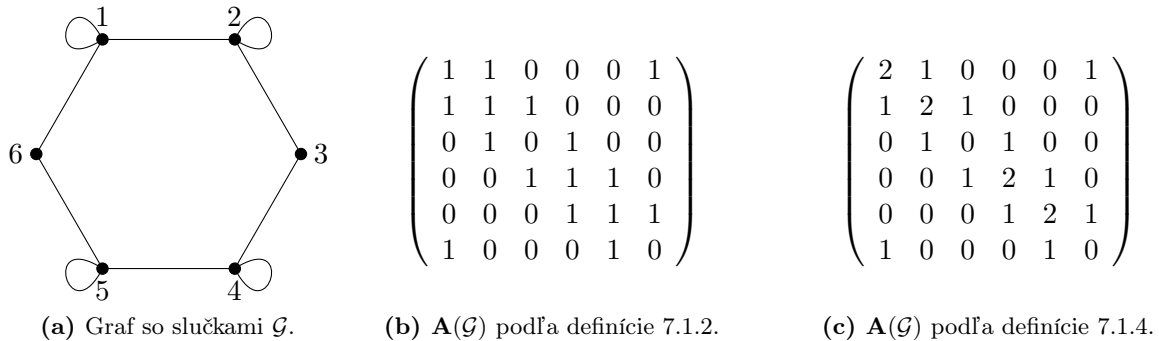
$$E(\mathcal{G}) = \{(\{u, v\}, k) \mid \{u, v\} \in \text{supp}(\varphi); k \in [\varphi(\{u, v\})]\}.$$

Definícia 7.1.4. Nech $n \in \mathbb{N}$ a $\mathcal{G} = ([n], \varphi)$ je neorientovaný multigraf. Maticou susednosti multigrafu $\mathcal{G}$ nazveme štvorcovú maticu $\mathbf{A}(\mathcal{G}) = (a_{i,j})_{n \times n} \in \mathbb{N}^{n \times n}$, kde $a_{i,j}$ je pre $i, j = 1, \dots, n$ dané ako

$$a_{i,j} = \begin{cases} \varphi(\{i, j\}) & \text{ak } i \neq j, \\ 2 \cdot \varphi(\{i, j\}) & \text{ak } i = j. \end{cases}$$

Každý neorientovaný graf – či už so slučkami alebo bez nich – môžeme s istou dávkou nepresnosti považovať aj za neorientovaný multigraf s násobnosťami všetkých hrán rovnými jednej. Pre daný neorientovaný graf $\mathcal{G}$ tak môže $\mathbf{A}(\mathcal{G})$ označovať dve rôzne matice susednosti;¹ v tejto kapitole ale obvyčajne budeme pracovať s grafmi bez slučiek a rozdiel medzi obidvoma definíciami matice $\mathbf{A}(\mathcal{G})$ sa tak stane nepodstatným.

Príklad 7.1.5. Na obrázku 7.1 je znázornený diagram neorientovaného grafu so slučkami $\mathcal{G}$ spoločne s jeho grafovou a multigrafovou maticou susednosti.



Obr. 7.1: Neorientovaný graf so slučkami $\mathcal{G}$ a jeho matice susednosti podľa definície pre grafy a pre multigrafy.

V rámci tejto kapitoly budeme používať štandardnú terminológiu teórie grafov, ktorá by predpokladanému čitateľovi mala byť dostatočne známa – na rozdiel od samotných definícií grafov a ich matic susednosti, ku ktorým sme pristupovali možno jemne nezvyčajným spôsobom, tu tak neexistuje podstatnejší dôvod na to, aby sme ju na tomto mieste pripomínali. Čitateľ so záujmom o jej osvieženie sa môže obrátiť na štandardnú literatúru o teórii grafov, napríklad na [28, 88, 27].

7.2 Spektrálne vlastnosti reálnych symetrických matic

Spektrálna teória grafov sa zaoberá predovšetkým vlastnými číslami matic prislúchajúcich ku grafom a ich súvisom so štrukturálnymi vlastnosťami grafov. Týmito maticami môžu byť *matice susednosti* grafov, ale napríklad aj takzvané *Laplaceove matice* grafov, ktoré zavedieme neskôr. V obidvoch prípadoch ide o *symetrické matice* reálnych čísel. V nasledujúcom preto dokážeme dve kľúčové vlastnosti spektier reálnych symetrických matic, o ktoré sa neskôr budeme opierať: vlastné čísla takýchto matic sú vždy reálne a každá takáto matica je ortogonálne diagonalizovateľná.² Prezentácia tohto oddielu je sčasti prevzatá z knihy [49].

Pre ľubovoľnú dvojicu vektorov $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, kde $n \in \mathbb{N}$, $\mathbf{x} = (x_1, \dots, x_n)^T$ a $\mathbf{y} = (y_1, \dots, y_n)^T$, reprezentuje hodnota

$$\mathbf{x}^T \mathbf{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

štandardný *skalárny súčin* vektorov $\mathbf{x}$ a $\mathbf{y}$ v euklidovskom priestore $\mathbb{R}^n$. *Euklidovská norma* (alebo dĺžka) ľubovoľného vektora $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ je pritom daná ako odmocnina skalárneho súčinu vektora $\mathbf{x}$ samého so sebou:

$$\|\mathbf{x}\|_2 := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \sqrt{\mathbf{x}^T \mathbf{x}}.$$

Vektory $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ sú navzájom kolmé – alebo *ortogonálne* – práve vtedy, keď $\mathbf{x}^T \mathbf{y} = 0$.

¹Všimnime si, že keby sme pre neorientovaný graf $\mathcal{G}$ v multigrafovej definícii 7.1.4 namiesto nad monoidom $(\mathbb{N}, +, 0)$ pracovali nad monoidom $(\mathbb{B}, \vee, 0)$, dostali by sme práve grafovú definíciu 7.1.2.

²Budeme hovoriť, že matica typu $n \times n$ je *ortogonálne diagonalizovateľná*, ak je diagonalizovateľná a súčasne existuje *ortonormálna báza* priestoru $\mathbb{R}^n$ zložená z jej vlastných vektorov. Pojem ortonormálnej bázy formálne zavedieme neskôr – pôjde o bázu zloženú z navzájom kolmých vektorov jednotkovej dĺžky.

V rámci prípravy na dôkaz reálnosti vlastných čísel reálnych symetrických matíc najprv dokážeme, že pre vlastné vektory $\mathbf{v}_1, \mathbf{v}_2$ prislúchajúce k dvom rôznym vlastným číslam reálnej symetrickej matice je vždy $\mathbf{v}_1^T \mathbf{v}_2 = 0$. Ak sú teda tieto vektory reálne, sú navzájom kolmé – pre nás však bude zaujímavý aj prípad, keď sú tieto vektory komplexné.

Lema 7.2.1. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. Nech λ_1, λ_2 sú rôzne vlastné čísla matice A a $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^n$ sú vlastné vektory matice A prislúchajúce k λ_1 resp. k λ_2 . Potom $\mathbf{v}_1^T \mathbf{v}_2 = 0$.*

Dôkaz. Je $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ a $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$. Zo symetrickosti matice A tak zrejme vyplýva

$$\lambda_2 \mathbf{v}_1^T \mathbf{v}_2 = \mathbf{v}_1^T A \mathbf{v}_2 = \mathbf{v}_1^T A^T \mathbf{v}_2 = (\mathbf{v}_2^T A \mathbf{v}_1)^T = (\lambda_1 \mathbf{v}_2^T \mathbf{v}_1)^T = \lambda_1 \mathbf{v}_1^T \mathbf{v}_2.$$

Pretože ale $\lambda_1 \neq \lambda_2$, môže táto rovnosť platiť len v prípade, že $\mathbf{v}_1^T \mathbf{v}_2 = 0$. $\square$

Pripomeňme si, že pre každé komplexné číslo $z = a + bi$ je jeho *komplexne združené číslo* $\bar{z}$ definované ako $\bar{z} = a - bi$. Pre vektory komplexných čísel $\mathbf{z} = (z_1, \dots, z_n)^T \in \mathbb{C}^n$ budeme túto notáciu interpretovať po zložkách: budeme teda písať $\bar{\mathbf{z}} = (\bar{z}_1, \dots, \bar{z}_n)^T$. Ak je $z \in \mathbb{C}$ dané ako $z = re^{i\theta}$ pre nejaké $r \geq 0$ a $\theta \in \mathbb{R}$, ľahko vidieť, že $\bar{z} = re^{-i\theta}$. Z toho možno okamžite odvodiť niekoľko vlastností komplexne združených čísel: pre ľubovoľné $z_1, z_2 \in \mathbb{C}$ je $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$; pre všetky $z \in \mathbb{C}$ ďalej $z \bar{z} = |z|^2$, z čoho bezprostredne vyplýva, že pre všetky $\mathbf{z} \in \mathbb{C}^n$ je hodnota $\mathbf{z}^T \bar{\mathbf{z}}$ nezáporná reálna; táto hodnota je navyše kladná, kedykoľvek je vektor $\mathbf{z}$ nenulový. Ak je napokon $\mathbf{z} \in \mathbb{C}^n$ a $A \in \mathbb{R}^{n \times n}$, môžeme vektor $\mathbf{z}$ pre nejakú dvojicu reálnych vektorov $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vyjadriť ako $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, z čoho $\bar{\mathbf{z}} = \mathbf{x} - i\mathbf{y}$. Potom $A\mathbf{z} = A\mathbf{x} + iA\mathbf{y}$ a $A\bar{\mathbf{z}} = A\mathbf{x} - iA\mathbf{y}$, kde vektory $A\mathbf{x}$ a $A\mathbf{y}$ sú reálne – musí teda byť $A\bar{\mathbf{z}} = \overline{A\mathbf{z}}$. Tieto vlastnosti komplexne združených čísel využijeme pri dôkaze nasledujúcej vety.

Veta 7.2.2. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. Potom $\sigma(A) \subseteq \mathbb{R}$ – všetky vlastné čísla matice A sú teda reálne. Ku každému vlastnému číslu λ matice A navyše prislúcha nejaký reálny vlastný vektor $\mathbf{w} \in \mathbb{R}^n$.*

Dôkaz. Nech je λ vlastné číslo matice A s prislúchajúcim vlastným vektorom $\mathbf{v}$. Potom $A\mathbf{v} = \lambda\mathbf{v}$, z čoho

$$A\bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}. \quad (7.1)$$

Za účelom sporu teraz predpokladajme, že $\lambda \in \mathbb{C} \setminus \mathbb{R}$. Vďaka (7.1) sú potom λ a $\bar{\lambda}$ dve rôzne vlastné čísla matice A . Podľa lemy 7.2.1 tak pre prislúchajúce vlastné vektory $\mathbf{v}$ a $\bar{\mathbf{v}}$ musí byť

$$\mathbf{v}^T \bar{\mathbf{v}} = 0.$$

Vlastný vektor $\mathbf{v}$ je však z definície nenulový, v dôsledku čoho je $\mathbf{v}^T \bar{\mathbf{v}}$ kladné reálne číslo: spor.

Na dôkaz druhej časti tvrdenia uvažujme reálne vlastné číslo λ matice A spolu s prislúchajúcim vlastným vektorom $\mathbf{v} \in \mathbb{C}^n$. Podľa (7.1) je potom aj $\bar{\mathbf{v}}$ vlastným vektorom prislúchajúcim k λ , pričom $\mathbf{v} = \mathbf{x} + i\mathbf{y}$ a $\bar{\mathbf{v}} = \mathbf{x} - i\mathbf{y}$ pre nejaké $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Ak $\mathbf{x} \neq \mathbf{0}$, z príslušnosti vektorov $\mathbf{v}$ a $\bar{\mathbf{v}}$ do vlastného priestoru $E(A, \lambda)$ vyplýva, že aj $\mathbf{v} + \bar{\mathbf{v}} = 2\mathbf{x} \in \mathbb{R}^n$ musí byť nenulovým prvkom $E(A, \lambda)$, a teda aj reálnym vlastným vektorom prislúchajúcim k λ . Ak $\mathbf{x} = \mathbf{0}$, nutne $\mathbf{y} \neq \mathbf{0}$ a rovnaké pozorovanie môžeme urobiť napríklad o vektore $(1/i)\mathbf{y} = -i\mathbf{y}$. $\square$

Ortogonalnou bázou podpriestoru $\mathcal{V}$ vektorového priestoru $\mathbb{R}^n$ pre nejaké $n \in \mathbb{N}$ nazveme ľubovoľnú jeho bázu $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ zloženú z po dvoch ortogonálnych vektorov – pre $i, j = 1, \dots, m$ je teda $\mathbf{v}_i^T \mathbf{v}_j = 0$ kedykoľvek $i \neq j$. *Ortonormálnou bázou* podpriestoru $\mathcal{V}$ nazveme jeho ortogonálnu bázu takú, že všetky jej vektory sú navyše jednotkovej dĺžky: teda $\|\mathbf{v}_k\|_2 = 1$ pre $k = 1, \dots, m$. Každý podpriestor $\mathcal{V}$ priestoru $\mathbb{R}^n$ má nejakú ortonormálnu bázu, ktorú možno nájsť takzvaným *Gramovým-Schmidtovým ortogonalizačným procesom* – ten v nasledujúcom stručne opíšeme.³

³Definície ortogonálnych a ortonormálnych báz ani Gramov-Schmidtov ortogonalizačný proces nie sú ani zďaleka limitované na podpriestory euklidovského priestoru $\mathbb{R}^n$. Rovnako dobre by sme mohli pracovať v ľubovoľnom konečnorozmernom vektorovom priestore s abstraktným skalárnym súčinom, čo by si ale vyžadovalo zaviesť ďalšie pojmy z oblasti lineárnej algebry. Obmedzíme sa tu preto len na podpriestory $\mathbb{R}^n$, pre ktoré možno celú problematiku podať na elementárnej úrovni. Výklad procesu Gramovej-Schmidtovej ortogonalizácie v spomínanom všeobecnejšom kontexte možno nájsť napríklad v [105].

Nech $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ je ľubovoľná báza priestoru $\mathcal{V}$. Pod Gramovým-Schmidtovým ortogonalizačným procesom chápeme konštrukciu, ktorá túto bázu prerobí na ortogonálnu bázu $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ priestoru $\mathcal{V}$ takto:

$$\begin{aligned}\mathbf{v}_1 &:= \mathbf{x}_1, \\ \mathbf{v}_2 &:= \mathbf{x}_2 - \frac{\mathbf{v}_1^T \mathbf{x}_2}{\|\mathbf{v}_1\|_2^2} \mathbf{v}_1, \\ \mathbf{v}_3 &:= \mathbf{x}_3 - \frac{\mathbf{v}_1^T \mathbf{x}_3}{\|\mathbf{v}_1\|_2^2} \mathbf{v}_1 - \frac{\mathbf{v}_2^T \mathbf{x}_3}{\|\mathbf{v}_2\|_2^2} \mathbf{v}_2, \\ &\vdots \\ \mathbf{v}_m &:= \mathbf{x}_m - \frac{\mathbf{v}_1^T \mathbf{x}_m}{\|\mathbf{v}_1\|_2^2} \mathbf{v}_1 - \frac{\mathbf{v}_2^T \mathbf{x}_m}{\|\mathbf{v}_2\|_2^2} \mathbf{v}_2 - \dots - \frac{\mathbf{v}_{m-1}^T \mathbf{x}_m}{\|\mathbf{v}_{m-1}\|_2^2} \mathbf{v}_{m-1}.\end{aligned}$$

Všimnime si najprv, že pre $k = 1, \dots, m$ generujú vektory $\mathbf{v}_1, \dots, \mathbf{v}_k$ rovnaký podpriestor priestoru $\mathbb{R}^n$ ako vektory $\mathbf{x}_1, \dots, \mathbf{x}_k$ – matica so stĺpcami $\mathbf{v}_1, \dots, \mathbf{v}_k$ totiž vznikne z matice so stĺpcami $\mathbf{x}_1, \dots, \mathbf{x}_k$ postupnosťou elementárnych stĺpcových operácií. Pomocou Gramovho-Schmidtovho ortogonalizačného procesu teda naozaj dostávame bázu $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ priestoru $\mathcal{V}$. Pre $k = 1, \dots, m$ a $j = 1, \dots, k-1$ ďalej udáva vektor

$$\frac{\mathbf{v}_j^T \mathbf{x}_k}{\|\mathbf{v}_j\|_2^2} \mathbf{v}_j$$

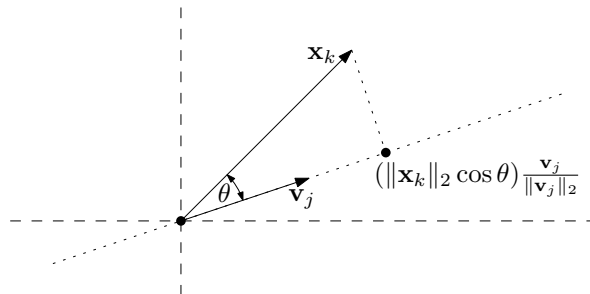
ortogonálnu projekciu vektora $\mathbf{x}_k$ na jednorozmerný vektorový priestor generovaný vektorom $\mathbf{v}_j$, čo je dôsledkom známej charakterizácie skalárneho súčinu, podľa ktorej

$$\mathbf{v}_j^T \mathbf{x}_k = \|\mathbf{v}_j\|_2 \|\mathbf{x}_k\|_2 \cos \theta,$$

kde θ je uhol zvieraný vektormi $\mathbf{v}_j$ a $\mathbf{x}_k$; z toho

$$\frac{\mathbf{v}_j^T \mathbf{x}_k}{\|\mathbf{v}_j\|_2^2} \mathbf{v}_j = (\|\mathbf{x}_k\|_2 \cos \theta) \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|_2}.$$

Táto situácia je znázornená na obrázku 7.2. Ak sú teda vektory $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ navzájom ortogonálne a od vektora $\mathbf{x}_k$ postupne odčítame jeho projekcie na jednorozmerné vektorové priestory generované vektormi $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$, zostane nám vo výsledku vektor $\mathbf{v}_k$, ktorý je na $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ kolmý.



Obr. 7.2: Ortogonálna projekcia vektora $\mathbf{x}_k$ na jednorozmerný priestor generovaný vektorom $\mathbf{v}_j$.

Ortogonalnosť vektorov $\mathbf{v}_1, \dots, \mathbf{v}_m$ však v skutočnosti môžeme overiť aj priamym výpočtom, kde indukciou vzhľadom $k = 1, \dots, m$ overíme, že vektory $\mathbf{v}_1, \dots, \mathbf{v}_k$ sú po dvoch ortogonálne. Tvrdenie evidentne platí pre $k = 1$. Predpokladajme teraz, že pre nejaké $s \in [m-1]$ sú vektory $\mathbf{v}_1, \dots, \mathbf{v}_s$

navzájom ortogonálne. Pre vektor $\mathbf{v}_{s+1}$ a $j = 1, \dots, s$ potom

$$\begin{aligned}\mathbf{v}_j^T \mathbf{v}_{s+1} &= \mathbf{v}_j^T \left(\mathbf{x}_{s+1} - \frac{\mathbf{v}_1^T \mathbf{x}_{s+1}}{\|\mathbf{v}_1\|_2^2} \mathbf{v}_1 - \frac{\mathbf{v}_2^T \mathbf{x}_{s+1}}{\|\mathbf{v}_2\|_2^2} \mathbf{v}_2 - \dots - \frac{\mathbf{v}_s^T \mathbf{x}_{s+1}}{\|\mathbf{v}_s\|_2^2} \mathbf{v}_s \right) = \\ &= \mathbf{v}_j^T \mathbf{x}_{s+1} - \mathbf{v}_j^T \left(\frac{\mathbf{v}_j^T \mathbf{x}_{s+1}}{\|\mathbf{v}_j\|_2^2} \mathbf{v}_j \right) = \mathbf{v}_j^T \mathbf{x}_{s+1} - \mathbf{v}_j^T \mathbf{x}_{s+1} = 0.\end{aligned}$$

Aj vektory $\mathbf{v}_1, \dots, \mathbf{v}_{s+1}$ sú teda po dvoch ortogonálne.

Skonstruovanú ortogonálnu bázu $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ priestoru $\mathcal{V}$ možno upraviť aj na jeho *ortonormálnu* bázu $(\mathbf{u}_1, \dots, \mathbf{u}_m)$:

$$\mathbf{u}_1 := \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|_2}, \quad \mathbf{u}_2 := \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|_2}, \quad \dots, \quad \mathbf{u}_m := \frac{\mathbf{v}_m}{\|\mathbf{v}_m\|_2}.$$

V prípade, že už v pôvodnej báze $(\mathbf{x}_1, \dots, \mathbf{x}_m)$ je prvých niekoľko vektorov $\mathbf{x}_1, \dots, \mathbf{x}_k$ pre $k \in [m]$ vzájomne ortogonálnych a pre $j = 1, \dots, k$ je súčasne $\|\mathbf{x}_j\|_2 = 1$, z uvedeného postupu ľahko vidieť, že týchto prvých k vektorov bude v nezmenenej podobe vystupovať aj vo výslednej ortonormálnej báze: pre $j = 1, \dots, k$ je $\mathbf{u}_j = \mathbf{x}_j$. Z toho okrem iného vyplýva, že každú ortonormálnu bázu priestoru $\mathcal{V}$ možno pridaním niekoľkých vektorov rozšíriť na ortonormálnu bázu celého priestoru $\mathbb{R}^n$.

