ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta jaderná a fyzikálně inženýrská

Kritický exponent balancovaných slov

Critical exponent of balanced sequences

Bakalářská práce

Autor: **Daniela Opočenská**

Vedoucí práce: **doc. Ing. L'ubomíra Dvořáková, Ph.D.**

Akademický rok: 2021/2022

Katedra: matematiky   Akademický rok: 2021/2022

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student:	Daniela Opočenská
Studijní program:	Aplikace přírodních věd
Studijní obor:	Matematické inženýrství
Studijní zaměření:	Matematické modelování
Název práce (česky):	Kritický exponent balancovaných slov
Název práce (anglicky):	Critical exponent of balanced sequences

Pokyny pro vypracování:

1) Studium kombinatoriky na slovech se zaměřením na balancovaná slova.
2) Tvorba programu pro výpočet asymptotického kritického exponentu balancovaných slov.
3) Tvorba programu pro výpočet kritického exponentu balancovaných slov.
4) Hledání minimálního kritického exponentu balancovaných slov nad danou abecedou - potvrzení či vyvrácení stávající domněnky.
5) Hledání minimálního asymptotického kritického exponentu balancovaných slov nad danou abecedou.

Doporučená literatura:

1) A. R. Baranwal, J. Shallit, Critical Exponent of Infinite Balanced Words via the Pell Number System. In 'Proceedings WORDS 2019, LNCS', Springer, 2019, 80-92.

2) F. Dolce, Ľ. Dvořáková, E. Pelantová, On balanced sequences and their asymptotic critical exponent. In 'Proceedings LATA 2021, LNCS', Springer, 2021, 293-304.

3) F. Dolce, Ľ. Dvořáková, E. Pelantová, Computation of critical exponent in balanced sequences. In 'Proceedings WORDS 2021, LNCS', Springer, 2021, 78-90.

4) M. Lothaire, Algebraic combinatorics on Words. Cambridge University Press, Cambridge, 2002.

5) N. Rampersad, J. Shallit, É. Vandomme, Critical exponents of infinite balanced words. Theoretical Computer Science 777, 2019, 454-463.

Jméno a pracoviště vedoucího bakalářské práce:

doc. Ing. Dvořáková Ľubomíra, Ph.D.
Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 120 00 Praha 2

Jméno a pracoviště konzultanta:

Datum zadání bakalářské práce: 31.10.2021

Datum odevzdání bakalářské práce: 7.7.2022

Doba platnosti zadání je dva roky od data zadání.

V Praze dne 21. října 2021

....
garant oboru

....
vedoucí katedry

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta jaderná a fyzikálně inženýrská
-1-

....
děkan

Poděkování:
Chtěla bych poděkovat doc. Ing. Ľubomíře Dvořákové, Ph.D., za ochotu, vedení, skvělé rady a připomínky a trpělivost při tvorbě této práce.

Čestné prohlášení:
Prohlašuji, že jsem tuto práci vypracovala samostatně a uvedla jsem všechnu použitou literaturu.

V Praze dne 7. července 2022 Daniela Opočenská

Název práce:

Kritický exponent balancovaných slov

Autor: Daniela Opočenská

Obor: Matematické inženýrství

Zaměření: Matematické modelování

Druh práce: Bakalářská práce

Vedoucí práce: doc. Ing. L'ubomíra Dvořáková, Ph.D., Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 120 00 Praha 2

Abstrakt: Práce spadá pod tematiku kombinatoriky na slovech. *Balancovaná slova* jsou skupinou slov, která dokážeme generovat pomocí dobře prozkoumaných slov sturmovských a zachovávají si mnoho podobných charakteristik. *Kritický exponent* udává horní hranici toho, kolikrát se faktory v daném slově opakují bezprostředně po sobě. Přirozeně se nabízí otázka, zda se dá nad danou abecedou určitému množství opakování vyhnout? Obecná dolní mez je dobře známa, pro balancovaná slova však byla vyslovena pouze domněnka na základě systematického procházení prefixů slov nad d-písmennými abecedami pro $d \leq 8$. Tato práce obsahuje ucelený popis vlastností sturmovských a balancovaných slov, algoritmus pro výpočet (asymptotického) kritického exponentu balancovaných slov, který je nezávislý na velikosti abecedy, a jeho následnou implementaci. Pomocí programu se nám dále povedlo původní domněnku vyvrátit, najít menší dolní mez pro $d \geq 11$ a pro sudé abecedy dokázat, že je nabývána.

Klíčová slova: asymptotický kritický exponent, balancované slovo, bispeciál, kombinatorika na slovech, kritický exponent, návratové slovo, slovo s konstantními mezerami, sturmovské slovo.