Naším najbližším cieľom bude dokázať, že každá reálna symetrická matica typu $n \times n$ je *ortogonálne diagonalizovateľná* – to jest, že existuje ortonormálna báza priestoru $\mathbb{R}^n$ zložená z jej vlastných vektorov. Alternatívne možno ortogonálnu diagonalizovateľnosť matice opísať ako diagonalizovateľnosť prostredníctvom *ortogonálnej matice*. Štvorcovú maticu $Q \in \mathbb{R}^{n \times n}$ nazveme *ortogonálnou*, ak jej stĺpce tvoria *ortonormálnu*⁴ bázu priestoru $\mathbb{R}^n$. Pre ľubovoľnú ortogonálnu maticu $Q \in \mathbb{R}^{n \times n}$ evidentne

$$QQ^T = Q^T Q = \mathbf{I}_n,$$

z čoho vyplýva, že každá ortogonálna matica Q je nesingulárna a $Q^{-1} = Q^T$. Uvedená definícia ortogonálnej diagonalizovateľnosti matice A je teda ekvivalentná existencii ortogonálnej matice P takej, že

$$P^{-1}AP = P^T AP$$

je diagonálna matica. V nasledujúcom budeme dokazovať, že každá reálna symetrická matica A má túto vlastnosť.

Definícia 7.2.3. Nech $n \in \mathbb{N}$ a $\mathcal{V}$ je podpriestor vektorového priestoru $\mathbb{R}^n$. *Ortogonálnym doplnkom* priestoru $\mathcal{V}$ v $\mathbb{R}^n$ nazveme vektorový priestor⁵

$$\mathcal{V}^\perp = \{ \mathbf{v} \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{v} = 0 \text{ pre všetky } \mathbf{x} \in \mathcal{V} \}.$$

Keďže pre všetky nenulové vektory $\mathbf{x} \in \mathbb{R}^n$ je $\mathbf{x}^T \mathbf{x} \neq 0$, pre ľubovoľný podpriestor $\mathcal{V}$ priestoru $\mathbb{R}^n$ musí byť $\mathcal{V} \cap \mathcal{V}^\perp = \{ \mathbf{0} \}$. Podľa tvrdenia 4.1.17 je teda lineárny súčet priestorov $\mathcal{V}$ a $\mathcal{V}^\perp$ priamy. Každý vektor $\mathbf{x} \in \mathbb{R}^n$ sa navyše dá vyjadriť ako súčet $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ pre $\mathbf{x}_1 \in \mathcal{V}$ a $\mathbf{x}_2 \in \mathcal{V}^\perp$: ortonormálnu bázu $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ priestoru $\mathcal{V}$ možno doplniť nejakými vektormi $\mathbf{v}_{m+1}, \dots, \mathbf{v}_n$ na ortonormálnu bázu priestoru $\mathbb{R}^n$. Každý z vektorov $\mathbf{v}_{m+1}, \dots, \mathbf{v}_n$ musí patriť do $\mathcal{V}^\perp$ – pre ľubovoľný vektor $\mathbf{v} \in \mathcal{V}$ totiž existujú $c_1, \dots, c_m$ také, že $\mathbf{v} = c_1 \mathbf{v}_1 + \dots + c_m \mathbf{v}_m$ a pre $j = m+1, \dots, n$ tak

$$\mathbf{v}_j^T \mathbf{v} = \mathbf{v}_j^T (c_1 \mathbf{v}_1 + \dots + c_m \mathbf{v}_m) = c_1 \mathbf{v}_j^T \mathbf{v}_1 + \dots + c_m \mathbf{v}_j^T \mathbf{v}_m = 0.$$

Každý vektor $\mathbf{x} \in \mathbb{R}^n$ navyše možno pre nejaké $a_1, \dots, a_n \in \mathbb{R}$ vyjadriť ako

$$\mathbf{x} = (a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m) + (a_{m+1} \mathbf{v}_{m+1} + \dots + a_n \mathbf{v}_n),$$

pričom $\mathbf{x}_1 := a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m \in \mathcal{V}$ a z uvedeného vyplýva, že $\mathbf{x}_2 := a_{m+1} \mathbf{v}_{m+1} + \dots + a_n \mathbf{v}_n \in \mathcal{V}^\perp$. Pre ľubovoľný podpriestor $\mathcal{V}$ priestoru $\mathbb{R}^n$ je teda $\mathbb{R}^n = \mathcal{V} \oplus \mathcal{V}^\perp$.

⁴Tento drobný nesúlad v terminológii je v literatúre štandardný.

⁵Ide tu o skryté, avšak ľahké tvrdenie: ak $\mathbf{u}, \mathbf{v} \in \mathcal{V}^\perp$ a $a, b \in \mathbb{R}$, tak pre všetky $\mathbf{x} \in \mathcal{V}$ je $\mathbf{x}^T (a\mathbf{u} + b\mathbf{v}) = a\mathbf{x}^T \mathbf{u} + b\mathbf{x}^T \mathbf{v} = 0$ a vektor $a\mathbf{u} + b\mathbf{v}$ tak tiež patrí do $\mathcal{V}^\perp$.

Lema 7.2.4. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. Ak je podpriestor $\mathcal{V}$ priestoru $\mathbb{R}^n$ invariantný vzhľadom na A , tak je vzhľadom na A invariantný aj priestor $\mathcal{V}^\perp$.*

Dôkaz. Z invariantnosti priestoru $\mathcal{V}$ vzhľadom na maticu A vyplýva, že pre všetky $\mathbf{x} \in \mathcal{V}$ je $A\mathbf{x} \in \mathcal{V}$. Pre všetky $\mathbf{x} \in \mathcal{V}$ a $\mathbf{y} \in \mathcal{V}^\perp$ preto

$$0 = (A\mathbf{x})^T \mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = \mathbf{x}^T A \mathbf{y} = \mathbf{x}^T (A\mathbf{y}).$$

Vektor $\mathbf{x} \in \mathcal{V}$ môže byť ľubovoľný. Zisťujeme teda, že pre všetky $\mathbf{y} \in \mathcal{V}^\perp$ je $A\mathbf{y} \in \mathcal{V}^\perp$ – priestor $\mathcal{V}^\perp$ je invariantný vzhľadom na A . $\square$

Lema 7.2.5. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. Každý netriviálny podpriestor priestoru $\mathbb{R}^n$ invariantný vzhľadom na A potom obsahuje aspoň jeden reálny vlastný vektor matice A .*

Dôkaz. Nech $\mathcal{V}$ je ľubovoľný netriviálny podpriestor priestoru $\mathbb{R}^n$ invariantný vzhľadom na A . Predpokladajme, že $\mathcal{V}$ je dimenzie $m \in \mathbb{N} \setminus \{0\}$. Musí potom existovať nejaká ortonormálna báza $(\mathbf{v}_1, \dots, \mathbf{v}_m)$ priestoru $\mathcal{V}$. Uvažujme maticu

$$R = \begin{pmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_m \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} \in \mathbb{R}^{n \times m}.$$

Evidentne potom

$$R^T R = \mathbf{I}_m. \quad (7.2)$$

Keďže je priestor $\mathcal{V}$ invariantný vzhľadom na A , nutne

$$AR = \begin{pmatrix} \uparrow & \uparrow & \dots & \uparrow \\ A\mathbf{v}_1 & A\mathbf{v}_2 & \dots & A\mathbf{v}_m \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow & \dots & \uparrow \\ R\mathbf{x}_1 & R\mathbf{x}_2 & \dots & R\mathbf{x}_m \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} = RB \quad (7.3)$$

pre nejaké $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^m$ a maticu

$$B = \begin{pmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_m \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} \in \mathbb{R}^{m \times m}.$$

Z rovností (7.3) a (7.2) potom

$$R^T AR = R^T RB = \mathbf{I}_m B = B.$$

Reálna matica B tak musí byť symetrická, pretože

$$B^T = (R^T AR)^T = R^T A^T R = R^T AR = B.$$

Podľa vety 7.2.2 sú teda všetky vlastné čísla matice B reálne a ku každému z nich prislúcha nejaký reálny vlastný vektor. Zoberme jedno z týchto reálnych vlastných čísel λ a príslušný reálny vlastný vektor $\mathbf{v}$. Potom

$$AR\mathbf{v} = RB\mathbf{v} = \lambda R\mathbf{v},$$

z čoho vyplýva, že $R\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ je vlastný vektor matice A prislúchajúci k λ .⁶ Keďže ale všetky stĺpce matice R patria do $\mathcal{V}$, musí byť $R\mathbf{v} \in \mathcal{V}$. $\square$

⁶Nenulovosť vektora $R\mathbf{v}$ je dôsledkom nenulovosti vlastného vektora $\mathbf{v}$ a lineárnej nezávislosti stĺpcov matice R .

Veta 7.2.6. *Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. Potom existuje ortonormálna báza $\mathbb{R}^n$ zložená z reálnych vlastných vektorov matice A – matica A je teda ortogonálne diagonalizovateľná.*

Dôkaz. Pre $n = 0$ je tvrdenie triviálne – nech teda $n \geq 1$. Vezmeme za $\mathbf{v}_1$ ľubovoľný reálny vlastný vektor matice A normovaný tak, aby $\|\mathbf{v}_1\|_2 = 1$; taký vektor musí existovať podľa vety 7.2.2. Predpokladajme, že pre nejaké $m \in [n - 1]$ už poznáme m lineárne nezávislých ortogonálnych reálnych vlastných vektorov $\mathbf{v}_1, \dots, \mathbf{v}_m$ matice A takých, že pre $k = 1, \dots, m$ je $\|\mathbf{v}_k\|_2 = 1$. Označme nimi generovaný m -rozmerný podpriestor priestoru $\mathbb{R}^n$ ako $\mathcal{V}$. Ortogonálny doplnok $\mathcal{V}^\perp$ priestoru $\mathcal{V}$ je netriviálny podpriestor priestoru $\mathbb{R}^n$, pretože $\mathcal{V} \oplus \mathcal{V}^\perp = \mathbb{R}^n$ a $m < n$. Ľahko vidieť, že priestor $\mathcal{V}$ je invariantný vzhľadom na A ; podľa lemy 7.2.4 tak musí byť vzhľadom na A invariantný aj priestor $\mathcal{V}^\perp$, ktorý v dôsledku toho musí podľa lemy 7.2.5 obsahovať aspoň jeden reálny vlastný vektor $\mathbf{v}_{m+1}$ matice A . Bez ujmy na všeobecnosti môžeme predpokladať, že $\|\mathbf{v}_{m+1}\|_2 = 1$ a keďže $\mathbf{v}_{m+1} \in \mathcal{V}^\perp$ a $\mathbf{v}_1, \dots, \mathbf{v}_m$ sú vzájomne ortogonálne vektory z $\mathcal{V}$, musia byť navzájom ortogonálne aj vektory $\mathbf{v}_1, \dots, \mathbf{v}_{m+1}$. S použitím indukcie vzhľadom na $m = 1, \dots, n$ tak zisťujeme, že naozaj existuje ortonormálna báza $\mathbf{v}_1, \dots, \mathbf{v}_n$ priestoru $\mathbb{R}^n$ zložená z reálnych vlastných vektorov matice A . $\square$

Poznámka 7.2.7. Spektrum reálnej symetrickej matice $A \in \mathbb{R}^{n \times n}$ budeme často zapisovať v tvare

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \quad \text{alebo} \quad \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n,$$

kde $\sigma(A) = \{\lambda_1, \dots, \lambda_n\} \subseteq \mathbb{R}$ a pre všetky $\lambda \in \sigma(A)$ je počet rôznych indexov $j \in [n]$ takých, že $\lambda_j = \lambda$, rovný algebraickej násobnosti $\alpha(A, \lambda)$ vlastného čísla λ . Z vyššie dokázaného vyplýva, že pre $n \in \mathbb{N}$ a každú reálnu symetrickú maticu $A \in \mathbb{R}^{n \times n}$ s vlastnými číslami $\lambda_1 \geq \dots \geq \lambda_n$ existuje ortogonálna matica $Q \in \mathbb{R}^{n \times n}$ taká, že

$$Q^{-1}AQ = Q^T A Q = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Z ortogonálnosti matice Q navyše $Q^{-1} = Q^T$ a $Q^{-1}Q = Q^T Q = \mathbf{I}_n$.

7.3 Rayleighove podiely a Courantova-Fischerova veta

Užitočná technika analýzy spektier reálnych symetrických matíc je založená na *Rayleighových podieloch*, ktorých úlohu v teórii takýchto matíc možno pripodobniť k úlohe Collatzovej-Wielandtovej funkcie v teórii nezáporných matíc. Takzvaný *Rayleighov princíp* nám umožní pomocou Rayleighových podielov charakterizovať najväčšie a najmenšie vlastné číslo reálnej symetrickej matice, ako aj vlastné vektory k týmto vlastným číslam prislúchajúce. Toto pomerne jednoduché pozorovanie následne zovšeobecníme do podoby *Courantovej-Fischerovej vety*, známej aj ako *veta o minimaxe*, ktorá pomocou Rayleighových podielov charakterizuje *všetky* vlastné čísla reálnej symetrickej matice.

Definícia 7.3.1. Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. *Rayleighovým podielom* nazveme pre každé $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ podiel

$$R_A(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}},$$

čím je tiež definovaná funkcia $R_A: \mathbb{R}^n \setminus \{\mathbf{0}\} \rightarrow \mathbb{R}$.

Podobne ako pri Collatzovej-Wielandtovej funkcii nie je ťažké vidieť, že pre všetky $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ a všetky $q \in \mathbb{R} \setminus \{0\}$ je $R_A(q\mathbf{x}) = R_A(\mathbf{x})$ – skutočne

$$R_A(q\mathbf{x}) = \frac{(q\mathbf{x})^T A (q\mathbf{x})}{(q\mathbf{x})^T (q\mathbf{x})} = \frac{q^2 \mathbf{x}^T A \mathbf{x}}{q^2 \mathbf{x}^T \mathbf{x}} = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = R_A(\mathbf{x}).$$

Všetky hodnoty funkcie R_A sú teda jednoznačne určené jej hodnotami na jednotkovej $(n-1)$ -sfére

$$S_{n-1} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 = 1\}.$$

Funkcia R_A je na kompakte S_{n-1} očividne spojitá a musí tam preto nadobúdať aspoň jedno maximum a minimum. Môžeme teda uzavrieť, že funkcia R_A nadobúda *globálne* maximum aj minimum. Ukážeme teraz, že pre reálnu symetrickú maticu $A \in \mathbb{R}^{n \times n}$ s vlastnými číslami

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

je globálnym maximom funkcie R_A najväčšie vlastné číslo λ_1 ; podobne globálnym minimom funkcie R_A je najmenšie vlastné číslo λ_n . Argumentmi maxima resp. minima funkcie R_A sú pritom práve všetky vlastné vektory matice A prislúchajúce k λ_1 resp. λ_n .

Veta 7.3.2 (Rayleighov princíp). *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica s vlastnými číslami*

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n.$$

Pre všetky $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ potom

$$\lambda_1 \geq R_A(\mathbf{x}) \geq \lambda_n,$$

pričom $R_A(\mathbf{x}) = \lambda_1$ práve vtedy, keď je $\mathbf{x}$ vlastným vektorom matice A prislúchajúcim k λ_1 a $R_A(\mathbf{x}) = \lambda_n$ práve vtedy, keď je $\mathbf{x}$ vlastným vektorom A prislúchajúcim k λ_n .

Dôkaz. Označme ako Λ diagonálnu maticu

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Musí potom existovať ortogonálna matica $Q \in \mathbb{R}^{n \times n}$ taká, že

$$Q^{-1}AQ = Q^T A Q = \Lambda,$$

a teda aj ortogonálna matica $P = Q^T$ taká, že

$$A = P^T \Lambda P.$$

Stĺpce matice P^T sú teda dané ako $\mathbf{v}_1, \dots, \mathbf{v}_n$, kde pre $k = 1, \dots, n$ je $\mathbf{v}_k$ vlastným vektorom matice A prislúchajúcim k vlastnému číslu λ_k a $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ je ortonormálna báza priestoru $\mathbb{R}^n$.

Nech je teraz vektor $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ daný ľubovoľne. Označme $P\mathbf{x} =: (y_1, \dots, y_n)^T$. Potom

$$\begin{aligned} R_A(\mathbf{x}) &= \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{1}{\|\mathbf{x}\|_2^2} (\mathbf{x}^T A \mathbf{x}) = \frac{1}{\|\mathbf{x}\|_2^2} (\mathbf{x}^T P^T \Lambda P \mathbf{x}) = \frac{1}{\|\mathbf{x}\|_2^2} ((P\mathbf{x})^T \Lambda (P\mathbf{x})) = \\ &= \frac{1}{\|\mathbf{x}\|_2^2} \sum_{j=1}^n y_j^2 \lambda_j. \end{aligned} \quad (7.4)$$

Dĺžka vektora $P\mathbf{x} = (y_1, \dots, y_n)^T$ pritom spĺňa

$$\|P\mathbf{x}\|_2 = \sqrt{(P\mathbf{x})^T (P\mathbf{x})} = \sqrt{\mathbf{x}^T P^T P \mathbf{x}} = \sqrt{\mathbf{x}^T \mathbf{x}} = \|\mathbf{x}\|_2, \quad (7.5)$$

vdďaka čomu zisťujeme, že

$$R_A(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|_2^2} \sum_{j=1}^n y_j^2 \lambda_j \leq \frac{\lambda_1}{\|\mathbf{x}\|_2^2} \sum_{j=1}^n y_j^2 = \lambda_1 \frac{\|P\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = \lambda_1$$

a podobne

$$R_A(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|_2^2} \sum_{j=1}^n y_j^2 \lambda_j \geq \frac{\lambda_n}{\|\mathbf{x}\|_2^2} \sum_{j=1}^n y_j^2 = \lambda_n \frac{\|P\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = \lambda_n.$$

Kedykoľvek je $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ vlastným vektorom matice A prislúchajúcim k niektorému jej vlastnému číslu λ , musí byť

$$R_A(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\lambda \mathbf{x}^T \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \lambda.$$

Zostáva teda ukázať, že v prípade rovnosti $R_A(\mathbf{x}) = \lambda_1$ resp. $R_A(\mathbf{x}) = \lambda_n$ musí byť $\mathbf{x}$ vlastným vektorom matice A prislúchajúcim k λ_1 resp. k λ_n .

Tvrdenie dokážeme pre vlastné číslo λ_1 , pričom argumentácia pre λ_n je symetrická. Uvažujme ľubovoľný vektor $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ taký, že $R_A(\mathbf{x}) = \lambda_1$. Z rovností (7.4) a (7.5) potom vidieť, že y_j môže byť nenulové iba pre $j \in [n]$ také, že $\lambda_j = \lambda_1$. Existuje teda $k \in [n]$ také, že $\lambda_1 = \dots = \lambda_k$ a $P\mathbf{x} = (y_1, \dots, y_k, 0, \dots, 0)^T$. V takom prípade je tiež prvých k stĺpcov $\mathbf{v}_1, \dots, \mathbf{v}_k$ matice P^T tvorených vlastnými vektormi matice A prislúchajúcimi k λ_1 , čiže prvkami vlastného priestoru $E(A, \lambda_1)$.

Keďže je matica P ortogonálna, nutne $P^T P = \mathbf{I}_n$, z čoho

$$\mathbf{x} = (P^T P) \mathbf{x} = P^T (P\mathbf{x}) = \sum_{j=1}^k y_j \mathbf{v}_j.$$

Ako lineárna kombinácia vektorov z $E(A, \lambda_1)$ tak aj vektor $\mathbf{x}$ musí byť prvkom $E(A, \lambda_1)$ a vďaka jeho nenulovosti teda musí ísť o vlastný vektor matice A prislúchajúci k vlastnému číslu λ_1 . $\square$

Práve dokázaný Rayleighov princíp teda charakterizuje *najväčšie* a *najmenšie* vlastné číslo reálnej symetrickej matice A ako maximum resp. minimum funkcie R_A ; príslušné vlastné vektory sú navyše dané ako argumenty tohto maxima resp. minima. Nasledujúca *Courantova-Fischerova veta*, niekedy nazývaná aj *vetou o minimaxe*, je zovšeobecnením časti Rayleighovho princípu hovoriacej o vlastných číslach: umožní nám pomocou Rayleighových podielov charakterizovať *všetky* vlastné čísla matice A .