Title:

Critical exponent of balanced sequences

Author: Daniela Opočenská

Abstract: This topic is actual in combinatorics on words. *Balanced sequences* can be generated using well-known Sturmian sequences, and they retain many similar characteristics. The *Critical exponent* is the upper limit of successive repetitions of factors in a given sequence. The obvious question is whether it is possible to avoid a certain number of repetitions in a sequence over a given alphabet? The general lower bound is well known, but only a conjecture was known for balanced sequences. The conjecture was based on a systematic search in prefixes of sequences over d-letter alphabets for $d \leq 8$. This thesis contains a comprehensive description of the properties of Sturmian and balanced sequences, an algorithm for computing the (asymptotic) critical exponent of balanced sequences, which is independent of the size of the alphabet, and its subsequent implementation. We managed to refute the original conjecture. We found a smaller lower bound for $d \geq 11$ and showed that it is attained for even-sized alphabets.

Key words: asymptotic critical exponent, balanced sequence, bispecial, combinatorics on words, constant gap sequence, critical exponent, return word, Sturmian word.

Obsah

Slovníček značení

$\mathbb{N}$ .... přirozená čísla
$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ .... nezáporná celá čísla
$\theta = [0, a_1, a_2, \dots]$ .... řetězový zlomek θ
$\theta = [0, d_1, \dots, d_h, (z_0, z_1, \dots, z_{M-1})^\omega]$ .... posléze periodický řetězový zlomek θ
h .... délka předperiody posléze periodického řetězového zlomku θ
M .... délka periody posléze periodického řetězového zlomku θ
$\mathbf{u}$ .... standardní sturmovské slovo příslušné θ
$\mathbf{y}, \mathbf{y}'$ .... slova s konstantními mezerami nad disjunktními abecedami
$\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$ .... balancované slovo
u_n .... n-tý bispeciál ve slově $\mathbf{u}$
r .... prefixové návratové slovo k u v $\mathbf{u}$
s .... neprefixové návratové slovo k u v $\mathbf{u}$
w .... bispeciál ve slově $\mathbf{v}$
(N, m) .... dvojice příslušná u_n, resp. w ve $\mathbf{v}$ pro $\pi(w) = u_n$, $n = a_1 + a_2 + \cdots + a_N + m$
H .... počet tříd ekvivalence $\sim$
$X_{n+1} = a_{n+1}X_n + X_{n-1}$ pro $n \geq 0$ .... rekurentní rovnice pro posloupnosti $(p_n), (q_n)$ a (Q_n)
p_n .... člen posloupnosti splňující rekurenci výše s poč. podm. $p_{-1} = 1$ a $p_0 = 0$
q_n .... člen posloupnosti splňující rekurenci výše s poč. podm. $q_{-1} = 0$ a $q_0 = 1$
Q_n .... člen posloupnosti splňující rekurenci výše s poč. podm. $Q_{-1} = 1$ a $Q_0 = 1$, tj. $Q_n = p_n + q_n$

Úvod

Kombinatorika na slovech je relativně novým odvětvím matematiky, které se věnuje studiu slov a formálních jazyků. První práce zabývající se kombinatorikou na slovech se datují kolem roku 1900, kdy matematik Axel Thue [4] sledoval opakování faktorů bezprostředně po sobě v nekonečných slovech a kladně zodpověděl následující dvě otázky. Existuje nekonečné slovo nad binární abecedou, ve kterém se žádný faktor neopakuje třikrát bezprostředně po sobě? Existuje nekonečné slovo nad ternární abecedou, ve kterém se žádný faktor neopakuje dvakrát bezprostředně za sebou?

Supremum z hodnot všech repetic faktorů bezprostředně po sobě nazýváme kritický exponent. Můžeme tedy říct, že nekonečná slova nalezená jako odpovědi na předchozí otázky mají kritický exponent menší, nebo roven 3, resp. 2. Tato práce se zaměřuje na nekonečná balancovaná slova, na výpočet jejich kritického exponentu a na minimální možnou hodnotu kritického exponentu pro balancovaná slova nad abecedou o pevně daném počtu písmen. Aktuálnost tématu ilustrujme faktem, že na seminářích *One World Combinatorics on Words Seminar* a *One World Numeration Seminar*, které fungovaly jako náhrada konferencí v době covidu, jsou témata exponentu, repetic, mocnin ve slovech a balancovanosti velmi častá a oblíbená.

Obecnou mez pro kritický exponent nám udává domněnka F. Dejeanové z roku 1972, která byla krok za krokem dokazována mnoha matematiky až do roku 2011. Tato domněnka říká, že minimální kritický exponent nad d-písmennou abecedou je $\frac{d}{d-1}$ pro d alespoň 5. (Asymptotický) kritický exponent se dále zkoumá pro konkrétní třídy nekonečných slov, my se zaměříme na slova balancovaná. Tato práce navazuje na výzkum balancovaných slov od N. Rampersada, J. Shallita a É. Vandomme [17] a A. R. Baranwala a J. Shallita [2].

Nad binární abecedou splývají aperiodická balancovaná slova s těmi sturmovskými a minimální kritický exponent sturmovských slov je dávno známý. Cílem dalšího zkoumání bylo vyslovit a dokázat pro balancovaná slova podobnou domněnku jako pro obecná slova.