Veta 7.3.3 (Courant, Fischer). *Nech $n \in \mathbb{N} \setminus \{0\}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica s vlastnými číslami*

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n.$$

Pre $k = 1, \dots, n$ potom

$$\lambda_k = \max_{\substack{\mathcal{V} \\ \mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=k}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) = \min_{\substack{\mathcal{V} \\ \mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=n-k+1}} \max_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}).$$

Dôkaz. Začnime opäť pozorovaním, že musí existovať ortogonálna matica $Q \in \mathbb{R}^{n \times n}$ taká, že

$$Q^{-1} A Q = Q^T A Q = \Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

a teda aj ortogonálna matica $P = Q^T$ taká, že

$$A = P^T \Lambda P.$$

Stĺpce matice P^T tvoria ortonormálnu bázu priestoru $\mathbb{R}^n$ zloženú z vlastných vektorov $\mathbf{v}_1, \dots, \mathbf{v}_n$ matice A prislúchajúcich postupne k vlastným číslam $\lambda_1, \dots, \lambda_n$.

Uvažujme pre $k = 1, \dots, n$ podpriestor $\mathcal{V}_k$ priestoru $\mathbb{R}^n$ generovaný vektormi $\mathbf{v}_1, \dots, \mathbf{v}_k$; tie tvoria jeho ortonormálnu bázu a evidentne teda $\dim(\mathcal{V}_k) = k$. Zvoľme $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathcal{V}_k \setminus \{\mathbf{0}\}$ ľubovoľne a označme $P\mathbf{x} = (y_1, \dots, y_n)^T$. Z ortogonálnosti matice P rovnako ako v (7.5) dostávame $\|P\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ a jej dôsledkom je aj rovnosť

$$\mathbf{x} = (P^T P) \mathbf{x} = P^T (P\mathbf{x}) = \sum_{j=1}^n y_j \mathbf{v}_j.$$

Z príslušnosti vektora $\mathbf{x}$ do $\mathcal{V}_k$ teda vyplýva, že $y_{k+1} = \dots = y_n = 0$. Pre hodnotu príslušného Rayleighovho podielu tak prichádzame k odhadu

$$\begin{aligned} R_A(\mathbf{x}) &= \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{1}{\|\mathbf{x}\|_2^2} (\mathbf{x}^T A \mathbf{x}) = \frac{1}{\|\mathbf{x}\|_2^2} (\mathbf{x}^T P^T \Lambda P \mathbf{x}) = \frac{1}{\|\mathbf{x}\|_2^2} ((P \mathbf{x})^T \Lambda (P \mathbf{x})) = \\ &= \frac{1}{\|\mathbf{x}\|_2^2} \sum_{j=1}^n y_j^2 \lambda_j = \frac{1}{\|\mathbf{x}\|_2^2} \sum_{j=1}^k y_j^2 \lambda_j \geq \frac{\lambda_k}{\|\mathbf{x}\|_2^2} \sum_{j=1}^k y_j^2 = \frac{\lambda_k}{\|\mathbf{x}\|_2^2} \sum_{j=1}^n y_j^2 = \lambda_k \frac{\|P \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = \lambda_k, \end{aligned}$$

z čoho

$$\lambda_k \leq \min_{\mathbf{x} \in \mathcal{V}_k \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) \leq \sup_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=k}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}); \quad (7.6)$$

existencia miním tu vyplýva zo spojitosti funkcie R_A na kompakte $\mathcal{V} \cap S_{n-1}$ pre ľubovoľný k -rozmerný podpriestor $\mathcal{V}$ priestoru $\mathbb{R}^n$ a zo skutočnosti, že hodnotami na $\mathcal{V} \cap S_{n-1}$ je R_A určená na celom $\mathcal{V} \setminus \{\mathbf{0}\}$.

Pre $k = 1, \dots, n$ ďalej uvažujme ľubovoľný k -rozmerný podpriestor $\mathcal{V}$ priestoru $\mathbb{R}^n$. Jeho prienik s $(n-k+1)$ -rozmerným podpriestorom $\mathcal{V}'_{n-k+1}$ priestoru $\mathbb{R}^n$ generovaným vektormi $\mathbf{v}_k, \dots, \mathbf{v}_n$ je nutne netriviálny. Vezmime ľubovoľný vektor

$$\mathbf{x}_{\mathcal{V}} \in (\mathcal{V} \cap \mathcal{V}'_{n-k+1}) \setminus \{\mathbf{0}\}.$$

Ak označíme $\mathbf{x}_{\mathcal{V}} =: (x_1, \dots, x_n)^T$ a $P \mathbf{x}_{\mathcal{V}} =: (y_1, \dots, y_n)^T$, rovnako ako vyššie zisťujeme, že

$$\mathbf{x}_{\mathcal{V}} = (P^T P) \mathbf{x}_{\mathcal{V}} = P^T (P \mathbf{x}_{\mathcal{V}}) = \sum_{j=1}^n y_j \mathbf{v}_j$$

a z príslušnosti vektora $\mathbf{x}_{\mathcal{V}}$ do $\mathcal{V}'_{n-k+1}$ dostávame $y_1 = \dots = y_{k-1} = 0$. Keďže tiež $\|P \mathbf{x}_{\mathcal{V}}\|_2 = \|\mathbf{x}_{\mathcal{V}}\|_2$, prichádzame k zisteniu, že

$$\begin{aligned} R_A(\mathbf{x}_{\mathcal{V}}) &= \frac{\mathbf{x}_{\mathcal{V}}^T A \mathbf{x}_{\mathcal{V}}}{\mathbf{x}_{\mathcal{V}}^T \mathbf{x}_{\mathcal{V}}} = \frac{1}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} (\mathbf{x}_{\mathcal{V}}^T A \mathbf{x}_{\mathcal{V}}) = \frac{1}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} (\mathbf{x}_{\mathcal{V}}^T P^T \Lambda P \mathbf{x}_{\mathcal{V}}) = \frac{1}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} ((P \mathbf{x}_{\mathcal{V}})^T \Lambda (P \mathbf{x}_{\mathcal{V}})) = \\ &= \frac{1}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} \sum_{j=1}^n y_j^2 \lambda_j = \frac{1}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} \sum_{j=k}^n y_j^2 \lambda_j \leq \frac{\lambda_k}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} \sum_{j=k}^n y_j^2 = \frac{\lambda_k}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} \sum_{j=1}^n y_j^2 = \lambda_k \frac{\|P \mathbf{x}_{\mathcal{V}}\|_2^2}{\|\mathbf{x}_{\mathcal{V}}\|_2^2} = \lambda_k, \end{aligned}$$

z čoho

$$\lambda_k \geq \sup_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=k}} R_A(\mathbf{x}_{\mathcal{V}}) \geq \sup_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=k}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}). \quad (7.7)$$

Ako bezprostredný dôsledok nerovností (7.6) a (7.7) tak pre $k = 1, \dots, n$ dostávame rovnosť

$$\lambda_k = \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=k}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}). \quad (7.8)$$

Rovnosť

$$\lambda_k = \min_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=n-k+1}} \max_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x})$$

napokon vyplynie z rovnosti (7.8) pre maticu $-A$, ktorá má podľa vety 4.8.2 vlastné čísla

$$-\lambda_n \geq -\lambda_{n-1} \geq \dots \geq -\lambda_1,$$

čo môžeme po preznačení $\lambda'_k := -\lambda_{n-k+1}$ pre $k = 1, \dots, n$ prepísať ako

$$\lambda'_1 \geq \lambda'_2 \geq \dots \geq \lambda'_n.$$

Pre všetky $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ navyše

$$R_{-A}(\mathbf{x}) = \frac{\mathbf{x}^T(-A)\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = -\frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = -R_A(\mathbf{x}).$$

Pre vlastné číslo $-\lambda_k = \lambda'_{n-k+1}$ matice $-A$ tak zo vzťahu (7.8) vyplýva

$$-\lambda_k = \lambda'_{n-k+1} = \max_{\substack{\mathcal{V} \\ \mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=n-k+1}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{-A}(\mathbf{x}) = \max_{\substack{\mathcal{V} \\ \mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=n-k+1}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} -R_A(\mathbf{x}),$$

a teda

$$\lambda_k = \min_{\substack{\mathcal{V} \\ \mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=n-k+1}} \max_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}).$$

Tým je veta dokázaná. □

7.4 Cauchyho veta o prepletaní vlastných čísel

Ako posledný z výsledkov o spektrách reálnych symetrických matíc, ktoré budeme využívať v súvislosti so spektrami grafov, teraz dokážeme tzv. *Cauchyho vetu o prepletaní vlastných čísel* (angl. *Cauchy's interlacing theorem*). Tá dáva spektrum reálnej symetrickej matice $A \in \mathbb{R}^{n \times n}$ pre prirodzené $n \geq 2$ do súvisu so spektrom matice $A[k]$, ktorá z nej vznikne odstránením k -teho riadku a stĺpca pre nejaké $k \in [n]$. Ukazuje sa, že vlastné čísla matice $A[k]$ sa „prepletajú pomedzi“ vlastné čísla matice A . Dôkaz Cauchyho vety o prepletaní bude založený na použití Courantovej-Fischerovej vety.

Veta 7.4.1 (Cauchyho o prepletaní). *Nech $n \in \mathbb{N} \setminus \{0, 1\}$ a $A = (a_{i,j})_{n \times n} \in \mathbb{R}^{n \times n}$ je symetrická matica s vlastnými číslami*

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n.$$

Pre $k = 1, \dots, n$ označme ako $A[k] \in \mathbb{R}^{(n-1) \times (n-1)}$ symetrickú maticu, ktorá vznikne z A odstránením k -teho riadku a k -teho stĺpca:

$$A[k] = \begin{pmatrix} a_{1,1} & \dots & a_{1,k-1} & a_{1,k+1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{k-1,1} & \dots & a_{k-1,k-1} & a_{k-1,k+1} & \dots & a_{k-1,n} \\ a_{k+1,1} & \dots & a_{k+1,k-1} & a_{k+1,k+1} & \dots & a_{k+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,k-1} & a_{n,k+1} & \dots & a_{n,n} \end{pmatrix}.$$

Nech sú vlastné čísla matice $A[k]$ dané ako

$$\kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_{n-1}.$$

Potom

$$\lambda_1 \geq \kappa_1 \geq \lambda_2 \geq \kappa_2 \geq \lambda_3 \geq \dots \geq \lambda_{n-1} \geq \kappa_{n-1} \geq \lambda_n.$$

Dôkaz. Nech $\mathbf{e}_1 = (1, 0, \dots, 0)^T, \mathbf{e}_2 = (0, 1, \dots, 0)^T, \dots, \mathbf{e}_n = (0, 0, \dots, 1)^T$ sú vektory štandardnej bázy priestoru $\mathbb{R}^n$. Pre $k = 1, \dots, n$ označme ako $\mathcal{V}_k$ podpriestor priestoru $\mathbb{R}^n$ generovaný vektormi $\mathbf{e}_1, \dots, \mathbf{e}_{k-1}, \mathbf{e}_{k+1}, \dots, \mathbf{e}_n$. Prirodzene potom môžeme definovať bijekciu $\varphi: \mathbb{R}^{n-1} \rightarrow \mathcal{V}_k$ tak, že pre všetky $\mathbf{x} = (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)^T \in \mathbb{R}^{n-1}$ položíme $\varphi(\mathbf{x}) = (x_1, \dots, x_{k-1}, 0, x_{k+1}, \dots, x_n)^T$. Je potom zrejmé, že pre všetky $\mathbf{x} \in \mathbb{R}^{n-1} \setminus \{\mathbf{0}\}$ je

$$R_{A[k]}(\mathbf{x}) = \frac{\mathbf{x}^T A[k] \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = \frac{\varphi(\mathbf{x})^T A \varphi(\mathbf{x})}{\varphi(\mathbf{x})^T \varphi(\mathbf{x})} = R_A(\varphi(\mathbf{x})).$$

Pretože je φ bijekcia zobrazujúca podpriestory $\mathbb{R}^{n-1}$ na podpriestory $\mathcal{V}_k$, s pomocou vety 7.3.3 zistujeme, že pre $j = 1, \dots, n-1$ je

$$\begin{aligned} \kappa_j &= \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^{n-1} \\ \dim(\mathcal{V})=j}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{A[k]}(\mathbf{x}) = \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathcal{V}_k \\ \dim(\mathcal{V})=j}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) \leq \\ &\leq \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=j}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) = \lambda_j. \end{aligned}$$

Z vety 7.3.3 na druhej strane vyplýva

$$\lambda_{j+1} = \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=j+1}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}).$$

Označme ako $\mathcal{W}$ podpriestor priestoru $\mathbb{R}^n$, na ktorom sa toto maximum nadobúda:

$$\mathcal{W} \in \arg \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=j+1}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}),$$

a teda

$$\min_{\mathbf{x} \in \mathcal{W} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) = \lambda_{j+1}.$$

Potom $\dim(\mathcal{W} \cap \mathcal{V}_k) \geq j$ a

$$\begin{aligned} \kappa_j &= \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^{n-1} \\ \dim(\mathcal{V})=j}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{A[k]}(\mathbf{x}) = \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathcal{V}_k \\ \dim(\mathcal{V})=j}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) \geq \\ &\geq \max_{\substack{\mathcal{V} \text{ je podpriestor } \mathcal{W} \cap \mathcal{V}_k \\ \dim(\mathcal{V})=j}} \min_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) \geq \min_{\mathbf{x} \in (\mathcal{W} \cap \mathcal{V}_k) \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) \geq \min_{\mathbf{x} \in \mathcal{W} \setminus \{\mathbf{0}\}} R_A(\mathbf{x}) = \lambda_{j+1}. \end{aligned}$$

Tým je dôkaz vety dokončený. □

7.5 Spektrálna charakterizácia regulárnych grafov

Pod *obyčajným spektrom* grafu⁷ $\mathcal{G}$ budeme rozumieť spektrum jeho matice susednosti $A = \mathbf{A}(\mathcal{G})$. Tá je vždy reálna a symetrická, takže obyčajné spektrum grafu o n vrcholoch možno zadať ako

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n,$$

kde tento zápis treba interpretovať ako v poznámke 7.2.7.

⁷V súlade s dohovorom z úvodu tejto kapitoly máme pod grafom na mysli neorientovaný graf *bez slučiek*. Nezáleží teda na tom, či matice susednosti takýchto grafov chápeme v zmysle definície 7.1.2, alebo v zmysle definície 7.1.4 (pokiaľ v prvom prípade interpretujeme booleovskú maticu ako maticu prirodzených čísel 0 a 1).

Ako ukážku relatívne elementárneho výsledku spektrálnej teórie grafov teraz pomocou vlastností obyčajných spektier charakterizujeme *regulárne grafy*; táto charakterizácia bude daná nasledujúcou vetou a predovšetkým jej dôsledkom 7.5.2. Pripomeňme, že graf $\mathcal{G} = (V, \varphi)$ je *d-regulárny* pre $d \in \mathbb{N}$, ak $\deg_{\mathcal{G}}(v) = d$ pre všetky $v \in V$ a *regulárny*, ak je *d-regulárny* pre *nejaké* $d \in \mathbb{N}$. Prezentácia tohto oddielu čiastočne vychádza z [26].

Veta 7.5.1 (Collatz, Sinogowitz). *Nech $n \in \mathbb{N} \setminus \{0\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s maticou susednosti $A = \mathbf{A}(\mathcal{G})$ a obyčajným spektrom $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Nech navyše*

$$\Delta := \max \{ \deg_{\mathcal{G}}(j) \mid j \in V(\mathcal{G}) \}$$

označuje maximálny a

$$\delta := \frac{1}{n} \sum_{j=1}^n \deg_{\mathcal{G}}(j)$$

priemerný stupeň vrcholu grafu $\mathcal{G}$. Potom

$$\delta \leq \lambda_1 \leq \Delta$$

a graf $\mathcal{G}$ je regulárny práve vtedy, keď $\delta = \lambda_1$. V takom prípade je λ_1 prirodzené číslo a graf $\mathcal{G}$ je λ_1 -regulárny. Ak je navyše graf $\mathcal{G}$ súvislý, je graf $\mathcal{G}$ regulárny práve vtedy, keď $\lambda_1 = \Delta$.

Dôkaz. Uvažujme ľubovoľný reálny vlastný vektor $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ matice A prislúchajúci k vlastnému číslu λ_1 . Potom $\lambda_1 \mathbf{x} = A\mathbf{x}$, z čoho pre $k = 1, \dots, n$ vyplýva

$$\lambda_1 x_k = \sum_{j \in N_{\mathcal{G}}(k)} x_j, \quad (7.9)$$

kde $N_{\mathcal{G}}(k)$ označuje množinu všetkých susedov vrcholu k v grafe $\mathcal{G}$. Bez ujmy na všeobecnosti môžeme predpokladať, že aspoň jedna zo zložiek $x_1, \dots, x_n$ vektora $\mathbf{x}$ je kladná – v opačnom prípade by sme totiž mohli namiesto neho uvažovať vlastný vektor $-\mathbf{x}$. Túto vlastnosť má potom špeciálne aj ľubovoľná spomedzi najväčších zložiek vektora $\mathbf{x}$; nech je jednou z nich zložka x_m pre $m \in [n]$. Z rovnosti (7.9) potom

$$\lambda_1 x_m = \sum_{j \in N_{\mathcal{G}}(m)} x_j \leq \sum_{j \in N_{\mathcal{G}}(m)} x_m \leq \Delta x_m$$

a vďaka kladnosti čísla x_m tak prichádzame k nerovnosti

$$\lambda_1 \leq \Delta. \quad (7.10)$$

Označme ďalej ako $\mathbf{1}$ vektor

$$\mathbf{1} := (1, 1, \dots, 1)^T \in \mathbb{R}^n.$$

Podľa vety 7.3.2 potom

$$\lambda_1 = \max_{\mathbf{y} \in \mathbb{R}^n \setminus \{0\}} R_A(\mathbf{y}) = \max_{\mathbf{y} \in \mathbb{R}^n \setminus \{0\}} \frac{\mathbf{y}^T A \mathbf{y}}{\mathbf{y}^T \mathbf{y}} \geq \frac{\mathbf{1}^T A \mathbf{1}}{\mathbf{1}^T \mathbf{1}} = \frac{2|E(\mathcal{G})|}{n} = \frac{1}{n} \sum_{j \in V(\mathcal{G})} \deg_{\mathcal{G}}(j) = \delta. \quad (7.11)$$

Spojením nerovností (7.10) a (7.11) tak skutočne dostávame

$$\delta \leq \lambda_1 \leq \Delta.$$

Z nerovnosti (7.11) tiež vyplýva, že $\lambda_1 = \delta$ práve vtedy, keď

$$\frac{\mathbf{1}^T A \mathbf{1}}{\mathbf{1}^T \mathbf{1}} = R_A(\mathbf{1}) = \lambda_1,$$

čo podľa vety 7.3.2 nastane práve vtedy, keď je $\mathbf{1}$ vlastným vektorom matice A prislúchajúcim k vlastnému číslu λ_1 , čiže keď

$$A\mathbf{1} = \lambda_1 \mathbf{1}.$$

To je ďalej ekvivalentné tomu, že pre $k = 1, \dots, n$ je

$$\sum_{j \in N_G(k)} 1 = \deg_G(k) = \lambda_1.$$

Rovnosť $\lambda_1 = \delta$ teda naozaj nastane práve vtedy, keď je graf $\mathcal{G}$ regulárny stupňa $\lambda_1 = \delta$.

Je zrejmé, že pre ľubovoľný regulárny graf $\mathcal{G}$ musí byť aj $\lambda_1 = \Delta$; pre regulárne grafy totiž $\Delta = \delta$. Predpokladajme teda napokon, že je graf $\mathcal{G}$ súvislý a $\lambda_1 = \Delta$. Dokážeme, že graf $\mathcal{G}$ je v takom prípade regulárny. Nech $\mathbf{x} = (x_1, \dots, x_n)^T$ je ľubovoľný reálny vlastný vektor prislúchajúci k λ_1 , o ktorom môžeme podobne ako vyššie bez ujmy na všeobecnosti predpokladať, že obsahuje aspoň jednu kladnú zložku.⁸ Uvažujme ľubovoľné

$$m \in \arg \max_{k \in [n]} x_k =: M;$$

nutne $x_m > 0$. Z rovnosti

$$A\mathbf{x} = \lambda_1 \mathbf{x} = \Delta \mathbf{x}$$

potom špeciálne

$$\sum_{j \in N_G(m)} x_j = \Delta x_m,$$

čo môže nastať iba v prípade, že $\deg_G(m) = \Delta$ a $j \in M$ pre všetky $j \in N_G(m)$. Zo súvislosti grafu $\mathcal{G}$ teda $M = [n]$, pričom pre $k = 1, \dots, n$ je $\deg_G(k) = \Delta$ a graf $\mathcal{G}$ je Δ -regulárny. $\square$

Z práve dokázanej vety možno odvodiť aj čisto spektrálnu charakterizáciu regulárnych grafov, ktorá je sformulovaná ako dôsledok 7.5.2 nižšie. Pri jeho dôkaze využijeme aj vetu 4.9.3 – pripomeňme si, že *stopou* štvorcovej matice $A = (a_{i,j})_{n \times n} \in \mathbb{C}^n$ nazývame súčet

$$\text{tr}(A) = a_{1,1} + a_{2,2} + \dots + a_{n,n}$$

jej diagonálnych prvkov a vďaka vete 4.9.3 je

$$\text{tr}(A) = \sum_{\lambda \in \sigma(A)} \alpha(A, \lambda) \lambda;$$

ide teda o súčet všetkých vlastných čísel matice A , kde každé z nich je zarátané aj viackrát podľa jeho algebraickej násobnosti.