Autoři [17] našli balancovaná slova s nejmenším kritickým exponentem nad abecedami o 3 a 4 písmenech. Dále vyslovili domněnku, že minimální kritický exponent balancovaných slov nad abecedou o $d \geq 5$ písmenech je $\frac{d-2}{d-3}$. Dokázali, že kritický exponent nemůže být pro $d \leq 10$ menší, a poskytli seznam balancovaných slov $\mathbf{x}_d$ pro $5 \leq d \leq 10$, u kterých se domnívali, že se hodnota $\frac{d-2}{d-3}$ nabývá. Jejich domněnku pro $5 \leq d \leq 8$ potvrdili A. R. Baranwal a J. Shallit [1, 2].

Další zkoumání ale narazilo na limity výpočetní techniky a bylo potřeba změnit přístup. My využíváme programy založené na výzkumu bispeciálních faktorů a jejich nejkratších návratových slov z článků od F. Dolceho, Ľ. Dvořákové a E. Pelantové [5], [6] a [7]. S těmito programy se nám napřed podařilo domněnku pro $d = 9, 10$ dokázat a následně původní domněnku pro větší abecedy vyvrátit. Ve spolupráci s A. M. Shurem se nám podařilo vyslovit novou domněnku a dokázat ji pro sudé abecedy o alespoň 12 písmenech. Tyto výsledky jsou publikovány v článku [9].

Vlastním příspěvkem bakalářské práce je implementace algoritmů [15] pro hledání asymptotického kritického exponentu a kritického exponentu a jejich následné využití při zkoumání minimálního kritického exponentu balancovaných slov nad d-písmennou abecedou. Dokázali jsme, že pro balancovaná slova o 11, resp. 12 písmenech platí, že minimální kritický exponent nabývá hodnoty $\frac{10}{9}$, resp. $\frac{11}{10}$. Tedy pro $d = 11, 12$ je minimální kritický exponent roven $\frac{d-1}{d-2}$, nikoliv $\frac{d-2}{d-3}$, což vyvrací původní domněnku. A. M. Shur následně dokázal, že minimální kritický exponent balancovaných slov nad d-písmennou abecedou pro $d \geq 11$ je větší nebo roven $\frac{d-1}{d-2}$. My jsme pak opět s využitím programu nalezli balancovaná slova s libovolným sudým počtem písmen d, kde $d \geq 12$, pro která kritický exponent nabývá minimální možné hodnoty $\frac{d-1}{d-2}$. Program intenzivně využíváme také pro zkoumání minimálního asymptotického kritického exponentu, který je známý pouze pro abecedy o nejvýše 5 písmenech.

Obsahem práce je v první části teoretické odvození algoritmů pro výpočet (asymptotického) kritického exponentu a dále pak popis implementace vlastního programu a výsledků, kterých jsme s jeho pomocí dosáhli. V prvé řadě se zaměříme na asymptotický kritický exponent, který odráží opakování faktorů při délce jdoucí do nekonečna, výpočet pak dále rozšíříme o zkoumání krátkých faktorů a umožníme tak výpočet kritického exponentu. Při výpočtu pro nás bude důležitý tvar bispeciálů a jejich návratových slov v balancovaných nekonečných slovech, ukážeme tedy, jak přesně bispeciály vypadají a jak se dají návratová slova hledat.

Celá práce je rozdělena do osmi kapitol. V kapitole 1 definujeme a vysvětlíme základní pojmy kombinatoriky na slovech, v kapitole 2 definujeme pojem exponent a ukážeme vzorec pro výpočet (asymptotického) kritického exponentu. Dále se budeme věnovat sturmovským slovům a jejich bispeciálům v kapitole 3 a následně z nich odvozeným balancovaným slovům v kapitole 4. Algoritmus výpočtu asymptotického kritického exponentu ukážeme s potřebnými vzorci v kapitole 5 a rozšíření výpočtu na krátké faktory je nakonec provedeno v kapitole 6. V kapitole 7 je popsán program s důrazem na změny oproti výpočtu, který jsme ukázali v předchozích kapitolách. S programem „v ruce“ se v kapitole 8 věnujeme otázkám minimálního (asymptotického) kritického exponentu pro balancovaná slova nad d-písmennou abecedou, které jsme s pomocí programu dokázali zodpovědět.

Na závěr shrneme všechny výsledky a zmíníme otevřené problémy, jimž se chceme v budoucnu s použitím programu a jeho dalších rozšíření věnovat.

Všechny kapitoly jsou doplněny příklady. Také pro lepší pochopení textu je v práci uveden slovníček nejdůležitějšího značení, které je v textu používáno a ne vždy znovu opakováno.

Kapitola 1

Základy kombinatoriky na slovech

Abeceda $\mathcal{A}$ je konečná množina znaků (symbolů), které nazýváme *písmena*. Z písmen dále můžeme skládat slova, a to jak konečná, tak nekonečná. Zřetězením konečného počtu písmen z abecedy získáme *konečné slovo* $u = u_0u_1 \dots u_{n-1}$, kde $u_i \in \mathcal{A}$ pro každé $i \in \{0, 1, \dots, n-1\}$. *Délku slova* pak značíme $|u| = n$. Dále pro každé písmeno a z abecedy značíme $|u|_a$ počet výskytů písmena a v u.

K operaci zřetězení existuje neutrální prvek – *prázdné slovo* značené ε. Pokud $u = xyz$ pro slova x, y, z nad abecedou $\mathcal{A}$ (mohou být i prázdná), pak x je *prefix*, z *sufix* a y je *faktor* slova u.

Slovo u nazveme *palindromem*, pokud při čtení pozpátku vznikne stejné slovo, tzn. $u_0u_1 \dots u_{n-1} = u_{n-1}u_{n-2} \dots u_0$.

Nekonečné slovo nad abecedou $\mathcal{A}$ je nekonečná posloupnost písmen z abecedy, tzn. $\mathbf{u} = u_0u_1u_2 \dots$, kde $u_i \in \mathcal{A}$ pro každé $i \in \mathbb{N}_0$. Nekonečné slovo budeme značit tučně. *Faktor* nekonečného slova je jakákoliv posloupnost po sobě následujících písmen ze slova, tzn. u je faktor $\mathbf{u}$, pokud $u = u_iu_{i+1} \dots u_{j-1}$, $i, j \in \mathbb{N}_0$, $i \le j$ (pro $i = j$ je u prázdné slovo). Číslo i nazýváme *výskyt* faktoru u. Můžeme vidět, že faktor nekonečného slova je vždy konečné slovo.

Jazykem slova $\mathbf{u}$ nazýváme množinu všech faktorů slova $\mathbf{u}$ a značíme ho $\mathcal{L}(\mathbf{u})$.

Ke slovu u nad abecedou obsahující d písmen přiřazujeme d-složkový *Parikhův vektor* $\vec{V}(u) \in \mathbb{N}_0^d$, jehož složky tvoří počet výskytů daného písmena ve slově, tj. $\left(\vec{V}(u)\right)_a = |u|_a$ pro každé $a \in \mathcal{A}$.

Řekneme, že nekonečné slovo je *rekurentní*, pokud se každý faktor ve slově vyskytuje nekonečněkrát. Pokud je pro každý faktor vzdálenost mezi jeho každými dvěma po sobě jdoucími výskyty omezená, pak $\mathbf{u}$ nazveme *stejnoměrně rekurentní*.

Nekonečné slovo nazveme *posléze periodické*, pokud je posloupnost písmen posléze periodická, tzn. $\mathbf{u} = uvvvvvvv \dots = u(v)^\omega$, kde ω značí nekonečné opakování faktoru v. Faktor u nazveme *předperioda*, v *perioda*. Nekonečné slovo se nazývá *periodické*, pokud $u = \varepsilon$. Pokud slovo není posléze periodické (tudíž ani periodické), nazýváme ho *aperiodické*.

Definice 1.1. *Slovo* $\mathbf{u}$ *nad abecedou* $\mathcal{A}$ *nazveme* balancované*, pokud pro každé dva faktory* $u, v \in \mathcal{L}(\mathbf{u})$ *stejné délky a pro každé písmeno* c *z abecedy platí, že se počet výskytů tohoto písmena v* u *a ve* v *liší maximálně o jedna, tzn.* $||u|_c - |v|_c| \le 1$.

Balancovaná slova mohou být jak periodická, tak aperiodická.

Návratové slovo k faktoru w slova $\mathbf{u}$ je slovo $v \in \mathcal{L}(\mathbf{u})$ takové, že $vw \in \mathcal{L}(\mathbf{u})$, vw má w jako prefix i jako sufix a mezi nimi se již w nevyskytuje. Jinak řečeno, jsou-li $i, j, i < j$, dva po sobě jdoucí výskyty faktoru w, pak návratové slovo k w je faktor $v = u_iu_{i+1} \dots u_{j-1}$. Množinu všech návratových slov k w značíme $\mathcal{R}_{\mathbf{u}}(w)$.

Faktor u nazveme *pravý speciál* slova $\mathbf{u}$, pokud existují alespoň dvě různá písmena $a, b \in \mathcal{A}$ taková, že $ua, ub \in \mathcal{L}(\mathbf{u})$ – faktor má dvě různá rozšíření doprava. Symetricky definujeme *levý speciál*, který má dvě různá rozšíření doleva, a *bispeciál*, který je zároveň levým i pravým speciálem.

Hustotou písmena $a \in \mathcal{A}$ v nekonečném slově $\mathbf{u}$ nad $\mathcal{A}$ nazveme limitu (pokud existuje)

$$\rho_a = \lim_{n \to +\infty} \frac{|u_0 u_1 \dots u_{n-1}|_a}{n}.$$

Například u periodických slov hustoty písmen vždy existují a jsou racionální.

Příklad 1.1. Zvolíme třípísmennou abecedu $\mathcal{A} = \{a, n, s\}$. Konečným slovem nad touto abecedou je například slovo *ananas*.