⁸Keďže je graf $\mathcal{G}$ súvislý, mohli by sme sa tiež odvolať na Perronovu-Frobeniovu vetu, podľa ktorej musí existovať *kladný* vlastný vektor matice A prislúchajúci k λ_1 ; matica susednosti A súvislého grafu $\mathcal{G}$ totiž musí byť ireducibilná. Odvolávka na Perronovu-Frobeniovu teóriu tu však nie je potrebná.

Dôsledok 7.5.2. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s maticou susednosti $A = \mathbf{A}(\mathcal{G})$ a obyčajným spektrom $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Graf $\mathcal{G}$ je potom regulárny práve vtedy, keď*

$$n\lambda_1 = \lambda_1^2 + \lambda_2^2 + \dots + \lambda_n^2$$

a v takom prípade je λ_1 -regulárny.

Dôkaz. Podľa vety 7.5.1 je graf $\mathcal{G}$ regulárny práve vtedy, keď

$$\lambda_1 = \frac{1}{n} \sum_{j=1}^n \deg_{\mathcal{G}}(j),$$

čiže keď

$$n\lambda_1 = \sum_{j=1}^n \deg_{\mathcal{G}}(j). \quad (7.12)$$

Položme $A = (a_{i,j})_{n \times n}$ a uvažujme maticu $A^2 = (a_{i,j}[2])_{n \times n}$. Pre $j = 1, \dots, n$ je diagonálny prvok $a_{j,j}[2]$ matice A^2 daný ako

$$a_{j,j}[2] = \sum_{k=1}^n a_{j,k} a_{k,j} = \sum_{k=1}^n a_{j,k} = \deg_{\mathcal{G}}(j).$$

Preto

$$\operatorname{tr}(A^2) = \sum_{j=1}^n \deg_{\mathcal{G}}(j)$$

a rovnosť (7.12) tak môžeme prepísať ako

$$n\lambda_1 = \operatorname{tr}(A^2).$$

Pokiaľ sú však $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ vlastné čísla matice A , sú vlastné čísla matice A^2 podľa vety 4.8.4 dané ako $\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2$, kde každé vlastné číslo λ je uvedené $\alpha(A^2, \lambda)$ -krát. Uvedenú rovnosť tak vďaka vete 4.9.3 skutočne možno prepísať do podoby

$$n\lambda_1 = \lambda_1^2 + \lambda_2^2 + \dots + \lambda_n^2,$$

pričom z pozorovaní vyššie vyplýva jej ekvivalencia s regulárnosťou grafu $\mathcal{G}$. Podľa vety 7.5.1 navyše graf $\mathcal{G}$ musí byť λ_1 -regulárny, kedykoľvek je regulárny. $\square$

7.6 Kladne semidefinitné matice

V súvislosti s Rayleighovými podielmi sme často pracovali s výrazmi $\mathbf{x}^T A \mathbf{x}$, kde $A = (a_{i,j})_{n \times n} \in \mathbb{R}^{n \times n}$ je symetrická matica a $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$. Ľahko pridáme k vzťahu

$$\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T (A \mathbf{x}) = \sum_{i=1}^n x_i \sum_{j=1}^n a_{i,j} x_j = \sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j. \quad (7.13)$$

Ak teraz o vektore $\mathbf{x}$ nebudeme uvažovať ako o pevne danom prvku $\mathbb{R}^n$, ale jeho zložky $x_1, \dots, x_n$ interpretujeme ako premenné, môžeme výraz na pravej strane (7.13) chápať ako polynóm o n premenných $x_1, \dots, x_n$ s reálnymi koeficientmi, ktorého všetky členy sú stupňa práve dva. Takéto polynómy sa zvyknú nazývať *kvadratickými formami*.⁹

⁹Podobným spôsobom určuje kvadratickú formu každá – čiže nie nutne iba symetrická – reálna štvorcová matica. Ľahko však vidieť, že každá kvadratická forma je určená práve jednou symetrickou maticou a naopak.

Reálnu symetrickú maticu $A \in \mathbb{R}^{n \times n}$ nazveme *kladne* (alebo *pozitívne*) *semidefinitnou* v prípade, že je polynomickeá funkcia prislúchajúca k ňou určenej kvadratickej forme nezáporná na $\mathbb{R}^n$ – ak teda do nej za premenné $x_1, \dots, x_n$ dosadíme ľubovoľné reálne čísla, dostaneme nezáporné reálne číslo.

Definícia 7.6.1. Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. Matica A je *kladne semidefinitná*, ak pre všetky $\mathbf{x} \in \mathbb{R}^n$ je $\mathbf{x}^T A \mathbf{x} \geq 0$.

Dokážeme teraz vetu, ktorá hovorí o dvoch užitočných ekvivalentných charakterizáciách kladne semidefinitných matic.

Veta 7.6.2. Nech $n \in \mathbb{N}$ a $A \in \mathbb{R}^{n \times n}$ je symetrická matica. Potom sú nasledujúce tvrdenia ekvivalentné:

- (i) Matica A je kladne semidefinitná.
- (ii) Všetky vlastné čísla matice A sú nezáporné.
- (iii) Existuje $m \in \mathbb{N}$ a matica $B \in \mathbb{R}^{m \times n}$ taká, že $A = B^T B$.

Dôkaz. Začnime ekvivalenciou tvrdení (i) a (ii). Uvažujme ľubovoľné – nutne reálne – vlastné číslo λ symetrickej matice A a k nemu prislúchajúci vlastný vektor $\mathbf{v} \in \mathbb{R}^n$. Potom $A\mathbf{v} = \lambda\mathbf{v}$, z čoho

$$\mathbf{v}^T A \mathbf{v} = \lambda \mathbf{v}^T \mathbf{v}.$$

Hodnota $\mathbf{v}^T \mathbf{v} = \|\mathbf{v}\|_2^2$ je pritom vždy kladná. Ak je teda matica A kladne semidefinitná, nutne $\lambda \geq 0$. Z vety 7.2.6 naopak vieme, že reálna symetrická matica A musí byť ortogonálne diagonalizovateľná, a teda existuje ortogonálna matica $P \in \mathbb{R}^{n \times n}$ taká, že pre vlastné čísla $\lambda_1 \geq \dots \geq \lambda_n$ matice A a diagonálnu maticu Λ danú ako

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

je

$$A = P^{-1} \Lambda P = P^T \Lambda P.$$

Ak sú teda všetky vlastné čísla $\lambda_1, \dots, \lambda_n$ nezáporné a $\mathbf{x} \in \mathbb{R}^n$ je ľubovoľný vektor s $P\mathbf{x} = (y_1, \dots, y_n)^T$, je

$$\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T P^T \Lambda P \mathbf{x} = (P\mathbf{x})^T \Lambda (P\mathbf{x}) = \sum_{k=1}^n y_k^2 \lambda_k \geq 0.$$

Matica A je teda kladne semidefinitná.

Zostáva dokázať, že *prvé dve tvrdenia sú ekvivalentné tvrdeniu (iii)*. Ak pre nejaké $m \in \mathbb{N}$ existuje matica $B \in \mathbb{R}^{m \times n}$ spĺňajúca $A = B^T B$, tak pre všetky $\mathbf{x} \in \mathbb{R}^n$ je

$$\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T B^T B \mathbf{x} = (B\mathbf{x})^T B \mathbf{x} = \|B\mathbf{x}\|_2^2 \geq 0$$

a matica A je kladne semidefinitná. Predpokladajme naopak, že sú všetky vlastné čísla $\lambda_1 \geq \dots \geq \lambda_n$ reálnej symetrickej matice A nezáporné. K matici Λ danej rovnako ako vyššie potom môžeme uvažovať maticu

$$\sqrt{\Lambda} = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\lambda_n} \end{pmatrix}.$$

Evidentne $(\sqrt{\Lambda})^2 = \Lambda$. Ak je teda $P \in \mathbb{R}^{n \times n}$ ortogonálna matica taká, že $A = P^{-1}\Lambda P = P^T\Lambda P$ a položíme

$$B := P^T\sqrt{\Lambda}P,$$

tak $B^T = B$ a

$$B^TB = B^2 = P^T\sqrt{\Lambda}PP^T\sqrt{\Lambda}P = P^T\Lambda P = A,$$

čo dokazuje pravdivosť tvrdenia (iii). $\square$

Podobne ako kladne semidefinitné matice možno definovať aj kladne *definitné* matice, pri ktorých je podmienka nezápornosti príslušnej kvadratickej formy nahradená podmienkou jej kladnosti na $\mathbb{R}^n \setminus \{\mathbf{0}\}$ – pre kladne definitnú maticu $A \in \mathbb{R}^{n \times n}$ a všetky $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ teda musí byť $\mathbf{x}^T A \mathbf{x} > 0$. Možno dokázať, že kladná definitnosť reálnej symetrickej matice je ekvivalentná kladnosti všetkých jej vlastných čísel, a to je ďalej ekvivalentné existencii $m \in \mathbb{N}$ a matice $B \in \mathbb{R}^{m \times n}$ plnej stĺpcovej hodnoty n takej, že $A = B^TB$. Dôkaz týchto tvrdení je náplňou cvičenia 7 na konci tejto kapitoly.

7.7 Laplaceova matica grafu

Popri matici susednosti prislúcha ku každému grafu aj jeho takzvaná *Laplaceova matica*, k definícii ktorej sa práve dostávame.¹⁰ Opäť pôjde pre každý graf o symetrickú maticu a jej spektrum budeme nazývať *laplaceovským spektrom* grafu. Po zvyšok tejto kapitoly sa budeme zaoberať predovšetkým súvisom vlastností Laplaceových matíc a laplaceovských spektier grafov s rôznymi ich štrukturálnymi a kombinatorickými vlastnosťami.

Definícia 7.7.1. Nech $n \in \mathbb{N}$ a $\mathcal{G} = ([n], \varphi)$ je graf. *Laplaceovou maticou* grafu $\mathcal{G}$ nazveme maticu $\mathbf{L}(\mathcal{G}) = (q_{i,j})_{n \times n} \in \mathbb{Z}^{n \times n}$ takú, že pre $i, j = 1, \dots, n$ je

$$q_{i,j} = \begin{cases} \deg_{\mathcal{G}}(i) & \text{ak } i = j, \\ -1 & \text{ak } i \neq j \text{ a } \{i, j\} \in E(\mathcal{G}), \\ 0 & \text{inak.} \end{cases}$$

Maticu $\mathbf{L}(\mathcal{G})$ tak možno vyjadriť ako rozdiel $\mathbf{L}(\mathcal{G}) := D - \mathbf{A}(\mathcal{G})$, kde $D = (d_{i,j})_{n \times n} \in \mathbb{N}^{n \times n}$ je diagonálna matica taká, že pre $k = 1, \dots, n$ je $d_{k,k} = \deg_{\mathcal{G}}(k)$.

V literatúre o spektrálnej teórii grafov sa často možno stretnúť aj s rôznymi variantmi takto definovaných Laplaceových matíc – za všetky spomeňme len *normalizované Laplaceove matice*, spektrám ktorých bola venovaná monografia [20].

Význam Laplaceovej matice spočíva okrem iného aj v tom, že ide o maticu veľmi prirodzenej kvadratickej formy. Pre všetky $\mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$ a $\mathbf{L}(\mathcal{G}) = (q_{i,j})_{n \times n}$ je totiž

$$\mathbf{x}^T \mathbf{L}(\mathcal{G}) \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n q_{i,j} x_i x_j = \sum_{k=1}^n \deg_{\mathcal{G}}(k) x_k^2 - \sum_{\{i,j\} \in E(\mathcal{G})} 2x_i x_j = \sum_{\{i,j\} \in E(\mathcal{G})} (x_i - x_j)^2. \quad (7.14)$$

Ľahko teda vidieť, že matica $\mathbf{L}(\mathcal{G})$ je vždy kladne semidefinitná. Z toho podľa vety 7.6.2 vyplýva, že všetky vlastné čísla matice $\mathbf{L}(\mathcal{G})$ sú nezáporné.

¹⁰Ide o zneužitie mena slávneho matematika – tentokrát prostredníctvom analógie s Laplaceovým operátorom, ktorú vysvetlíme nižšie.

Keďže nás budú zaujímať predovšetkým menšie vlastné čísla matice $\mathbf{L}(\mathcal{G})$, budeme laplaceovské spektrum grafu $\mathcal{G}$ o n vrchoch písať ako

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$$

pri interpretácii z poznámky 7.2.7. Dokážeme teraz, že najmenšie vlastné číslo λ_1 matice $\mathbf{L}(\mathcal{G})$ je vždy rovné nule, pričom prislúchajúcim vlastným vektorom je $\mathbf{1} = (1, 1, \dots, 1)^T$.

Tvrdenie 7.7.2. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s Laplaceovou maticou $L = \mathbf{L}(\mathcal{G})$ a laplaceovským spektrom $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Potom $\lambda_1 = 0$, pričom $\mathbf{1} = (1, 1, \dots, 1)^T$ je vlastným vektorom matice L prislúchajúcim k λ_1 .*

Dôkaz. Nech $L = (q_{i,j})_{n \times n}$ a $L\mathbf{1} = (y_1, \dots, y_n)^T$. Pre $i = 1, \dots, n$ potom

$$y_i = \sum_{j=1}^n q_{i,j} = \deg_{\mathcal{G}}(i) + \sum_{\substack{j \in V(\mathcal{G}) \\ \{i,j\} \in E(\mathcal{G})}} -1 = \deg_{\mathcal{G}}(i) - \deg_{\mathcal{G}}(i) = 0.$$

Teda $L\mathbf{1} = (0, 0, \dots, 0)^T = 0 \cdot \mathbf{1}$ a $\mathbf{1}$ je vlastným vektorom matice L prislúchajúcim k vlastnému číslu 0. Keďže sú všetky vlastné čísla kladne semidefinitnej matice L nezáporné, musí byť $0 = \lambda_1$. $\square$

Najmenším zaujímavým spomedzi vlastných čísel $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ Laplaceovej matice grafu $\mathcal{G}$ je tak vždy vlastné číslo λ_2 . Dokážeme teraz, že $\lambda_2 > 0$ práve vtedy, keď je graf $\mathcal{G}$ súvislý; získame tak dôležitú spektrálnu charakterizáciu súvislých grafov.

Veta 7.7.3. *Nech $n \in \mathbb{N} \setminus \{0, 1\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s Laplaceovou maticou $L = \mathbf{L}(\mathcal{G})$ a laplaceovským spektrom $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Potom je graf $\mathcal{G}$ súvislý práve vtedy, keď $\lambda_2 > 0$.*

Dôkaz. Ak je graf $\mathcal{G}$ nesúvislý, možno vrcholy grafu $\mathcal{G}$ prečíslovať tak, aby pre nejaké $m \in [n-1]$ v $\mathcal{G}$ nevedla žiadna hrana spájajúca vrcholy z množín $V_1 = \{1, \dots, m\}$ a $V_2 = \{m+1, \dots, n\}$. Matica L je tak podobná matici L' , ktorú pre nejaké štvorcové bloky L_1, L_2 možno zapísať ako

$$L' = \begin{pmatrix} L_1 & \mathbf{0} \\ \mathbf{0} & L_2 \end{pmatrix}.$$

Matice L_1 resp. L_2 pritom sú – pri vhodnom prečíslovaní vrcholov – Laplaceovými maticami grafov

$$\mathcal{G}_1 = (V_1, \varphi|_{\binom{V_1}{2}}) \quad \text{resp.} \quad \mathcal{G}_2 = (V_2, \varphi|_{\binom{V_2}{2}})$$

a podľa tvrdenia 7.7.2 tak nula musí byť vlastným číslom obidvoch týchto matíc. Algebraická násobnosť vlastného čísla 0 matice L je preto rovná aspoň dvom, z čoho $\lambda_2 = 0$.

Predpokladajme naopak, že je graf $\mathcal{G}$ súvislý a $\mathbf{x} = (x_1, \dots, x_n)^T$ je vlastný vektor matice L prislúchajúci k vlastnému číslu 0. Potom aj $\mathbf{x}^T L \mathbf{x} = 0$ a z rovnosti (7.14) vidieť, že pre $i, j = 1, \dots, n$ musí byť $x_i = x_j$ kedykoľvek $\{i, j\} \in E(\mathcal{G})$. Vďaka súvislosti grafu $\mathcal{G}$ tak musí byť vlastný vektor $\mathbf{x}$ konštantný, z čoho $\gamma(L, 0) = 1$ a z diagonalizovateľnosti reálnej symetrickej matice L aj $\alpha(L, 0) = 1$. Vlastné číslo λ_2 tak nemôže byť nulové, a teda $\lambda_2 > 0$. $\square$

Druhé najmenšie vlastné číslo λ_2 Laplaceovej matice grafu je jeho dôležitou charakteristikou a býva nazývané aj *algebraickým stupňom súvislosti* grafu. Detailnejšie tento parameter grafov preskúmame v rámci oddielu 7.9.

Laplaceova matica každého grafu $\mathcal{G}$ musí byť – pretože je kladne semidefinitná – podľa vety 7.6.2 vyjadriteľná v tvare $\mathbf{L}(\mathcal{G}) = B^T B$ pre nejaké $m \in \mathbb{N}$ a maticu $B \in \mathbb{R}^{m \times n}$. V nasledujúcom ukážeme, že za túto maticu B možno vziať ďalšiu z matíc prirodzene prislúchajúcich ku grafu $\mathcal{G}$ – konkrétne *incidenčnú maticu* ľubovoľnej orientácie grafu $\mathcal{G}$.

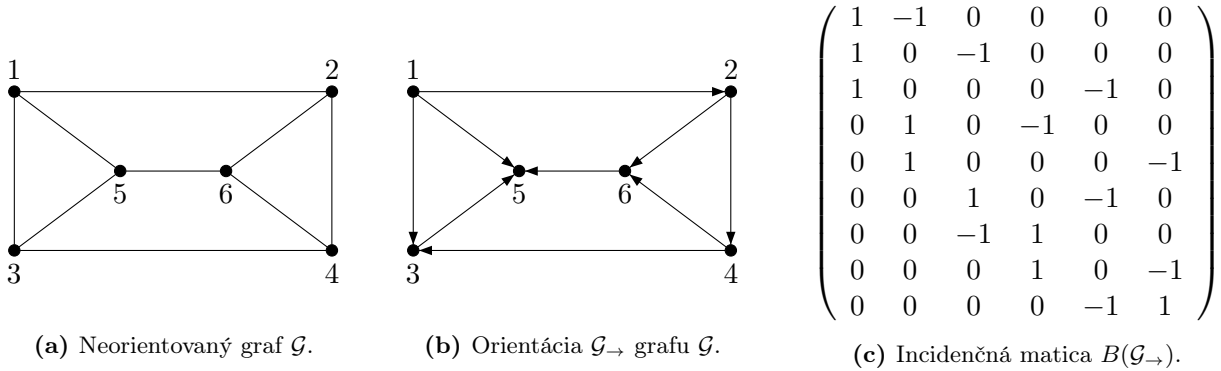
Pod *orientáciou* neorientovaného grafu¹¹ $\mathcal{G} = (V, \varphi)$ budeme rozumieť orientovaný graf, ktorý z $\mathcal{G}$ vznikne zamenou každej jeho neorientovanej hrany za *jednu* orientovanú hranu vedúcu ľubovoľným smerom – čiže orientovaný graf $\mathcal{G}_{\rightarrow} = (V, \varphi_{\rightarrow})$, kde $\varphi_{\rightarrow}: V^2 \rightarrow \mathbb{B}$ je zobrazenie také, že pre všetky $i, j \in V$ je buď $\varphi_{\rightarrow}(i, j) = \varphi(\{i, j\})$ a $\varphi_{\rightarrow}(j, i) = 0$, alebo $\varphi_{\rightarrow}(j, i) = \varphi(\{i, j\})$ a $\varphi_{\rightarrow}(i, j) = 0$.