Jeho Parikhův vektor (při zachování abecedního pořadí písmen) je:

$$\vec{V}(ananas) = \begin{pmatrix} |ananas|_a \\ |ananas|_n \\ |ananas|_s \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Prefixem tohoto slova je například *ana*, sufixem *anas*. Podíváme-li se na nekonečné periodické slovo $\mathbf{u} = (ananas)^\omega$, pak můžeme vidět, že návratová slova k faktoru *na* jsou $\mathcal{R}_\mathbf{u}(na) = \{na, nasa\}$, k *sa* je návratové slovo pouze jedno, a to *sanana*. Návratové slovo může být i kratší než samotný faktor. Nastává to, když se po sobě jdoucí výskyty překrývají – například v našem příkladu pro faktor *ana*. $\mathcal{R}_\mathbf{u}(ana) = \{an, anas\}$.

Můžeme se také podívat na pravé a levé speciály. Bispeciálem je například a, protože $as, an, na, sa \in \mathcal{L}(\mathbf{u})$. Delším bispeciálem je poté *ana*. U posléze periodických slov nalezneme pouze konečný počet bispeciálů, nikdy totiž nenajdeme bispeciál, který je delší než součet délky předperiody a délky periody.

Dále lze vidět, že $\rho_a = \frac{1}{2}$, $\rho_n = \frac{1}{3}$ a $\rho_s = \frac{1}{6}$ jsou příslušné hustoty písmen.

Morfismy

Morfismus χ je funkce zobrazující slovo nad abecedou $\mathcal{A}$ na jiné slovo nad abecedou $\mathcal{B}$ splňující $\chi(uv) = \chi(u)\chi(v)$ pro každá dvě slova u, v nad abecedou $\mathcal{A}$. Obraz slova při morfismu tedy můžeme konstruovat aplikací morfismu na jednotlivá písmena. Morfismus lze tímto způsobem aplikovat i na nekonečná slova, $\chi(\mathbf{u}) = \chi(u_0 u_1 u_2 \dots) = \chi(u_0)\chi(u_1)\chi(u_2)\dots$.

Nekonečné slovo $\mathbf{v}$ nazveme *pevným bodem* morfismu χ, pokud $\chi(\mathbf{v}) = \mathbf{v}$.

Příklad 1.2. Nejznámějším aperiodickým slovem je slovo *Fibonacciho*, které je pevným bodem morfismu

$$\begin{aligned} \varphi : \mathtt{a} &\to \mathtt{ab}, \\ \mathtt{b} &\to \mathtt{a}. \end{aligned}$$

Prefix slova pak můžeme generovat postupnou aplikací φ na písmeno $\mathtt{a}$.

$$\mathtt{a} \to \varphi(\mathtt{a}) = \mathtt{ab} \to \varphi(\mathtt{ab}) = \mathtt{aba} \to \varphi(\mathtt{aba}) = \mathtt{abaab} \to \varphi(\mathtt{abaab}) = \mathtt{abaababa} \quad \cdots$$

Kapitola 2

Kritický exponent

V předchozí kapitole jsme již využili značení $(v)^\omega$ pro nekonečné opakování slova v. Nyní zavedeme pojem *mocnina* a *exponent*. Máme-li přirozené číslo $n \in \mathbb{N}$ a konečné slovo s, pak s^n znamená zřetězení slova s n-krát za sebou, například $s^4 = ssss$. Říkáme, že s^n je n-tá *mocnina* slova s. Mocnina může být i racionální, je-li $l = |s| > 0$, pak pro $\frac{k}{l} \in \mathbb{Q}$ výraz $s^{\frac{k}{l}}$ značí prefix délky k z nekonečného slova s^ω.

Naopak je-li z neprázdným prefixem nekonečného slova u^ω, kde u je minimální délky, pak můžeme psát $z = u^e$ a racionální číslo $e = \frac{|z|}{|u|}$ nazýváme *exponentem*.

Pro neprázdný faktor u nekonečného slova $\mathbf{u}$ definujeme *index* u v $\mathbf{u}$ jako supremum ze všech možných exponentů e takových, aby u^e bylo stále ještě faktorem $\mathbf{u}$:

$$\mathrm{ind}_{\mathbf{u}}(u) = \sup\{e \in \mathbb{Q} : u^e \in \mathcal{L}(\mathbf{u})\}.$$

Pozorování. Index faktoru u nekonečného slova $\mathbf{u}$ je vždy racionální, nebo $+\infty$.

Důkaz. Pokud se ve slově vyskytuje u^n pro všechna $n \in \mathbb{N}$, pak je index faktoru u roven $+\infty$. V opačném případě existuje pouze konečný počet přirozených čísel $k \in \mathbb{N}$ tak, že $u^{\frac{k}{|u|}} \in \mathcal{L}(\mathbf{u})$. Supremum je pak maximem z konečné množiny racionálních čísel. $\square$

Příklad 2.1. Opět využijeme slovo *ananas*:

$$(ananas)^3 = ananasananasananas,$$
$$(ananas)^{\frac{1}{6}} = a,$$
$$(ananas)^{\frac{5}{3}} = (ananas)^{\frac{10}{6}} = ananasanan.$$

Vezmeme-li $\mathbf{u} = (\mathtt{abb})^\omega$, pak $\mathrm{ind}_{\mathbf{u}}(\mathtt{b}) = 2$, $\mathrm{ind}_{\mathbf{u}}(\mathtt{bb}) = 1$, $\mathrm{ind}_{\mathbf{u}}(\mathtt{abbabb}) = +\infty$, $\mathrm{ind}_{\mathbf{u}}(\mathtt{abbab}) = 1$.

Můžeme vidět, že pro $\mathbf{u}$ periodické, tzn. $\mathbf{u} = w^\omega$, má každý faktor délky $n \cdot |w|$ pro $n \in \mathbb{N}$ nekonečný index. I v aperiodických slovech může existovat neprázdný faktor, který má nekonečný index, ve *stejnoměrně rekurentních aperiodických* slovech to však již nenastává. Jinak by totiž musel existovat neprázdný faktor u takový, že pro všechna $n \in \mathbb{N}$ platí $u^n \in \mathcal{L}(\mathbf{u})$. Takové slovo pak ale buď není aperiodické, nebo není stejnoměrně rekurentní.

Příklad 2.2. Například slovo $\mathbf{u} = \mathtt{ababbabbbabbbb}\ldots$, které je tvořeno zřetězením faktorů $\mathtt{ab}^n$ pro všechna $n \in \mathbb{N}$, je aperiodické, ale není stejnoměrně rekurentní (mezera mezi výskyty písmena a nabývá všech přirozených čísel, není proto stejně omezená). Navíc $\mathrm{ind}_{\mathbf{u}}(\mathtt{b}) = +\infty$.

Kritický exponent nekonečného slova pak definujeme jako supremum ze všech exponentů všech neprázdných faktorů slova

$$\mathrm{E}(\mathbf{u}) = \sup\{e \in \mathbb{Q} : \text{ existují neprázdné faktory } x, y \in \mathcal{L}(\mathbf{u}) \text{ a } x^e = y\}.$$

Pokud je $\mathbf{u}$ stejnoměrně rekurentní, všechny neprázdné faktory mají konečné indexy a lze psát $\mathrm{E}(\mathbf{u}) = \sup\{\mathrm{ind}_{\mathbf{u}}(u) : u \in \mathcal{L}(\mathbf{u}) \setminus \{\varepsilon\}\}$. I v tomto případě, dokonce i pro sturmovská slova, může vyjít kritický exponent nekonečno. Jsou-li indexy všech neprázdných faktorů konečné, definujeme *asymptotický kritický exponent* jako

$$\mathrm{E}^*(\mathbf{u}) = \limsup_{n\to+\infty}(\max\{\mathrm{ind}_{\mathbf{u}}(u) : u \in \mathcal{L}(\mathbf{u}) \text{ a } |u| = n\}).$$

V opačném případě klademe $\mathrm{E}^*(\mathbf{u}) = +\infty$. Je-li kritický exponent konečný, lze asymptotický kritický exponent počítat jako

$$\mathrm{E}^*(\mathbf{u}) = \lim_{n\to\infty} \sup\{e \in \mathbb{Q} : \text{ existují faktory } x, y \in \mathcal{L}(\mathbf{u}), \text{ kde } |x| > n \text{ a } y = x^e\}.$$

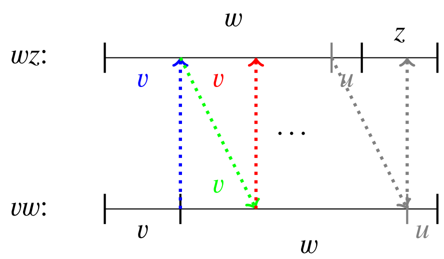
Z definic je vidět, že vždy platí $\mathrm{E}^*(\mathbf{u}) \leq \mathrm{E}(\mathbf{u})$. Pokud nastává, že $\mathrm{E}(\mathbf{u}) = +\infty$, pak i $\mathrm{E}^*(\mathbf{u}) = +\infty$.

Dále se budeme věnovat výpočtu asymptotického kritického exponentu a kritického exponentu pro stejnoměrně rekurentní aperiodická slova. Inspiraci a výsledky čerpáme z [5]. Pro stejnoměrně rekurentní aperiodická slova totiž existuje jednodušší vzorec pro výpočet, k jeho odvození je zásadní následující věta.

Věta 2.1. *Máme-li w, v neprázdné faktory rekurentního slova $\mathbf{u}$, kde v je návratové slovo k w, pak lze psát $w = v^e$ pro nějaké $e \in \mathbb{Q}$.*

Důkaz.