Pre *orientovaný graf bez slučiek* $\mathcal{G} = (V, \varphi)$ s $|E(\mathcal{G})| = m$ a $E(\mathcal{G}) = \{(u_1, v_1), \dots, (u_m, v_m)\}$ definujeme jeho *incidenčnú maticu*¹² $B(\mathcal{G})$ ako $B(\mathcal{G}) = (b_{i,j})_{m \times n} \in \mathbb{Z}^{m \times n}$, kde $b_{i,j}$ je pre $i = 1, \dots, m$ a $j = 1, \dots, n$ dané ako

$$b_{i,j} = \begin{cases} 1 & \text{ak } j = u_i, \\ -1 & \text{ak } j = v_i, \\ 0 & \text{inak.} \end{cases}$$

Definícia incidenčnej matice tak závisí nielen od usporiadania vrcholov, ale aj od usporiadania hrán uvažovaného grafu. Podstatné vlastnosti tejto matice však zostanú pri ľubovoľnej voľbe usporiadania hrán nezmenené.¹³

Príklad 7.7.4. Na obrázku 7.3 je znázornený diagram neorientovaného grafu $\mathcal{G}$, diagram jednej z jeho orientácií $\mathcal{G}_{\rightarrow}$ a incidenčná matica grafu $\mathcal{G}_{\rightarrow}$.



Obr. 7.3: Neorientovaný graf $\mathcal{G}$, jedna z jeho orientácií $\mathcal{G}_{\rightarrow}$ a incidenčná matica grafu $\mathcal{G}_{\rightarrow}$.

Veta 7.7.5. Nech $n \in \mathbb{N}$, nech $\mathcal{G} = ([n], \varphi)$ je graf a $\mathcal{G}_{\rightarrow} = ([n], \varphi_{\rightarrow})$ je ľubovoľná jeho orientácia. Potom

$$\mathbf{L}(\mathcal{G}) = B(\mathcal{G}_{\rightarrow})^T B(\mathcal{G}_{\rightarrow}).$$

Dôkaz. Označme $m := |E(\mathcal{G})|$, $L = (q_{i,j})_{n \times n} := \mathbf{L}(\mathcal{G})$, $B = (b_{i,j})_{m \times n} := B(\mathcal{G}_{\rightarrow})$ a $B^T B =: (c_{i,j})_{n \times n}$. Dokážeme, že pre $i, j = 1, \dots, n$ je $c_{i,j} = q_{i,j}$.

Uvažujme najprv $i, j \in [n]$ také, že $i \neq j$. Potom

$$c_{i,j} = \sum_{k=1}^m b_{k,i} b_{k,j} = \begin{cases} -1 & \text{ak } \{i, j\} \in E(\mathcal{G}), \\ 0 & \text{inak,} \end{cases}$$

pretože existencia hrany $\{i, j\} \in E(\mathcal{G})$ je ekvivalentná existencii (práve jedného) indexu $k \in [m]$ takého, že $b_{k,i}$ a $b_{k,j}$ sú obidve nenulové; v takom prípade je jeden z týchto prvkov rovný jednej a druhý mínus jednej, z čoho $b_{k,i} b_{k,j} = -1$. Pre $i \neq j$ teda skutočne $c_{i,j} = q_{i,j}$.

¹¹Stále pracujeme výhradne s neorientovanými grafmi *bez slučiek*.

¹²Podobne by sme mohli definovať aj incidenčnú maticu neorientovaného grafu – jediným rozdielom by bolo, že obidvom koncovým vrcholom každej hrany by v matici prislúchala hodnota 1.

¹³Z tohto dôvodu budeme používať označenie $B(\mathcal{G})$ aj bez toho, aby sme usporiadanie hrán explicitne určili – v takom prípade ho bude možné zvoliť ľubovoľne.

Zostáva preskúmať diagonálne prvky matice $B^T B$. Tu ale pre $i = 1, \dots, n$ zrejme dostávame

$$c_{i,i} = \sum_{k=1}^m b_{k,i}^2 = \deg_{\mathcal{G}}(i) = q_{i,i}$$

a veta je dokázaná. $\square$

Práve dokázanou vetou možno vysvetliť pomenovanie *Laplaceova matica*, ktoré sa pre maticu $\mathbf{L}(\mathcal{G})$ používa napriek tomu, že Pierre-Simon Laplace sa pravdepodobne grafmi nikdy nezaoberal. Vďaka vete 7.7.5 možno Laplaceovu maticu grafu chápať ako diskretnú analógiu *Laplaceovho operátora* Δ definovaného pre dvakrát diferencovateľnú funkciu $f: \mathbb{R}^n \rightarrow \mathbb{R}$ premenných $x_1, \dots, x_n$ ako

$$\Delta f = \sum_{k=1}^n \frac{\partial^2 f}{\partial x_k^2} = \operatorname{div}(\operatorname{grad} f),$$

kde operátor *gradientu* je pre diferencovateľnú funkciu $f: \mathbb{R}^n \rightarrow \mathbb{R}$ premenných $x_1, \dots, x_n$ definovaný predpisom

$$\operatorname{grad} f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)^T$$

a operátor *divergencie* je pre vektorové pole $\mathbf{F} = (f_1, \dots, f_n)^T$ pozostávajúce z diferencovateľných funkcií $f_1, \dots, f_n: \mathbb{R}^n \rightarrow \mathbb{R}$ premenných $x_1, \dots, x_n$ definovaný ako

$$\operatorname{div} \mathbf{F} = \sum_{k=1}^n \frac{\partial f_k}{\partial x_k}.$$

Ako notačná skratka sa tu často používa operátor ∇ – takzvaná *nabla*¹⁴ – pod ktorým sa v $\mathbb{R}^n$ rozumie vektor

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right)^T.$$

Možno potom písať

$$\operatorname{grad} f = \nabla f \quad \text{a} \quad \operatorname{div} \mathbf{F} = \nabla^T \mathbf{F}.$$

Pre Laplaceov operátor tak dostávame vyjadrenie

$$\Delta f = \nabla^T \nabla f.$$

Označme teraz ako $B = (b_{i,j})_{m \times n}$ incidenčnú maticu ľubovoľnej orientácie grafu $\mathcal{G}$ o n vrchoch a m hranách a uvažujme vektor $\mathbf{f} = (f_1, \dots, f_n)^T \in \mathbb{R}^n$, ktorý môžeme chápať ako funkciu priradujúcu každému z vrcholov $k \in [n]$ grafu $\mathcal{G}$ nejakú reálnu hodnotu f_k . Pre vektor $B\mathbf{f} = (f'_1, \dots, f'_m)^T \in \mathbb{R}^m$ a $i = 1, \dots, m$ potom

$$f'_i = \sum_{j=1}^n b_{i,j} f_j = f_u - f_v$$

pre $u, v \in [n]$ také, že i -ta hrana uvažovanej orientácie grafu $\mathcal{G}$ vedie z vrcholu u do vrcholu v . Jednotlivé zložky vektora $B\mathbf{f}$ teda udávajú diferencie „funkcie“ $\mathbf{f}$ na jednotlivých hranách a vektor $B\mathbf{f}$ možno chápať ako diskretnú obdobu gradientu funkcie; samotná matica B sa tak stáva analógiou operátora ∇ . Matica $\mathbf{L}(\mathcal{G}) = B^T B$ potom zodpovedá Laplaceovmu operátoru $\Delta = \nabla^T \nabla$.

¹⁴Staroveký hebrejský strunový hudobný nástroj, ktorého tvar symbol ∇ pripomína.

7.8 Enumerácia kostier grafu

Ukážeme si teraz klasickú kombinatorickú aplikáciu Laplaceových matíc grafov – uvidíme, že s ich pomocou možno ľahko nájsť *počet všetkých kostier* daného grafu. Ten bude daný determinantom matice, ktorá vznikne z Laplaceovej matice grafu $\mathcal{G}$ odstránením k -teho riadku a stĺpca pre nejaké $k \in V(\mathcal{G})$.

Pre každé $n \in \mathbb{N}$, štvorcovú maticu $A \in \mathbb{R}^{n \times n}$ a množinu $S \subseteq [n]$ označíme ako $A[S]$ maticu, ktorá vznikne z matice A odobraním všetkých riadkov a stĺpcov s indexmi z množiny S . Ak teda $A = (a_{i,j})_{n \times n}$, je $A[S] = (b_{i,j})_{m \times m}$, kde $m = n - |S|$ a pre $i, j = 1, \dots, m$ je $b_{i,j} = a_{\psi(i), \psi(j)}$, pričom $\psi(k) \in [n]$ je pre $k = 1, \dots, m$ dané ako index taký, že $\psi(k) \notin S$ a $|\{\psi(k)\} \setminus S| = k$. Pre $S = \{s_1, \dots, s_k\}$ budeme namiesto $A[\{s_1, \dots, s_k\}]$ písať len $A[s_1, \dots, s_k]$.

Pripomeňme si tiež skutočnosť, že determinant matice typu 0×0 je rovný jednej – determinant matice $A = (a_{i,j})_{n \times n} \in \mathbb{R}^{n \times n}$ je totiž daný ako

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{k=1}^n a_{k, \pi(k)}, \quad (7.15)$$

kde S_n je symetrická grupa n -prvkovej množiny a $\operatorname{sgn}(\pi)$ označuje znamienko permutácie π . Pre $n = 0$ je symetrická grupa triviálna a pre jediné $\pi_0 \in S_0$ je $\operatorname{sgn}(\pi_0) = 1$. Z výrazu na pravej strane (7.15) sa tak stáva

$$\sum_{\pi \in S_0} \operatorname{sgn}(\pi) \prod_{k=1}^0 a_{k, \pi(k)} = \operatorname{sgn}(\pi_0) \prod_{k=1}^0 a_{k, \pi_0(k)} = 1 \cdot 1 = 1.$$

Veta 7.8.1. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s Laplaceovou maticou $L = \mathbf{L}(\mathcal{G})$. Pre ľubovoľné $k \in [n]$ má potom graf $\mathcal{G}$ práve $\det(L[k])$ rôznych kostier.*

Dôkaz. Nech $k \in [n]$; ľubovoľnú kostru grafu $\mathcal{G}$ potom možno chápať ako zakorenenú vo vrchole k a všetky hrany ako orientované smerom ku koreňu k . Označme množinu takýchto *orientovaných kostier* grafu $\mathcal{G}$ ako $\mathcal{T}_k(\mathcal{G})$. Dokážeme, že $|\mathcal{T}_k(\mathcal{G})| = \det(L[k])$.

Označme ďalej ako $\mathcal{M}_k(\mathcal{G})$ množinu všetkých zobrazení $f: [n] \setminus \{k\} \rightarrow [n]$ takých, že pre všetky $v \in [n] \setminus \{k\}$ je $f(v)$ niektorý spomedzi susedov vrcholu v v grafe $\mathcal{G}$ – je teda $\{v, f(v)\} \in E(\mathcal{G})$. Každé takéto zobrazenie určuje *orientovaný* graf $\mathcal{D}_f$ s množinou vrcholov $[n]$ a s hranami $(v, f(v))$ pre všetky $v \in [n] \setminus \{k\}$: to znamená, že $\mathcal{D}_f = ([n], \psi)$, pričom pre všetky $u, v \in [n]$ je $\psi(u, v) = 1$ práve vtedy, keď $u \neq k$ a $v = f(u)$. Evidentne

$$|\mathcal{M}_k(\mathcal{G})| = \prod_{v \in [n] \setminus \{k\}} \deg_{\mathcal{G}}(v). \quad (7.16)$$

Každá orientovaná kostra z $\mathcal{T}_k(\mathcal{G})$ je jednoznačne daná zobrazením $f \in \mathcal{M}_k(\mathcal{G})$ takým, že pre všetky $v \in [n] \setminus \{k\}$ je $f(v)$ jediný sused vrcholu v ležiaci v príslušnej kostre na ceste z v do k ; v takom prípade je touto orientovanou kostrou graf $\mathcal{D}_f$.

Hoci má pre všetky $f \in \mathcal{M}_k(\mathcal{G})$ graf $\mathcal{D}_f$ presne $n - 1$ hrán, nemusí vždy ísť o orientovanú kostru grafu $\mathcal{G}$. Ľahko ale vidieť, že *graf $\mathcal{D}_f$ reprezentuje orientovanú kostru grafu $\mathcal{G}$ práve vtedy, keď je acyklický*. Stačí teda vyčíslieť počet všetkých zobrazení $f \in \mathcal{M}_k(\mathcal{G})$ s acyklickým grafom $\mathcal{D}_f$ – označme množinu všetkých takýchto zobrazení f ako $\mathcal{Q}_k(\mathcal{G})$. Z uvedeného vyplýva $|\mathcal{T}_k(\mathcal{G})| = |\mathcal{Q}_k(\mathcal{G})|$.

Pre možné *orientované* kružnice v grafoch prislúchajúcich k zobrazeniam $f \in \mathcal{M}_k(\mathcal{G})$ budeme používať nasledujúcu terminológiu a notáciu inšpirovanú permutáciami: pod *prípustným cyklom* budeme rozumieť postupnosť γ zloženú z po dvoch rôznych vrcholov $v_1, \dots, v_s \in [n] \setminus \{k\}$ takú, že $s \in \mathbb{N} \setminus \{0\}$, pre $j = 1, \dots, s - 1$ je $\{v_j, v_{j+1}\} \in E(\mathcal{G})$ a $\{v_s, v_1\} \in E(\mathcal{G})$; keďže uvažujeme grafy bez slučiek, nutne $s \geq 2$. V takom prípade budeme pre cyklus γ písať $\gamma = (v_1 v_2 \dots v_s)$ a $V(\gamma) = \{v_1, \dots, v_s\}$. Pre $f \in \mathcal{M}_k(\mathcal{G})$ budeme navyše hovoriť, že $\gamma = (v_1 v_2 \dots v_s)$ je *prípustným cyklom* v f – alebo že *obsahuje prípustný cyklus* γ – ak pre $j = 1, \dots, s - 1$ je $v_{j+1} = f(v_j)$ a $v_1 = f(v_s)$.

Ďalej budeme hovoriť, že prípustné cykly $\gamma_1, \dots, \gamma_t$, kde $t \in \mathbb{N}$, sú *po dvoch disjunktné*, ak pre všetky $i, j \in [t]$ je $V(\gamma_i) \cap V(\gamma_j) = \emptyset$ kedykoľvek $i \neq j$. Pre $m \in \mathbb{N}$ definujeme $\mathcal{C}_k(\mathcal{G}, m)$ ako množinu všetkých m -prvkových množín $X = \{\gamma_1, \dots, \gamma_m\}$ po dvoch disjunktných prípustných cyklov $\gamma_1, \dots, \gamma_m$; pre každú takúto množinu $X = \{\gamma_1, \dots, \gamma_m\}$ tiež píšeme $V(X) = V(\gamma_1) \cup \dots \cup V(\gamma_m)$. Pre všetky $m \in \mathbb{N}$ a $X \in \mathcal{C}_k(\mathcal{G}, m)$ napokon definujeme $\mathcal{M}_k(\mathcal{G}, X)$ ako množinu práve všetkých $f \in \mathcal{M}_k(\mathcal{G})$ takých, že f obsahuje všetky prípustné cykly z X (a prípadne aj ďalšie prípustné cykly).

Keďže evidentne

$$\mathcal{Q}_k(\mathcal{G}) = \mathcal{M}_k(\mathcal{G}) \setminus \left(\bigcup_{m=1}^n \bigcup_{X \in \mathcal{C}_k(\mathcal{G}, m)} \mathcal{M}_k(\mathcal{G}, X) \right),$$

vďaka princípu zapojenia a vypojenia dostávame

$$|\mathcal{Q}_k(\mathcal{G})| = |\mathcal{M}_k(\mathcal{G})| + \sum_{m=1}^n (-1)^m \sum_{X \in \mathcal{C}_k(\mathcal{G}, m)} |\mathcal{M}_k(\mathcal{G}, X)|. \quad (7.17)$$

Pre všetky $m \in \mathbb{N} \setminus \{0\}$ a $X \in \mathcal{C}_k(\mathcal{G}, m)$ ale zrejme

$$|\mathcal{M}_k(\mathcal{G}, X)| = \prod_{v \in ([n] \setminus \{k\}) \setminus V(X)} \deg_{\mathcal{G}}(v);$$

z rovností (7.16) a (7.17) teda dostávame

$$\begin{aligned} |\mathcal{T}_k(\mathcal{G})| &= |\mathcal{Q}_k(\mathcal{G})| = \prod_{v \in [n] \setminus \{k\}} \deg_{\mathcal{G}}(v) + \sum_{m=1}^n \sum_{X \in \mathcal{C}_k(\mathcal{G}, m)} (-1)^m \prod_{v \in ([n] \setminus \{k\}) \setminus V(X)} \deg_{\mathcal{G}}(v) = \\ &= \sum_{m=0}^n \sum_{X \in \mathcal{C}_k(\mathcal{G}, m)} (-1)^m \prod_{v \in ([n] \setminus \{k\}) \setminus V(X)} \deg_{\mathcal{G}}(v). \end{aligned}$$

Ak ďalej označíme

$$\mathcal{C}_k(\mathcal{G}) = \bigcup_{m=0}^n \mathcal{C}_k(\mathcal{G}, m),$$

je každá množina $X \in \mathcal{C}_k(\mathcal{G})$ prvkom $\mathcal{C}_k(\mathcal{G}, |X|)$; uvedenú rovnosť tak možno prepísať aj ako

$$|\mathcal{T}_k(\mathcal{G})| = |\mathcal{Q}_k(\mathcal{G})| = \sum_{X \in \mathcal{C}_k(\mathcal{G})} (-1)^{|X|} \prod_{v \in ([n] \setminus \{k\}) \setminus V(X)} \deg_{\mathcal{G}}(v). \quad (7.18)$$

Ak ale na druhej strane $L = (q_{i,j})_{n \times n}$ a ak $S'_{n,k}$ označuje množinu všetkých permutácií $\pi: [n] \rightarrow [n]$ s pevným bodom $k = \pi(k)$, je

$$\det(L[k]) = \sum_{\pi \in S'_{n,k}} \operatorname{sgn}(\pi) \prod_{j \in [n] \setminus \{k\}} q_{j, \pi(j)}. \quad (7.19)$$

Nenulovou hodnotou k tomuto súčtu evidentne prispievajú iba tie permutácie $\pi \in S'_{n,k}$, ktoré pozostávajú z niekoľkých pevných bodov a z niekoľkých cyklov zodpovedajúcich prípustným cyklom v zmysle zavedenom vyššie. Označme množinu všetkých takýchto permutácií ako $S''_{n,k}$. Ak $j \in [n] \setminus \{k\}$ nie je pevným bodom takejto permutácie, nutne $q_{j, \pi(j)} = -1$. Z toho vyplýva, že cykly permutácie $\pi \in S''_{n,k}$ párnej dĺžky prispievajú do súčtu v uvedenom výraze faktorom 1; cykly nepárnej dĺžky rôzne od pevných bodov prispievajú faktorom -1 . Cykly párnej dĺžky ďalej prispievajú faktorom -1 k znamienku permutácie $\operatorname{sgn}(\pi)$ a cykly nepárnej dĺžky k tomuto znamienku prispievajú faktorom 1. Vo výsledku tak každý cyklus permutácie π rôzny od pevného bodu prispieje k príslušnému členu sumy faktorom -1 . Keďže ďalej

pevné body permutácie π – označme ich množinu ako $\text{Fix}(\pi)$ – nemenia znamienko $\text{sgn}(\pi)$ a k súčinu v uvedenom výraze evidentne prispievajú faktorom

$$\prod_{v \in \text{Fix}(\pi) \setminus \{k\}} \deg_{\mathcal{G}}(v),$$

možno rovnosť (7.19) prepísať ako

$$\det(L[k]) = \sum_{\pi \in S''_{n,k}} (-1)^{c(\pi)} \prod_{v \in \text{Fix}(\pi) \setminus \{k\}} \deg_{\mathcal{G}}(v),$$

kde $c(\pi)$ je počet cyklov permutácie π rôznych od pevných bodov. Keďže je napokon každá permutácia $\pi \in S''_{n,k}$ jednoznačne určená množinou jej cyklov rôznych od pevných bodov a tie zodpovedajú príпустným cyklom zavedeným vyššie, dostávame bijekciu $\Phi: S''_{n,k} \rightarrow \mathcal{C}_k(\mathcal{G})$, pričom pre všetky $\pi \in S''_{n,k}$ je $c(\pi) = |\Phi(\pi)|$ a $\text{Fix}(\pi) \setminus \{k\} = ([n] \setminus \{k\}) \setminus V(\Phi(\pi))$. To znamená, že

$$\begin{aligned} \det(L[k]) &= \sum_{\pi \in S''_{n,k}} (-1)^{|\Phi(\pi)|} \prod_{v \in ([n] \setminus \{k\}) \setminus V(\Phi(\pi))} \deg_{\mathcal{G}}(v) = \\ &= \sum_{X \in \mathcal{C}_k(\mathcal{G})} (-1)^{|X|} \prod_{v \in ([n] \setminus \{k\}) \setminus V(X)} \deg_{\mathcal{G}}(v), \end{aligned}$$

čo je podľa (7.18) rovné hľadanej hodnote $|\mathcal{T}_k(\mathcal{G})| = |\mathcal{Q}_k(\mathcal{G})|$. $\square$

Ukážme si ešte jeden dôkaz vety 7.8.1, pri ktorom budeme nútení pracovať s *multigrafmi*, využívajúc fakt, že grafy sú ich špeciálnym prípadom – operácia *kontrakcie* podľa hrany, ktorú v nasledujúcom budeme hojne používať, totiž môže vyrobiť multigraf aj z jednoduchého grafu. Vo výsledku tak ale okrem alternatívneho dôkazu vety 7.8.1 dostaneme aj jej zovšeobecnenie na prípad multigrafu. Prezentácia nasledujúcich partií tohto oddielu je čiastočne založená na [49] a na [82].