- Protože v je návratové slovo k w, je w prefixem vw. Proto existuje $z \in \mathcal{L}(\mathbf{u})$ tak, že $vw = wz$.
  - Pokud $|v| \geq |w|$, pak je w jistě prefixem v, a proto $w = v^e$, kde $e = \frac{|w|}{|v|} \leq 1$.
  - Uvažujme dále případ $|v| < |w|$.
    Nutně platí $|v| = |z|$. Dále z rovnosti $vw = wz$ získáme, že wz má v jako prefix (viz obrázek 2.1). Tuto úvahu lze provádět opakovaně. Najdeme tedy $n \in \mathbb{N}$ tak, že $(n-1)|v| < |w| \leq n|v|$. Pak na začátku w najdeme $n-1$ faktorů v a zbytek tvoří $u \in \mathcal{L}(\mathbf{u})$. Obdobným postupem od konce získáme, že na konci w je $(n-1)$ faktorů z a zbytek tvoří nějaký faktor ℓ. Proto $w = v^{n-1}u = \ell z^{n-1}$, kde $|u| = |\ell| = |w| - (n-1)|v|$. Dále z rovnosti $vw = wz$ získáme $vw = v^n u = \ell z^n$, a proto se musí rovnat i jejich prefixy a sufixy. Proto $v = \ell k$ a $z = ku$ pro nějaké $k \in \mathcal{L}(\mathbf{u})$. Nakonec si uvědomíme, že v posledním kroku dojde k překryvu v a z tak, že $v\ell = \ell k \ell = uz = uku$, proto $\ell = u$.
    Celkem tedy získáme $v = uk$, $z = ku$ a $w = v^{n-1}u = (uk)^{n-1}u$. Z toho získáme, že w je prefixem nekonečného slova $(uk)^\omega = v^\omega$, a proto $w = v^e$, kde $e = n - 1 + \frac{|u|}{|v|}$.



Obrázek 2.1: $vw = wz$

□

Příklad 2.3. Prefix aperiodického stejnoměrně rekurentního slova nad binární abecedou může vypadat například takto

$$\mathtt{ab\underline{aba}ababab\underline{aba}babaababa}.$$

Prvním bispeciálem je písmeno $\mathtt{a}$ a přísluší k němu návratová slova $\mathtt{a}, \mathtt{ab}$. Zřejmě $\mathtt{a} = \mathtt{a}^1 = (\mathtt{ab})^{\frac{1}{2}}$. Dalším bispeciálem je například $\mathtt{aba}$, v prefixu jsou označeny dva jeho výskyty lišící se předcházejícím a následujícím písmenem. Tento faktor má návratová slova $\mathtt{ab}, \mathtt{aba}$ a opět platí $\mathtt{aba} = (\mathtt{ab})^{\frac{3}{2}} = (\mathtt{aba})^1$.

Lemma 2.1. *Nechť* $\mathbf{u}$ *je aperiodické stejnoměrně rekurentní slovo. Pak každé návratové slovo je návratovým slovem pouze ke konečnému počtu faktorů.*

Důkaz. Pro spor předpokládejme, že u je návratovým slovem nekonečně mnoha faktorů. Protože existuje pouze konečný počet faktorů dané délky a tyto faktory jsou mocninami u, získáme $u^n \in \mathcal{L}(\mathbf{u})$ pro všechna $n \in \mathbb{N}$. Pak je $\mathbf{u}$ buď periodické, nebo není stejnoměrně rekurentní, což je spor. □

Lemma 2.2. *Nechť* $\mathbf{u}$ *je stejnoměrně rekurentní aperiodické slovo a* $f \in \mathcal{L}(\mathbf{u})$ *je neprázdný faktor takový, že* $\mathrm{ind}_{\mathbf{u}}(f) > 1$. *Pak existuje faktor* $u \in \mathcal{L}(\mathbf{u})$ *a bispeciál* $w \in \mathcal{L}(\mathbf{u})$ *splňující* $|f| = |u|$, *dále* $\mathrm{ind}_{\mathbf{u}}(f) \leq \mathrm{ind}_{\mathbf{u}}(u) = 1 + \frac{|w|}{|u|}$ *a navíc* u *je zřetězením návratových slov k* w.

Důkaz. Uvažujme faktor $u \in \mathcal{L}(\mathbf{u})$ splňující

$$\mathrm{ind}_{\mathbf{u}}(u) = \max\{\mathrm{ind}_{\mathbf{u}}(v) : v \in \mathcal{L}(u) \wedge |v| = |f|\} > 1.$$

Dále k faktoru u nalezneme bispeciál w. Označíme $\mathrm{ind}_{\mathbf{u}}(u) = k + \alpha$, kde $k \in \mathbb{N}$, $\alpha \in [0, 1)$. Pak faktor $u := u'u''$, kde $u' = u^\alpha$ a u'' je neprázdný, protože $\alpha < 1$. Dále $u^{k+\alpha} = (u'u'')^k u'$.