Keďže slučky nijako neovplyvňujú počet kostier grafu, *budeme odteraz aj pod multigrafom vždy rozumieť multigraf bez slučiek*. Pre multigraf $\mathcal{G} = (V, \varphi)$ tak budeme písať aj $\varphi: \binom{V}{2} \rightarrow \mathbb{N}$.

Pre daný (neorientovaný) multigraf $\mathcal{G} = (V, \varphi)$ a hranu $e \in E(\mathcal{G})$ budeme ako $\mathcal{G} \setminus e$ označovať multigraf, ktorý z $\mathcal{G}$ vznikne odobraním hrany e – pre $e = (\{u, v\}, k)$ je teda $\mathcal{G} \setminus e = (V, \varphi')$, kde $\varphi': \binom{V}{2} \rightarrow \mathbb{N}$ je dané ako $\varphi'(\{u, v\}) = \varphi(\{u, v\}) - 1$ a $\varphi'(\{x, y\}) = \varphi(\{x, y\})$ pre zvyšné $\{x, y\}$ z definičného oboru.

Ako $\mathcal{G} : e$ označíme *kontrakciu* multigrafu $\mathcal{G}$ podľa e – čiže multigraf, ktorý vznikne stotožnením koncových vrcholov hrany e , pričom hrana e je spolu so všetkými s ňou súbežnými hranami z multigrafu odstránená a všetky ostatné hrany incidentné s niektorým z jej koncových vrcholov sú presmerované do jediného vrcholu, ktorý je ich stotožnením.¹⁵ Formálne je multigraf $\mathcal{G} : e$ pre $e = (\{u, v\}, k)$ daný ako dvojica¹⁶ $\mathcal{G} : e = (U, \varphi'')$, kde $U = (V \setminus \{u, v\}) \cup \{e\}$ a $\varphi'': \binom{U}{2} \rightarrow \mathbb{N}$ je pre všetky $\{x, y\} \in \binom{U}{2}$ dané takto:

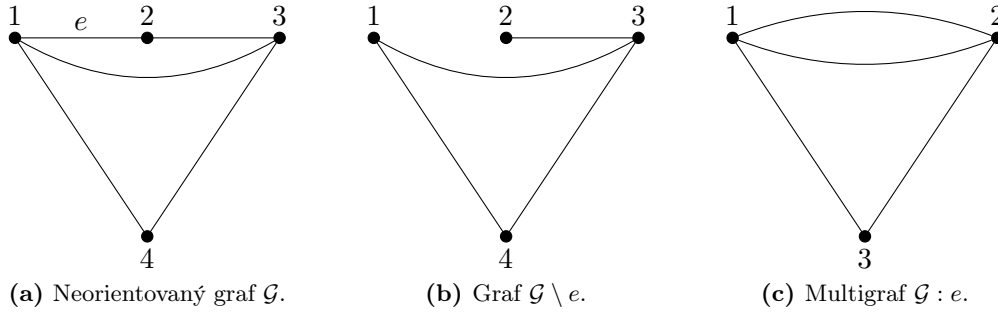
$$\varphi''(\{x, y\}) = \begin{cases} \varphi(\{x, u\}) + \varphi(\{x, v\}) & \text{ak } x \neq y = e, \\ \varphi(\{x, y\}) & \text{ak } x \neq e \text{ a } y \neq e. \end{cases}$$

Keďže na názvoch vrcholov typicky nezáleží, budeme obvykle predpokladať, že z multigrafu o vrchoch $1, \dots, n$ vznikne po kontrakcii podľa niektorej hrany multigraf s vrcholmi $1, \dots, n-1$. Napríklad môžeme v opísanej konštrukcii premenovať vrchol e na $\min\{u, v\}$ a vrcholy $k > \max\{u, v\}$ na $k-1$.

Príklad 7.8.2. Na obrázku 7.4 je znázornený diagram neorientovaného grafu $\mathcal{G}$ s vyznačenou hranou e , ako aj graf $\mathcal{G} \setminus e$ a multigraf $\mathcal{G} : e$.

¹⁵Keby sme uvažovali multigrafy so slučkami, bolo by prirodzené transformovať hrany súbežné s hranou e na slučky v novovytvorenom vrchole.

¹⁶Bez ujmy na všeobecnosti predpokladáme, že $V \cap E = \emptyset$.



Obr. 7.4: Odobranie hrany e z grafu a kontrakcia grafu podľa hrany e .

V nasledujúcom tiež budeme pracovať s Laplaceovými maticami multigrafov. Tie možno pre každý multigraf $\mathcal{G}$ o n vrchoch definovať v súlade s definíciou 7.7.1 ako

$$\mathbf{L}(\mathcal{G}) = D - \mathbf{A}(\mathcal{G}),$$

kde $D = (d_{i,j})_{n \times n}$ je diagonálna matica taká, že pre $k = 1, \dots, n$ je $d_{k,k} = \deg_{\mathcal{G}}(k)$ pre prirodzené multigrafové rozšírenie pojmu stupňa vrcholu. Ak teda $\mathbf{A}(\mathcal{G}) = (a_{i,j})_{n \times n}$ a $\mathbf{L}(\mathcal{G}) = (q_{i,j})_{n \times n}$, je $q_{i,j} = -a_{i,j}$ pre $i, j = 1, \dots, n$ také, že $i \neq j$ a pre $k = 1, \dots, n$ je

$$q_{k,k} = \sum_{i=1}^n a_{i,k} = \sum_{j=1}^n a_{k,j}.$$

Nasledujúca veta sformulovaná pre multigrafy (bez slučiek) tak špeciálne hovorí aj o grafoch.

Veta 7.8.3. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $\mathcal{G} = ([n], \varphi)$ je multigraf s Laplaceovou maticou $L = \mathbf{L}(\mathcal{G})$. Pre ľubovoľné $k \in [n]$ má potom multigraf $\mathcal{G}$ práve $\det(L[k])$ rôznych kostier.*

Dôkaz. Položme $L = (q_{i,j})_{n \times n}$ a počet kostier multigrafu $\mathcal{G}$ označme ako $\tau(\mathcal{G})$. Vetu dokážeme indukciou vzhľadom na počet hrán multigrafu $\mathcal{G}$, čiže na $m = |E(\mathcal{G})|$.

Pre $m = 0$ a $n \geq 2$ neexistuje v multigrafe $\mathcal{G}$ žiadna kostra, čo súhlasí s tým, že pre každé $k \in [n]$ je $L[k]$ nulová matica typu $(n-1) \times (n-1)$ a jej determinant je tým pádom nulový. Pre $m = 0$ a $n = 1$ existuje v $\mathcal{G}$ jediná kostra, ktorá je rovnako ako multigraf $\mathcal{G}$ daná samotným vrcholom 1. Matica $L[1]$ je v takom prípade typu 0×0 a jej determinant je tiež rovný jednej.

Predpokladajme teraz, že dokazované tvrdenie platí pre všetky multigrafy s $m \leq s \in \mathbb{N}$ hranami a uvažujme multigraf $\mathcal{G} = ([n], \varphi)$ taký, že $|E(\mathcal{G})| = s + 1$. Keďže $\mathcal{G}$ obsahuje aspoň jednu hranu, nutne $n \geq 2$. Zvoľme $k \in [n]$ ľubovoľne. Dokážeme, že $\tau(\mathcal{G}) = \det(L[k])$.

Ak v $\mathcal{G}$ neexistuje žiadna hrana incidentná s vrcholom k , je multigraf $\mathcal{G}$ nutne nesúvislý a nemá tak žiadnu kostru. Na druhej strane je z definície Laplaceovej matice multigrafu ľahko vidieť, že pre všetky $j \in [n]$ je súčet prvkov v j -tom stĺpci matice L nulový:

$$\sum_{i=1}^n q_{i,j} = 0. \quad (7.20)$$

Ak v multigrafe $\mathcal{G}$ neexistuje žiadna hrana incidentná s vrcholom k , musí byť k -ty riadok matice L nulový a z rovnosti (7.20) vyplýva, že pre všetky $j \in [n-1]$ je nulový aj súčet prvkov v j -tom stĺpci matice $L[k]$. To znamená, že súčet všetkých riadkov matice $L[k]$ je nulový vektor a matica $L[k]$ je tým pádom singulárna. Preto $\det(L[k]) = 0 = \tau(\mathcal{G})$ a tvrdenie je v tomto prípade dokázané.

Predpokladajme teda, že v multigrafe $\mathcal{G}$ existuje aspoň jedna hrana e incidentná s vrcholom k , ktorá vrchol k spája s nejakým ďalším vrcholom $\ell \in [n]$. Očividne potom existuje bijekcia medzi kostrami

multigrafu $\mathcal{G}$ neobsahujúcimi hranu e a kostrami multigrafu $\mathcal{G} \setminus e$. Podobne môžeme pozorovať bijekciu medzi kostrami $\mathcal{G}$ obsahujúcimi hranu e a kostrami $\mathcal{G} : e$. Je teda

$$\tau(\mathcal{G}) = \tau(\mathcal{G} \setminus e) + \tau(\mathcal{G} : e), \quad (7.21)$$

pričom oba multigrafy $\mathcal{G} \setminus e$ a $\mathcal{G} : e$ majú aspoň o jednu hranu menej, než multigraf $\mathcal{G}$. Na počty ich kostier sa teda vzťahuje indukčný predpoklad, čo je pozorovanie, ktoré zakrátko využijeme.

Laplaceova matica multigrafu $\mathcal{G} \setminus e$ sa od Laplaceovej matice L multigrafu $\mathcal{G}$ líši iba na pozíciách (k, k) , (k, ℓ) , (ℓ, k) a (ℓ, ℓ) . Ľahko teda vidieť, že

$$L[k] = \mathbf{L}(\mathcal{G})[k] = \mathbf{L}(\mathcal{G} \setminus e)[k] + E_\ell[k],$$

kde $E_\ell = (e_{i,j})_{n \times n}$ je matica taká, že $e_{\ell,\ell} = 1$ a $e_{i,j} = 0$ pre všetky ostatné $i, j \in [n]$. Z toho

$$\begin{aligned} \det(L[k]) &= \det(\mathbf{L}(\mathcal{G} \setminus e)[k]) + \det(\mathbf{L}(\mathcal{G} \setminus e)[k, \ell]) = \\ &= \det(\mathbf{L}(\mathcal{G} \setminus e)[k]) + \det(L[k, \ell]) = \\ &= \det(\mathbf{L}(\mathcal{G} \setminus e)[k]) + \det(\mathbf{L}(\mathcal{G} : e)[j]), \end{aligned} \quad (7.22)$$

kde $j = \min\{k, \ell\}$ – posledná rovnosť teda platí za predpokladu, že hranu e kontrahujeme „do vrcholu s menším číslom“, pričom vrcholy s pôvodnými číslami väčšími ako $\max\{k, \ell\}$ dostanú číslo o jedna menšie. Vďaka indukčnému predpokladu a (7.21) potom možno rovnosť (7.22) prepísať ako

$$\det(L[k]) = \tau(\mathcal{G} \setminus e) + \tau(\mathcal{G} : e) = \tau(\mathcal{G}),$$

čo bolo treba dokázať. □

Jednoduchým dôsledkom vety o počte kostier grafu, ktorú sme práve dvoma rôznymi spôsobmi dokázali, je aj nasledujúci *Cayleyho vzorec* udávajúci počet stromov o n (označených) vrchoch. Takéto stromy totiž možno chápať ako kostry úplného grafu o n vrchoch.

Dôsledok 7.8.4 (Cayleyho vzorec). *Nech $n \in \mathbb{N} \setminus \{0\}$. Potom existuje práve n^{n-2} stromov s množinou vrcholov $[n]$.*

Dôkaz. Tvrdenie je očividne pravdivé pre $n = 1$; predpokladajme teda, že $n \geq 2$. Podľa vety 7.8.1 je hľadaný počet stromov daný ako

$$\tau(K_n) = \det(L[1]),$$

kde K_n je úplný graf o n vrchoch, $\tau(K_n)$ označuje jeho počet kostier a $L = \mathbf{L}(K_n)$. Potrebujeme teda vypočítať determinant štvorcovej matice $L[1]$ typu $(n-1) \times (n-1)$, ktorá je očividne daná ako

$$L[1] = \begin{pmatrix} n-1 & -1 & -1 & \dots & -1 \\ -1 & n-1 & -1 & \dots & -1 \\ -1 & -1 & n-1 & \dots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \dots & n-1 \end{pmatrix}.$$

Pripočítanie konštantného násobku niektorého riadku resp. stĺpca k *inému* riadku resp. stĺpcu matice nemení jej determinant. Bezo zmeny determinantu tak môžeme od riadkov 2, ..., $n-1$ matice $L[1]$ odpočítať jej prvý riadok, čím dostávame maticu

$$L'[1] = \begin{pmatrix} n-1 & -1 & -1 & \dots & -1 \\ -n & n & 0 & \dots & 0 \\ -n & 0 & n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -n & 0 & 0 & \dots & n \end{pmatrix}.$$

Ďalej môžeme k prvému stĺpcu matice $L'[1]$ postupne pripočítať jej stĺpce $2, \dots, n-1$, čím dostávame maticu

$$L''[1] = \begin{pmatrix} 1 & -1 & -1 & \dots & -1 \\ 0 & n & 0 & \dots & 0 \\ 0 & 0 & n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n \end{pmatrix}.$$

Počet stromov o n vrcholoch je teda daný ako determinant tejto hornej trojuholníkovej matice:

$$\tau(K_n) = \det(L''[1]) = 1 \cdot n^{n-2} = n^{n-2},$$

čo bolo treba dokázať. $\square$

Cayleyho vzorec je oproti vete 7.8.1 o dosť elementárnejšie tvrdenie a možno ho dokázať aj jednoduchšími metódami. Päť rôznych dôkazov tohto výsledku možno nájsť v [82].

V závere tohto oddielu ešte z vety 7.8.1 odvodíme alternatívny vzorec pre počet kostier grafu $\mathcal{G}$, v ktorom budú vystupovať iba vlastné čísla jeho Laplaceovej matice – vrátime sa tak na pôdu spektrálnej teórie grafov v pravom slova zmysle. Uvažujme najprv ľubovoľnú dvojicu matíc $A, B \in \mathbb{R}^{n \times n}$ pre nejaké $n \in \mathbb{N}$. Pre každé $S \subseteq [n]$ označme ako A_S maticu, ktorá vznikne z A nahradením riadkov s indexmi z množiny S príslušnými riadkami matice B ; maticu B_S môžeme definovať symetricky. Ľahko potom vidieť, že

$$\det(A + B) = \sum_{S \subseteq [n]} \det(A_S) = \sum_{S \subseteq [n]} \det(B_S).$$

Ak teraz pre ľubovoľnú maticu $A \in \mathbb{R}^{n \times n}$ zvolíme za B maticu $B = -z\mathbf{I}$, z uvedeného vzťahu vyplynie, že koeficient pri z^k v charakteristickom polynóme $\text{ch}_A(z)$ matice A bude pre $k = 0, \dots, n$ až na znamienko daný súčtom k -tych minorov matice A :

$$[z^k] \text{ch}_A(z) = [z^k] ((-1)^n \det(A - z\mathbf{I})) = (-1)^{n+k} \sum_{\substack{S \subseteq [n] \\ |S|=k}} \det(A[S]).$$

Toto pozorovanie využijeme pri dôkaze nasledujúcej vety.

Veta 7.8.5. *Nech $n \in \mathbb{N} \setminus \{0\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s Laplaceovou maticou $L = \mathbf{L}(\mathcal{G})$ a laplaceovským spektrom $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Počet kostier $\tau(\mathcal{G})$ grafu $\mathcal{G}$ je potom daný ako*

$$\tau(\mathcal{G}) = \frac{1}{n} \prod_{k=2}^n \lambda_k.$$

Dôkaz. Uvažujme charakteristický polynóm Laplaceovej matice L , ktorý vďaka tvrdeniu 4.2.7 možno vyjadriť ako

$$\text{ch}_L(z) = \det(z\mathbf{I} - L) = (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_n).$$

Podľa tvrdenia 7.7.2 je $\lambda_1 = 0$ a konštantný koeficient polynómu $\text{ch}_L(z)$ je tým pádom nulový. Koeficient pri z je navyše daný ako

$$[z^1] \text{ch}_L(z) = (-1)^{n-1} \prod_{k=2}^n \lambda_k.$$

Z pozorovaní učených bezprostredne pred dokazovanou vetou a z vety 7.8.1 však vyplýva, že súčasne

$$[z^1] \text{ch}_L(z) = (-1)^{n+1} \sum_{\substack{S \subseteq [n] \\ |S|=1}} \det(L[S]) = (-1)^{n+1} \sum_{k \in [n]} \det(L[k]) = (-1)^{n+1} n \tau(\mathcal{G}).$$

Preto

$$(-1)^{n+1} n \tau(\mathcal{G}) = (-1)^{n-1} \prod_{k=2}^n \lambda_k,$$

z čoho

$$\tau(\mathcal{G}) = \frac{1}{n} \prod_{k=2}^n \lambda_k$$

a veta je dokázaná. $\square$

7.9 Algebraický stupeň súvislosti grafu

S *algebraickým stupňom súvislosti* grafu – čiže druhým najmenším vlastným číslom λ_2 jeho Laplaceovej matice – sme sa už stretli v súvislosti s vetou 7.7.3, podľa ktorej je graf súvislý práve vtedy, keď $\lambda_2 > 0$. Toto pozorovanie možno o trochu zosilniť a dokázať, že počet komponentov súvislosti grafu je vždy rovný počtu nulových vlastných čísel jeho Laplaceovej matice; túto nie príliš zložitú úlohu prenechávame čitateľovi ako cvičenie 9. Algebraický stupeň súvislosti grafu, podľa českého matematika Miroslava Fiedlera nazývaný aj *Fiedlerovým číslom*, však nemožno redukovať len na „binárny indikátor“ súvislosti grafu. V nasledujúcom ukážeme, že hodnota vlastného čísla λ_2 obsahuje cennú informáciu o *miere* súvislosti grafu; neskôr preskúmame aj príslušné vlastné vektory.

Najmenším možným algebraickým stupňom súvislosti grafu je teda nula – a to v prípade, že je graf nesúvislý. V nasledujúcich dvoch tvrdeniach identifikujeme *najväčší* možný algebraický stupeň súvislosti n -vrcholového grafu: najprv nájdeme algebraický stupeň súvislosti *úplného* grafu o n vrchoch a následne dokážeme, že faktory¹⁷ grafov nikdy nemôžu mať vyšší algebraický stupeň súvislosti, než pôvodný graf. Druhé z týchto tvrdení je okrem iného zásadné aj pre interpretáciu vlastného čísla λ_2 ako miery súvislosti grafu.

Tvrdenie 7.9.1. *Nech $n \in \mathbb{N} \setminus \{0, 1\}$ a nech $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ tvoria laplaceovské spektrum úplného grafu K_n . Potom $\lambda_1 = 0$ a $\lambda_2 = \dots = \lambda_n = n$. Algebraický stupeň súvislosti grafu K_n je teda $\lambda_2 = n$.*

Dôkaz. Budeme postupovať veľmi podobne ako pri dôkaze Cayleyho vzorca – ten sme vyslovili ako dôsledok 7.8.4. Nech $L := \mathbf{L}(K_n)$. Potom

$$\text{ch}_L(z) = \det(z\mathbf{I} - L) = \det \begin{pmatrix} z - n + 1 & 1 & 1 & \dots & 1 \\ 1 & z - n + 1 & 1 & \dots & 1 \\ 1 & 1 & z - n + 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & z - n + 1 \end{pmatrix},$$

kde matica napravo je typu $n \times n$. Bezo zmeny determinantu môžeme najprv odpočítať prvý riadok tejto matice od všetkých ostatných a následne pripočítať k prvému stĺpcu postupne všetky ostatné. Zisťujeme tak, že

$$\text{ch}_L(z) = \det \begin{pmatrix} z - n + 1 & 1 & 1 & \dots & 1 \\ n - z & z - n & 0 & \dots & 0 \\ n - z & 0 & z - n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n - z & 0 & 0 & \dots & z - n \end{pmatrix} =$$

¹⁷Podgrafy s nezmenenou množinou vrcholov.

$$= \det \begin{pmatrix} z & 1 & 1 & \dots & 1 \\ 0 & z-n & 0 & \dots & 0 \\ 0 & 0 & z-n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & z-n \end{pmatrix}$$

a keďže v poslednom prípade ide o determinant hornej trojuholníkovej matice, je

$$\text{ch}_L(z) = z(z-n)^{n-1}.$$

Skutočne teda $\lambda_1 = 0$ a $\lambda_2 = \dots = \lambda_n = n$. □

Tvrdenie 7.9.2. *Nech $n \in \mathbb{N} \setminus \{0, 1\}$, $\mathcal{G} = ([n], \varphi)$ je graf s algebraickým stupňom súvislosti λ_2 a $\mathcal{G}' = ([n], \varphi')$ je jeho faktor s algebraickým stupňom súvislosti λ'_2 . Potom $\lambda'_2 \leq \lambda_2$.*

Dôkaz. Označme $L := \mathbf{L}(\mathcal{G})$ a $L' := \mathbf{L}(\mathcal{G}')$. Z rovnosti (7.14) ľahko vidieť, že pre všetky $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ musí byť

$$R_L(\mathbf{x}) = \frac{\mathbf{x}^T L \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \geq \frac{\mathbf{x}^T L' \mathbf{x}}{\mathbf{x}^T \mathbf{x}} = R_{L'}(\mathbf{x}).$$

Podľa Courantovej-Fischerovej vety¹⁸ je potom

$$\lambda_2 = \min_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=2}} \max_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_L(\mathbf{x}) \geq \min_{\substack{\mathcal{V} \text{ je podpriestor } \mathbb{R}^n \\ \dim(\mathcal{V})=2}} \max_{\mathbf{x} \in \mathcal{V} \setminus \{\mathbf{0}\}} R_{L'}(\mathbf{x}) = \lambda'_2,$$

čo bolo treba dokázať. □

Keďže je každý graf $\mathcal{G}$ s n vrcholmi faktorom úplného grafu K_n , z predchádzajúcich dvoch tvrdení vyplýva, že pre ľubovoľný takýto graf jeho algebraický stupeň súvislosti λ_2 spĺňa nerovnosť $0 \leq \lambda_2 \leq n$.

Nasledujúca lema, ktorú vzápätí využijeme pri porovnaní algebraického stupňa súvislosti so stupňom vrcholovej súvislosti grafu, dáva do súvisu algebraický stupeň súvislosti grafu $\mathcal{G} = ([n], \varphi)$ s algebraickým stupňom súvislosti jeho indukovaného podgrafu, ktorý z grafu $\mathcal{G}$ vznikne odobraním jedného vrcholu. Na príklade grafu pozostávajúceho z úplného podgrafu a jedného izolovaného vrcholu ľahko vidieť, že odobraním jedného vrcholu z n -vrcholového grafu možno jeho algebraický stupeň súvislosti zvýšiť až o maximálnu možnú hodnotu $n-1$. V nasledujúcej leme však ukážeme, že algebraický stupeň súvislosti takto nikdy príliš *neznížime*. Dôkaz je prevzatý z knihy [17].

Lema 7.9.3. *Nech $n \geq 3$ je prirodzené číslo a $\mathcal{G} = ([n], \varphi)$ je graf s Laplaceovou maticou $L = \mathbf{L}(\mathcal{G})$ a algebraickým stupňom súvislosti λ_2 . Nech $\mathcal{G}_k$ je indukovaný podgraf grafu $\mathcal{G}$, ktorý vznikne z grafu $\mathcal{G}$ odobraním vrcholu $k \in [n]$ a všetkých s ním incidentných hrán.¹⁹ Označme algebraický stupeň súvislosti grafu $\mathcal{G}_k$ ako $\lambda_2^{(k)}$. Potom $\lambda_2^{(k)} \geq \lambda_2 - 1$.*

Dôkaz. Predpokladajme najprv, že vrchol k je v $\mathcal{G}$ spojený hranou so všetkými ostatnými vrcholmi – pre všetky $j \in V(\mathcal{G}) \setminus \{k\}$ je teda $\{k, j\} \in E(\mathcal{G})$. Laplaceova matica L_k grafu $\mathcal{G}_k$ v takom prípade vznikne z Laplaceovej matice L grafu $\mathcal{G}$ odobraním k -teho riadku a k -teho stĺpca a znížením všetkých diagonálnych prvkov o jedna, čo zodpovedá o jedna nižšiemu stupňu všetkých vrcholov. Je teda

$$L_k := \mathbf{L}(\mathcal{G}_k) = \mathbf{L}(\mathcal{G})[k] - \mathbf{I} = L[k] - \mathbf{I}.$$

Preto

$$L_k + \mathbf{I} = L[k]$$

¹⁸Nami uvažované vlastné číslo λ_2 je vo vete 7.3.3 označené ako λ_{n-1} .

¹⁹Predpokladáme pritom, že vrcholy $k+1, \dots, n$ pôvodného grafu sú po odobraní vrcholu k prečíslované na $k, \dots, n-1$ tak, aby množina vrcholov výsledného grafu bola $[n-1]$ a mohli sme tak uvažovať k nemu prislúchajúcej matice.

a vlastné čísla $\lambda'_1 \leq \lambda'_2 \leq \dots \leq \lambda'_{n-1}$ matice $L_k + \mathbf{I}$ tak možno pomocou Cauchyho vety o prepletaní odhadnúť vlastnými číslami $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ matice L :

$$\lambda_1 \leq \lambda'_1 \leq \lambda_2 \leq \lambda'_2 \leq \lambda_3 \leq \dots \leq \lambda_{n-1} \leq \lambda'_{n-1} \leq \lambda_n. \quad (7.23)$$

Podľa vety 4.8.3 ale súčasne

$$\text{ch}_{L_k + \mathbf{I}}(z) = \text{ch}_{L_k}(z - 1),$$

z čoho ďalej vyplýva, že vlastné čísla $\lambda_1^{(k)} \leq \lambda_2^{(k)} \leq \dots \leq \lambda_{n-1}^{(k)}$ Laplaceovej matice L_k grafu $\mathcal{G}_k$ sú pre $j = 1, \dots, n$ dané ako $\lambda_j^{(k)} = \lambda'_j - 1$. Nerovnosti (7.23) tak možno prepísať ako

$$\lambda_1 \leq \lambda_1^{(k)} + 1 \leq \lambda_2 \leq \lambda_2^{(k)} + 1 \leq \lambda_3 \leq \dots \leq \lambda_{n-1} \leq \lambda_{n-1}^{(k)} + 1 \leq \lambda_n,$$

z čoho špeciálne

$$\lambda_2^{(k)} \geq \lambda_2 - 1. \quad (7.24)$$

Všetky ostatné grafy o n vrcholoch môžeme z grafov uvažovaného typu získať vymazaním niekoľkých hrán incidentných s vrcholom k . Podľa tvrdenia 7.9.2 touto procedúrou určite nemôžeme zvýšiť hodnotu λ_2 , pričom odobranie týchto hrán evidentne nemá žiaden vplyv na hodnotu $\lambda_2^{(k)}$. Nerovnosť (7.24) tak bude platiť aj vo všeobecnom prípade. $\square$

Stupeň vrcholovej súvislosti grafu $\mathcal{G}$ definujeme ako minimálny počet prvkov $\kappa(\mathcal{G})$ množiny vrcholov $S \subseteq V(\mathcal{G})$ takej, že indukovaný podgraf grafu $\mathcal{G}$ na vrcholoch $V(\mathcal{G}) \setminus S$ je nesúvislý; ide tak o *najmenší možný* počet vrcholov, ktorých odobraním z grafu $\mathcal{G}$ vznikne nesúvislý graf. Pri úplných grafoch, kde žiadna takáto množina neexistuje, budeme trochu neštandardne klásť $\kappa(K_n) = n$.²⁰ Dokážeme teraz, že algebraický stupeň súvislosti grafu nemôže presiahnuť jeho stupeň vrcholovej súvislosti.

Veta 7.9.4. *Nech $n \in \mathbb{N} \setminus \{0, 1\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s algebraickým stupňom súvislosti λ_2 a stupňom vrcholovej súvislosti $\kappa(\mathcal{G})$. Potom $\lambda_2 \leq \kappa(\mathcal{G})$.*

Dôkaz. Pre úplný graf K_n je pri našej definícii stupňa vrcholovej súvislosti $\lambda_2 = \kappa(\mathcal{G}) = n$ a tvrdenie platí. Predpokladajme teda, že graf $\mathcal{G}$ nie je úplný. Ak potom $\kappa(\mathcal{G}) = m$ pre nejaké $m \in \mathbb{N}$, musí existovať množina vrcholov $S \subseteq V(\mathcal{G})$ taká, že $|S| = m$ a indukovaný podgraf grafu $\mathcal{G}$ s množinou vrcholov $V(\mathcal{G}) \setminus S$ je nesúvislý. V takom prípade zrejme $n \geq m + 2$. Ak teraz $S = \{v_1, \dots, v_m\}$, môže tento indukovaný podgraf vzniknúť z grafu $\mathcal{G}$ tak, že pre $k = 1, \dots, m$ z grafu postupne odoberieme vrchol v_k aj s incidentnými hranami. Pre každé k pritom podľa lemy 7.9.3 znížime algebraický stupeň súvislosti najviac o 1 a algebraický stupeň súvislosti výsledného nesúvislého grafu musí byť rovný nule. Nutne teda $\lambda_2 \leq m = \kappa(\mathcal{G})$, čo bolo treba dokázať. $\square$

7.10 Fiedlerove vektory

Po rozložení množiny vrcholov $V = V(\mathcal{G})$ súvislého grafu $\mathcal{G}$ na dve neprázdne disjunktné podmnožiny V_1, V_2 také, že $V = V_1 \cup V_2$, nemusia byť indukované podgrafy grafu $\mathcal{G}$ na množinách vrcholov V_1 a V_2 súvislé. Nájdenie rozkladov $V = V_1 \cup V_2$ zodpovedajúcich rozkladu súvislého grafu na dva *súvislé* indukované podgrafy je prakticky užitočná úloha, ktorú možno vyriešiť na základe znalosti vlastných vektorov Laplaceovej matice grafu $\mathcal{G}$ prislúchajúcich k jej vlastnému číslu λ_2 určujúcemu algebraický stupeň súvislosti grafu $\mathcal{G}$. Tieto vlastné vektory nazveme *Fiedlerovými vektormi*.

Definícia 7.10.1. *Nech $n \in \mathbb{N} \setminus \{0, 1\}$ a $\mathcal{G} = ([n], \varphi)$ je graf s Laplaceovou maticou $L = \mathbf{L}(\mathcal{G})$ a laplaceovským spektrom $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. *Fiedlerovým vektorom* grafu $\mathcal{G}$ nazveme ľubovoľný reálny vlastný vektor $\mathbf{v}_2 \in \mathbb{R}^n$ matice L prislúchajúci k vlastnému číslu λ_2 .*

²⁰Častejšia je definícia, pri ktorej $\kappa(K_n) = n - 1$. Stupeň vrcholovej súvislosti vtedy možno chápať ako najmenší počet vrcholov, ktorých odobraním z grafu môže vzniknúť nesúvislý graf *alebo* graf o jednom vrchole. Pri našej definícii by sme graf o jednom vrchole v princípe nahradili prázdny grafom.

Použitie Fiedlerových vektorov pri hľadaní rozkladov súvislých grafov na súvislé grafy je založené na nasledujúcej vete, ktorú dokážeme len čiastočne metódou prebratou z [17].

Veta 7.10.2 (Fiedler). *Nech $n \in \mathbb{N} \setminus \{0, 1\}$, $\mathcal{G} = ([n], \varphi)$ je súvislý graf, $\mathbf{v}_2 = (x_1, \dots, x_n)^T$ je Fiedlerov vektor grafu $\mathcal{G}$ a $r \in \mathbb{R}$. Položme*

$$V_1 = \{k \in V(\mathcal{G}) \mid x_k \geq r\} \quad a \quad V_2 = \{k \in V(\mathcal{G}) \mid x_k < r\}.$$

Ak $V_1 \neq \emptyset$ a $V_2 \neq \emptyset$, sú indukované podgrafy grafu $\mathcal{G}$ na množinách vrcholov V_1 a V_2 súvislé.

Dôkaz. Vetu dokážeme iba pre $r = 0$ a podgraf indukovaný množinou $V_1 = \{k \in V(\mathcal{G}) \mid x_k \geq 0\}$ – dokážeme teda, že graf $\mathcal{G}_1 = (V_1, \varphi_1)$ s $\varphi_1(u, v) = \varphi(u, v)$ pre všetky $u, v \in V_1$ musí byť súvislý. Za účelom sporu predpokladajme, že je tento graf nesúvislý a množiny vrcholov jeho komponentov súvislosti sú dané ako $U_1, \dots, U_s$ pre nejaké $s \geq 2$.

Bez ujmy na všeobecnosti predpokladajme, že vrcholy grafu $\mathcal{G}$ sú očíslované tak, že pre $i = 1, \dots, s$ a $j = i+1, \dots, s$ majú všetky vrcholy z U_j väčšie číslo ako všetky vrcholy z U_i a zároveň všetky vrcholy z V_2 majú väčšie číslo ako všetky vrcholy z $V_1 = U_1 \cup \dots \cup U_s$. Laplaceovu maticu $L = \mathbf{L}(\mathcal{G})$ grafu $\mathcal{G}$ potom môžeme písať ako

$$L = \begin{pmatrix} M_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & R_1 \\ \mathbf{0} & M_2 & \mathbf{0} & \dots & \mathbf{0} & R_2 \\ \mathbf{0} & \mathbf{0} & M_3 & \dots & \mathbf{0} & R_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & M_s & R_s \\ R_1^T & R_2^T & R_3^T & \dots & R_s^T & N_2 \end{pmatrix},$$

kde diagonálne bloky $M_1, \dots, M_s$ prislúchajú množinám vrcholov $U_1, \dots, U_s$ a diagonálny blok N_2 prislúcha množine vrcholov V_2 . Vektor $\mathbf{v}_2$ budeme potom tiež písať ako

$$\mathbf{v}_2 = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_s \\ \mathbf{y}_2 \end{pmatrix},$$

kde počet zložiek každého z vektorov $\mathbf{x}_i$ je pre $i = 1, \dots, s$ rovnaký ako počet prvkov množiny U_i a počet zložiek vektora $\mathbf{y}_2$ je rovnaký ako počet prvkov množiny V_2 . Z definície množín vrcholov V_1 a V_2 pritom vyplýva, že vektory $\mathbf{x}_1, \dots, \mathbf{x}_s$ sú nezáporné a vektor $\mathbf{y}_2$ je záporný.

Predpokladajme teraz, že sú vlastné čísla matice L dané ako $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ a uvažujme vlastné čísla $\lambda'_1 \leq \lambda'_2 \leq \dots \leq \lambda'_{n-|V_2|}$ podmaticy

$$L[V_2] = \begin{pmatrix} M_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & M_2 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & M_3 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & M_s \end{pmatrix}.$$

Určite $\lambda_1 = 0$ a vďaka súvislosti grafu $\mathcal{G}$ musí byť $\lambda_2 > 0$. Maticu $L[V_2]$ navyše možno z matice L získať postupným odobraním k -teho riadku a stĺpca pre $k = n, \dots, n - |V_2| + 1$. Niekoľkonásobným použitím Cauchyho vety o prepletaní vlastných čísel teda zisťujeme, že $\lambda'_2 \geq \lambda_2 > 0$.²¹

²¹Vlastné číslo λ'_1 pritom môže, ale nemusí byť nulové, pretože $L[V_2]$ nemusí byť Laplaceovou maticou žiadneho grafu.

Matica $L[V_2]$ je blokovo diagonálna a jej spektrum je tak „zložením“ spektier matíc $M_1, \dots, M_s$. Bez ujmy na všeobecnosti môžeme predpokladať, že $\lambda'_1 \notin \sigma(M_1)$; v opačnom prípade sme totiž mohli očíslovanie vrcholov zvoliť inak. Rovnosť $L\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ môžeme prepísať ako

$$\begin{pmatrix} M_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & R_1 \\ \mathbf{0} & M_2 & \mathbf{0} & \dots & \mathbf{0} & R_2 \\ \mathbf{0} & \mathbf{0} & M_3 & \dots & \mathbf{0} & R_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & M_s & R_s \\ R_1^T & R_2^T & R_3^T & \dots & R_s^T & N_2 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \vdots \\ \mathbf{x}_s \\ \mathbf{y}_2 \end{pmatrix} = \begin{pmatrix} \lambda_2\mathbf{x}_1 \\ \lambda_2\mathbf{x}_2 \\ \lambda_2\mathbf{x}_3 \\ \vdots \\ \lambda_2\mathbf{x}_s \\ \lambda_2\mathbf{y}_2 \end{pmatrix},$$

z čoho

$$\begin{pmatrix} M_1 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & R_1 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_s \\ \mathbf{y}_2 \end{pmatrix} = M_1\mathbf{x}_1 + R_1\mathbf{y}_2 = \lambda_2\mathbf{x}_1.$$

Preto

$$(M_1 - \lambda_2\mathbf{I})\mathbf{x}_1 = -R_1\mathbf{y}_2. \quad (7.25)$$

Podobne ako v dôkaze lemy 7.9.3 teraz zisťujeme, že vlastné čísla matice $M_1 - \lambda_2\mathbf{I}$ sú práve rozdielmi vlastných čísel matice M_1 a čísla λ_2 . Keďže navyše $\lambda'_1 \notin \sigma(M_1)$, sú všetky vlastné čísla matice M_1 rovné aspoň λ'_2 – pričom vieme, že $\lambda'_2 \geq \lambda_2$. Všetky vlastné čísla matice $M_1 - \lambda_2\mathbf{I}$ sú teda nezáporné a matica $M_1 - \lambda_2\mathbf{I}$ je kladne semidefinitná podľa vety 7.6.2.

Keďže je ďalej matica R_1 mimodiagonálnym blokom Laplaceovej matice grafu $\mathcal{G}$, nemôže obsahovať žiaden kladný prvok. Podobne vektor $\mathbf{y}_2$ z definície množiny V_2 obsahuje iba záporné čísla. Z toho vyplýva, že vektor $R_1\mathbf{y}_2$ je nezáporný. Po prenásobení rovnosti (7.25) zľava nezáporným vektorom $\mathbf{x}_1^T$ tak zisťujeme, že

$$\mathbf{x}_1^T(M_1 - \lambda_2\mathbf{I})\mathbf{x}_1 = -\mathbf{x}_1^T R_1\mathbf{y}_2 \leq 0,$$

čo môže vďaka kladnej semidefinitnosti matice $M_1 - \lambda_2\mathbf{I}$ nastať iba ak

$$\mathbf{x}_1^T(M_1 - \lambda_2\mathbf{I})\mathbf{x}_1 = 0. \quad (7.26)$$

Potom však buď $\mathbf{x}_1 = \mathbf{0}$, alebo

$$R_{M_1 - \lambda_2\mathbf{I}}(\mathbf{x}_1) = \frac{\mathbf{x}_1^T(M_1 - \lambda_2\mathbf{I})\mathbf{x}_1}{\mathbf{x}_1^T \mathbf{x}_1} = 0$$

a z Rayleighovho princípu vyplýva, že $\mathbf{x}_1$ je vlastným vektorom kladne semidefinitnej matice $M_1 - \lambda_2\mathbf{I}$ prislúchajúcim k vlastnému číslu 0. V oboch prípadoch zisťujeme, že

$$(M_1 - \lambda_2\mathbf{I})\mathbf{x}_1 = \mathbf{0},$$

čo podľa (7.25) možno prepísať ako

$$-R_1\mathbf{y}_2 = \mathbf{0}.$$

Vektor $\mathbf{y}_2$ je tu záporný a matica R_1 neobsahuje žiaden kladný prvok – nutne teda $R_1 = \mathbf{0}$. Z toho ďalej vyplýva, že vrcholy z množiny U_1 nie sú v grafe $\mathcal{G}$ spojené hranou so žiadnym vrcholom množiny V_2 . Keďže tieto vrcholy nie sú spojené hranou ani so žiadnym vrcholom z množiny $U_2 \cup \dots \cup U_s$, je graf $\mathcal{G}$ nesúvislý, čo je spor s naším pôvodným predpokladom. $\square$

Fiedlerove vektory teda obsahujú informáciu umožňujúcu rozložiť daný súvislý graf na dva súvislé podgrafy a opakovaním tohto procesu na ľubovoľný prípustný počet súvislých podgrafov. Empirické zistenia ukazujú, že pri vhodných voľbách konštanty r zo znenia predchádzajúcej vety takto získané podgrafy často zodpovedajú prirodzeným zhlukom vrcholov, čo je základom metódy analýzy grafových dát známej ako *spektrálne zhľukovanie* [40].

Cvičenia

1. Pozorovali sme, že každá reálna symetrická matica je ortogonálne diagonalizovateľná. Existuje ortogonálne diagonalizovateľná reálna matica, ktorá nie je symetrická?
2. Uvažujme maticu $A \in \mathbb{R}_{\geq 0}^{n \times n}$ pre nejaké $n \in \mathbb{N} \setminus \{0\}$, ktorá je súčasne ireducibilná a symetrická. Charakterizujte všetky možné indexy primitívnosti matice A .
3. Nájdite príklad nesúvislého grafu $\mathcal{G} = (V, \varphi)$ takého, že najväčšie vlastné číslo λ_1 jeho matice susednosti je rovné maximálnemu stupňu vrcholu v grafe $\mathcal{G}$, no graf $\mathcal{G}$ napriek tomu nie je regulárny.
4. Nájdite všetky grafy $\mathcal{G}$, ktorých matica susednosti $\mathbf{A}(\mathcal{G})$ má aspoň jeden vlastný vektor

$$\mathbf{c} = (c, \dots, c)^T$$

pre nejaké $c \in \mathbb{R} \setminus \{0\}$.

5. Charakterizujte všetky súvislé grafy $\mathcal{G}$ s obyčajným spektrom $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ takým, že $\lambda_n = -\lambda_1$.
6. Charakterizujte všetky grafy $\mathcal{G}$, pre ktoré je matica susednosti $\mathbf{A}(\mathcal{G})$ kladne semidefinitná.
7. Symetrickú maticu $A \in \mathbb{R}^{n \times n}$ nazveme *kladne definitnou*, ak pre všetky $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ je $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$. Dokážte, že sú nasledujúce tvrdenia ekvivalentné:
 - (i) Matica A je kladne definitná.
 - (ii) Všetky vlastné čísla matice A sú kladné.
 - (iii) Existuje $m \in \mathbb{N}$ a matica $B \in \mathbb{R}^{m \times n}$ plnej stĺpcovej hodnoty n taká, že $A = B^T B$.
8. Dajte do vzájomného vzťahu obyčajné a laplaceovské spektrum *regulárnych* grafov.
9. Dokážte, že počet komponentov súvislosti grafu je vždy rovný počtu nulových vlastných čísel jeho Laplaceovej matice.
10. Dokážte, že počet kostier úplného bipartitného grafu $K_{m,n}$ s $m, n \in \mathbb{N} \setminus \{0\}$ je $m^{n-1} n^{m-1}$.

Nech $\mathcal{G}$ je neorientovaný graf bez slučiek s množinou vrcholov $[n]$. *Kladnou Laplaceovou maticou* (angl. *signless Laplacian matrix*) grafu $\mathcal{G}$ nazveme maticu

$$\mathbf{Q}(\mathcal{G}) = D + \mathbf{A}(\mathcal{G}),$$

kde $\mathbf{A}(\mathcal{G})$ je matica susednosti grafu $\mathcal{G}$ a $D = (d_{i,j})_{n \times n}$ je diagonálna matica taká, že pre $k = 1, \dots, n$ je $d_{k,k} = \deg_{\mathcal{G}}(k)$.

11. Dokážte, že kladná Laplaceova matica grafu $\mathcal{G}$ s množinou vrcholov $[n]$ je vždy kladne semidefinitná a nájdite čo možno najprirodzenejšie definovanú maticu B typu $m \times n$ takú, že $\mathbf{Q}(\mathcal{G}) = B^T B$.
12. Vlastné čísla kladne semidefinitnej matice $\mathbf{Q}(\mathcal{G})$ sú vždy nezáporné. Nájdite triedu všetkých grafov $\mathcal{G}$, pre ktoré je nula vlastným číslom matice $\mathbf{Q}(\mathcal{G})$.
13. Dokážte, že Laplaceova a kladná Laplaceova matica *bipartitného grafu* majú rovnaké charakteristické polynómy (a teda aj rovnaké spektrá).

Poznámky a odkazy na samoštúdium

Rozsiahlej problematike spektrálnej teórie grafov sa venuje pomerne veľké množstvo špecializovanej literatúry, často s iba veľmi malými obsahovými prienikmi – spomeňme napríklad [26, 25, 16, 20, 100]. O spektrách grafov sa tiež možno dočítať v klasických učebniciach algebraickej teórie grafov [49, 12], do ktorej spektrálna teória grafov spadá. Veľmi užitočným zdrojom je tu tiež text [17] o kombinatorickej teórii matíc. Výklad spektrálnej teórie grafov viac orientovaný na aplikácie možno nájsť v literatúre, v ktorej sa grafy skúmajú pod názvom *komplexné siete* – napríklad v [40, 39].

Päť rôznych dôkazov Cayleyho vzorca, ako aj ďalší z alternatívnych dôkazov vety 7.8.1, možno nájsť v učebnici Matouška a Nešetřila [82].

Spektrá matíc prislúchajúcich ku grafom sa sprvoti skúmali predovšetkým implicitne na poli kvantovej chémie. Za počiatok spektrálnej teórie grafov ako matematickej disciplíny sa však považuje až článok [21] z päťdesiatych rokov 20. storočia, v ktorom bola okrem iného dokázaná aj veta 7.5.1. Pojem algebraického stupňa súvislosti grafu zaviedol M. Fiedler [42], ktorý neskôr dokázal aj vetu 7.10.2 [43].

Literatúra

- [1] Agarwal, R. P.: *Difference Equations and Inequalities*. New York: Marcel Dekker, druhé vydanie, 2000.
- [2] Aho, A. V.; Hopcroft, J. E.; Ullman, J. D.: *The Design and Analysis of Computer Algorithms*. Reading: Addison-Wesley, 1974.
- [3] Albert, J.; Kari, J.: Digital Image Compression. In *Handbook of Weighted Automata*, ed. M. Droste; W. Kuich; H. Vogler, kapitola 11, Springer, 2009, s. 453–479.
- [4] Almagor, S.; Boker, U.; Kupferman, O.: What’s Decidable About Weighted Automata? In *ATVA 2011*, Springer, 2011, s. 482–491.
- [5] Almagor, S.; Boker, U.; Kupferman, O.: What’s Decidable About Weighted Automata? *Information and Computation*, ročník 282, 2022, článok 104651.
- [6] Arden, D. N.: Delayed-Logic and Finite-State Machines. In *SWCT (FOCS) 1961*, IEEE Computer Society, 1961, s. 133–151.
- [7] Axler, S.: *Linear Algebra Done Right*. Cham: Springer, tretie vydanie, 2015.
- [8] Béal, M.-P.; Lombardy, S.; Sakarovitch, J.: Conjugacy and Equivalence of Weighted Automata and Functional Transducers. In *CSR 2006*, Springer, 2006, s. 58–69.
- [9] Berman, A.; Plemmons, R. J.: *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia: SIAM, 1994.
- [10] Berstel, J.: *Transductions and Context-Free Languages*. Stuttgart: B. G. Teubner, 1979.
- [11] Berstel, J.; Reutenauer, C.: *Noncommutative Rational Series with Applications*. Cambridge: Cambridge University Press, 2011.
- [12] Biggs, N.: *Algebraic Graph Theory*. Cambridge: Cambridge University Press, 1974.
- [13] Bloom, S. L.; Ésik, Z.; Kuich, W.: Partial Conway and Iteration Semirings. *Fundamenta Informaticae*, ročník 86, č. 1–2, 2008: s. 19–40.
- [14] Bohner, M.; Peterson, A.: *Dynamic Equations on Time Scales*. Boston: Birkhäuser, 2001.
- [15] Boole, G.: *A Treatise on the Calculus of Finite Differences*. Londýn: MacMillan and Co., 1860.
- [16] Brouwer, A. E.; Haemers, W. H.: *Spectra of Graphs*. New York: Springer, 2012.
- [17] Brualdi, R. A.; Ryser, H. J.: *Combinatorial Matrix Theory*. Cambridge: Cambridge University Press, 1991.
- [18] Cameron, P. J.: *Notes on Counting: An Introduction to Enumerative Combinatorics*. Cambridge: Cambridge University Press, 2017.
- [19] Cardon, A.; Crochemore, M.: Détermination de la Représentation Standard d’une Série Reconnaisable. *Informatique Théorique et Applications*, ročník 14, č. 4, 1980: s. 371–379.
- [20] Chung, F. R. K.: *Spectral Graph Theory*. Providence: American Mathematical Society, 1997.
- [21] Collatz, L.; Sinogowitz, U.: Spektren endlicher grafen. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, ročník 21, č. 1, 1957: s. 63–77.
- [22] Conway, J. H.: *Regular Algebra and Finite Machines*. Londýn: Chapman and Hall, 1971.
- [23] Cormen, T. H.; Leiserson, C. E.; Rivest, R. L.; et al.: *Introduction to Algorithms*. Cambridge: MIT Press, tretie vydanie, 2009.
- [24] Culik II, K.; Kari, J.: Image compression using weighted finite automata. *Computers & Graphics*, ročník 17, č. 3, 1993: s. 305–313.
- [25] Cvetković, D.; Doob, M.; Sachs, H.: *Spectra of Graphs*. New York: Academic Press, 1980.
- [26] Cvetković, D.; Rowlinson, P.; Simić, S.: *An Introduction to the Theory of Graph Spectra*. Cambridge: Cambridge University Press, 2010.
- [27] Demel, J.: *Grafy a jejich aplikace*. Praha: Academia, 2002.

- [28] Diestel, R.: *Graph Theory*. Berlín: Springer, piate vydanie, 2017.
- [29] Droste, M.; Kuich, W.: Semirings and Formal Power Series. In *Handbook of Weighted Automata*, ed. M. Droste; W. Kuich; H. Vogler, kapitola 1, Springer, 2009, s. 3–28.
- [30] Droste, M.; Kuich, W.; Vogler, H. (editori): *Handbook of Weighted Automata*. Heidelberg: Springer, 2009.
- [31] Droste, M.; Kuske, D.: Weighted Automata. In *Handbook of Automata Theory, Vol. 1*, ed. J.-É. Pin, kapitola 4, European Mathematical Society, 2021, s. 113–150.
- [32] Droste, M.; Stüber, T.; Vogler, H.: Weighted Finite Automata over Strong Bimonoids. *Information Sciences*, ročník 180, č. 1, 2010: s. 156–166.
- [33] Eilenberg, S.: *Automata, Languages, and Machines, Vol. A*. New York: Academic Press, 1974.
- [34] Elaydi, S.: *An Introduction to Difference Equations*. New York: Springer, tretie vydanie, 2005.
- [35] Ěsik, Z.; Kuich, W.: Locally Closed Semirings. *Monatshefte für Mathematik*, ročník 137, 2002: s. 21–29.
- [36] Ěsik, Z.; Kuich, W.: Finite Automata. In *Handbook of Weighted Automata*, ed. M. Droste; W. Kuich; H. Vogler, kapitola 3, Springer, 2009, s. 69–104.
- [37] Ěsik, Z.; Kuich, W.: A Unifying Kleene Theorem for Weighted Finite Automata. In *Maurer Festschrift*, Springer, 2011, s. 76–89.
- [38] Ěsik, Z.; Kuich, W.: Modern Automata Theory. Elektronický učebný text, 2012.
URL <http://dmg.tuwien.ac.at/kuich/mat.pdf>
- [39] Estrada, E.: *The Structure of Complex Networks*. Oxford: Oxford University Press, 2011.
- [40] Estrada, E.; Knight, P. A.: *A First Course in Network Theory*. Oxford: Oxford University Press, 2015.
- [41] Everest, G.; van der Poorten, A.; Shparlinski, I.; et al.: *Recurrence Sequences*. Providence: American Mathematical Society, 2003.
- [42] Fiedler, M.: Algebraic Connectivity of Graphs. *Czechoslovak Mathematical Journal*, ročník 23, č. 2, 1973: s. 298–305.
- [43] Fiedler, M.: A Property of Eigenvectors of Nonnegative Symmetric Matrices and Its Application to Graph Theory. *Czechoslovak Mathematical Journal*, ročník 25, č. 4, 1975: s. 619–633.
- [44] Flajolet, P.; Sedgewick, R.: *Analytic Combinatorics*. Cambridge: Cambridge University Press, 2009.
- [45] Floyd, R. W.: Algorithm 97: Shortest Path. *Communications of the ACM*, ročník 5, č. 6, 1962: str. 345.
- [46] Frobenius, G.: Über Matrizen aus nicht negativen Elementen. *Sitzungsberichte Königlich Preussischen Akademie der Wissenschaft*, 1912: s. 456–477.
- [47] Gallian, J. A.: *Contemporary Abstract Algebra*. Lexington: D. C. Heath and Company, druhé vydanie, 1990.
- [48] Gantmacher, F. R.: *The Theory of Matrices, Vol. 2*. Providence: AMS Chelsea Publishing, 2000.
- [49] Godsil, C.; Royle, G.: *Algebraic Graph Theory*. New York: Springer, 2001.
- [50] Golan, J. S.: *Semirings and Their Applications*. Dordrecht: Springer, 1999.
- [51] Graham, R. L.; Knuth, D. E.; Patashnik, O.: *Concrete Mathematics*. Reading: Addison-Wesley, druhé vydanie, 1994.
- [52] Hebisch, U.; Weinert, H. J.: *Semirings*. Singapur: World Scientific, 1998.
- [53] Henrici, P.: *Applied and Computational Complex Analysis, Vol. 1*. New York: Wiley, 1974.
- [54] Hirschfeld, J. W. P.; Korchmáros, G.; Torres, F.: *Algebraic Curves over a Finite Field*. Princeton: Princeton University Press, 2008.
- [55] Hladík, M.: *Lineární algebra (nejen) pro informatiky*. Praha: MatfyzPress, 2019.
- [56] Hopcroft, J. E.; Ullman, J. D.: *Introduction to Automata Theory, Languages, and Computation*. Reading: Addison-Wesley, 1979.
- [57] Horn, R. A.; Johnson, C. R.: *Matrix Analysis*. New York: Cambridge University Press, druhé vydanie, 2013.
- [58] Jerri, A. J.: *Linear Difference Equations with Discrete Transform Methods*. Dordrecht: Springer, 1996.
- [59] Jordan, C.: *Calculus of Finite Differences*. New York: Chelsea Publishing Company, druhé vydanie, 1950.
- [60] Kalman, D.: The Generalized Vandermonde Matrix. *Mathematics Magazine*, ročník 57, č. 1, 1984: s. 15–21.
- [61] Kari, J.: Image Processing Using Finite Automata. In *Recent Advances in Formal Languages and Applications*, ed. Z. Ěsik; C. Martín-Vide; V. Mitrana, kapitola 7, Springer, 2006, s. 171–208.
- [62] Kelley, W. G.; Peterson, A. C.: *Difference Equations*. San Diego: Academic Press, druhé vydanie, 2001.
- [63] Kemeny, J. G.; Snell, J. L.: *Finite Markov Chains*. New York: Springer, 1960.

- [64] Kleene, S. C.: Representation of Events in Nerve Nets and Finite Automata. Technická správa RM-704, U. S. Air Force Project RAND, 1951.
- [65] Knight, K.; May, J.: Applications of Weighted Automata in Natural Language Processing. In *Handbook of Weighted Automata*, ed. M. Droste; W. Kuich; H. Vogler, kapitola 14, Springer, 2009, s. 571–596.
- [66] Knuth, D. E.: *The Art of Computer Programming, Vol. 1*. Reading: Addison-Wesley, tretie vydanie, 1997.
- [67] Kostolányi, P.: A Unifying Approach to Algebraic Systems Over Semirings. *Theory of Computing Systems*, ročník 63, 2019: s. 615–633.
- [68] Kostolányi, P.: Rational Elements of Summation Semirings. *Theoretical Computer Science*, ročník 867, 2021: s. 101–127.
- [69] Kostolányi, P.: On Some Decidability Questions Concerning Supports of Rational Series. *Information Processing Letters*, ročník 179, 2023, článok 106290.
- [70] Kostolányi, P.: Analytická a enumeratívna kombinatorika. 2024.
URL <http://www.dcs.fmph.uniba.sk/~kostolanyi/aek/skripta.pdf>
- [71] Kozen, D. C.: *The Design and Analysis of Algorithms*. New York: Springer, 1992.
- [72] Kreyszig, E.: *Introductory Functional Analysis with Applications*. New York: Wiley, 1978.
- [73] Krob, D.: The Equality Problem for Rational Series with Multiplicities in the Tropical Semiring is Undecidable. *International Journal of Algebra and Computation*, ročník 4, č. 3, 1994: s. 405–425.
- [74] Kuich, W.: Semirings and Formal Power Series. In *Handbook of Formal Languages, Vol. 1*, ed. G. Rozenberg; A. Salomaa, kapitola 9, Springer, 1997, s. 609–677.
- [75] Kuich, W.; Salomaa, A.: *Semirings, Automata, Languages*. Heidelberg: Springer, 1986.
- [76] Lancaster, P.; Tismenetsky, M.: *The Theory of Matrices*. San Diego: Academic Press, druhé vydanie, 1985.
- [77] Lax, P. D.: *Linear Algebra and its Applications*. Hoboken: Wiley, druhé vydanie, 2007.
- [78] Lind, D.; Marcus, B.: *Symbolic Dynamics and Coding*. Cambridge: Cambridge University Press, 1995.
- [79] Lombardy, S.; Mairesse, J.: Max-Plus Automata. In *Handbook of Automata Theory, Vol. 1*, ed. J.-É. Pin, kapitola 5, European Mathematical Society, 2021, s. 151–188.
- [80] Maletti, A.: Survey: Finite-State Technology in Natural Language Processing. *Theoretical Computer Science*, ročník 679, 2017: s. 2–17.
- [81] Mariconda, C.; Tonolo, A.: *Discrete Calculus*. Cham: Springer, 2016.
- [82] Matoušek, J.; Nešetřil, J.: *Kapitoly z diskrétní matematiky*. Karolinum: Praha, tretie vydanie, 2007.
- [83] McNaughton, R.; Yamada, H.: Regular Expressions and State Graphs for Automata. *IRE Transactions on Electronic Computers*, ročník 9, č. 1, 1960: s. 39–47.
- [84] Meyer, C. D.: *Matrix Analysis and Applied Linear Algebra*. Philadelphia: SIAM, 2000.
- [85] Mickens, R. E.: *Difference Equations*. Boca Raton: CRC Press, tretie vydanie, 2015.
- [86] Minc, H.: *Nonnegative Matrices*. New York: Wiley, 1988.
- [87] Mohri, M.: Weighted Automata Algorithms. In *Handbook of Weighted Automata*, ed. M. Droste; W. Kuich; H. Vogler, kapitola 6, Springer, 2009, s. 213–254.
- [88] Nešetřil, J.: *Teorie grafů*. Praha: SNTL, 1979.
- [89] Niven, I.: Formal Power Series. *The American Mathematical Monthly*, ročník 76, č. 8, 1969: s. 871–889.
- [90] Perron, O.: Zur Theorie der Matrices. *Mathematische Annalen*, ročník 64, č. 2, 1907: s. 248–263.
- [91] Petre, I.; Salomaa, A.: Algebraic Systems and Pushdown Automata. In *Handbook of Weighted Automata*, ed. M. Droste; W. Kuich; H. Vogler, kapitola 7, Springer, 2009, s. 257–289.
- [92] Pin, J.-É.: Tropical Semirings. In *Idempotency*, ed. J. Gunawardena, kapitola 2, Cambridge University Press, 1998, s. 50–69.
- [93] Roy, B.: Transitivité et connexité. *Comptes rendus de l'Académie des Sciences Paris*, ročník 249, 1959: s. 216–218.
- [94] Sakarovitch, J.: *Elements of Automata Theory*. Cambridge: Cambridge University Press, 2009.
- [95] Sakarovitch, J.: Rational and Recognisable Power Series. In *Handbook of Weighted Automata*, ed. M. Droste; W. Kuich; H. Vogler, kapitola 4, Springer, 2009, s. 105–174.
- [96] Šalát, T.; et al.: *Malá encyklopédia matematiky*. Bratislava: Obzor, tretie vydanie, 1981.
- [97] Salomaa, A.; Soittola, M.: *Automata-Theoretic Aspects of Formal Power Series*. New York: Springer, 1978.
- [98] Schützenberger, M.-P.: On the definition of a family of automata. *Information and Control*, ročník 4, č. 2–3, 1961: s. 245–270.

- [99] Simmons, G. F.: *Introduction to Topology and Modern Analysis*. New York: McGraw-Hill, 1963.
- [100] Van Mieghem, P.: *Graph Spectra for Complex Networks*. Cambridge: Cambridge University Press, 2011.
- [101] Varga, R. S.: *Matrix Iterative Analysis*. Heidelberg: Springer, druhé vydanie, 2000.
- [102] Warshall, S.: A Theorem on Boolean Matrices. *Journal of the ACM*, ročník 9, č. 1, 1962: s. 11–12.
- [103] Wielandt, H.: Unzerlegbare, nicht negative Matrizen. *Mathematische Zeitschrift*, ročník 52, č. 1, 1950: s. 642–648.
- [104] Wilf, H. S.: *generatingfunctionology*. New York: Academic Press, druhé vydanie, 1994.
- [105] Zlatoš, P.: *Lineárna algebra a geometria*. Bratislava: Albert Marenčin PT, 2011.