Označme a první a b poslední písmeno faktoru u''. Protože $\mathbf{u}$ je rekurentní, existují písmena x a y z abecedy tak, že $xu^{k+\alpha}y \in \mathcal{L}(\mathbf{u})$. Pak $xu^{k+\alpha}y = x(u'u'')^k u'y \in \mathcal{L}(\mathbf{u})$, a proto faktory $x(u'u'')^{k-1}u'a$ a $b(u'u'')^{k-1}u'y$ jsou faktory slova $\mathbf{u}$. Dále si uvědomíme, že jistě

- $y \neq a$, jinak $(u'u'')^k u'a = u^e$, kde $e > k + \alpha$.
- $x \neq b$, jinak $b(u'u'')^k u' = v^k bu' = v^e$, kde $e > k + \alpha$. Navíc v získáme z u přidáním písmena b na začátek a odebráním b z konce, tzn. $v = bub^{-1}$, proto $|v| = |u|$. Dostáváme spor s maximem $\mathrm{ind}_{\mathbf{u}}(u)$.

Faktor $w := (u'u'')^{k-1}u'$ je proto bispeciálem v $\mathbf{u}$. Navíc $u^{k+\alpha} = uw$, a proto je u zřetězením návratových slov k w a $\mathrm{ind}_{\mathbf{u}}(u) = k + \alpha = 1 + \frac{|w|}{|u|}$. □

Již máme vše potřebné k dokázání následující věty, která bude pro náš výpočet klíčová.

Věta 2.2. *Nechť* $\mathbf{u}$ *je stejnoměrně rekurentní aperiodické slovo a* $(w_n)_{n\in\mathbb{N}}$ *je posloupnost všech bispeciálních faktorů* $\mathbf{u}$ *seřazených dle délky. Dále nechť pro každé* $n \in \mathbb{N}$ *je* v_n *nejkratší návratové slovo k* w_n *v* $\mathbf{u}$. *Pak*

$$\mathrm{E}(\mathbf{u}) = 1 + \sup_{n\in\mathbb{N}}\left\{\frac{|w_n|}{|v_n|}\right\} \qquad a \qquad \mathrm{E}^*(\mathbf{u}) = 1 + \limsup_{n\to\infty}\frac{|w_n|}{|v_n|}.$$

Důkaz.

- Z věty 2.1 vyplývá, že pro všechny bispeciály platí, že $v_n w_n = v_n^{e_n} \in \mathcal{L}(\mathbf{u})$ pro $e_n = 1 + \frac{|w_n|}{|v_n|}$. Proto $\mathrm{ind}_{\mathbf{u}}(v_n) \geq 1 + \frac{|w_n|}{|v_n|}$ a $\mathrm{E}(\mathbf{u}) \geq 1 + \sup\left\{\frac{|w_n|}{|v_n|}\right\} > 1$.

  Z lemmatu 2.1 pak plyne, že $\lim_{n\to+\infty}|v_n| = +\infty$, proto i $\mathrm{E}^*(\mathbf{u}) \geq 1 + \limsup_{n\to+\infty}\frac{|w_n|}{|v_n|} \geq 1$.

- Nechť $\delta > 0$ taková, že $\mathrm{E}(\mathbf{u}) - \delta > 1$. Pak z definice suprema nalezneme $f \in \mathcal{L}(\mathbf{u})$ tak, že $1 < \mathrm{E}(\mathbf{u}) - \delta < \mathrm{ind}_{\mathbf{u}}(f)$ a z lemmatu 2.2 vyplývá, že existují alespoň jeden bispeciál w a faktor u takové, že $\mathrm{ind}_{\mathbf{u}}(f) \leq \mathrm{ind}_{\mathbf{u}}(u) = 1 + \frac{|w|}{|u|}$ a u je zřetězením návratových slov k w. Proto existuje $m \in \mathbb{N}$ tak, že $w = w_m$ a $|u| \geq |v_m|$. Celkem tedy máme

$$\mathrm{E}(\mathbf{u}) - \delta < \mathrm{ind}_{\mathbf{u}}(f) \leq \mathrm{ind}_{\mathbf{u}}(u) = 1 + \frac{|w|}{|u|} \leq 1 + \frac{|w_m|}{|v_m|} \leq 1 + \sup\left\{\frac{|w_n|}{|v_n|}\right\}.$$

  Protože δ můžeme volit libovolně malé, platí $\mathrm{E}(\mathbf{u}) \leq 1 + \sup\left\{\frac{|w_n|}{|v_n|}\right\}$. V předchozím bodě jsme dokázali opačnou nerovnost, rovnost ve větě proto platí.

- Pokud je $\mathrm{E}^*(\mathbf{u}) = 1$, pak s využitím prvního bodu $\mathrm{E}^*(\mathbf{u}) \geq 1 + \limsup_{n\to+\infty} \frac{|w_n|}{|v_n|} \geq 1 = \mathrm{E}^*(\mathbf{u})$ a rovnost je splněna.

  Pokud $\mathrm{E}^*(\mathbf{u}) > 1$, pak existuje posloupnost faktorů $(f^{(n)})_{n\in\mathbb{N}} \subset \mathcal{L}(\mathbf{u})$, kde $\mathrm{ind}_{\mathbf{u}}(f^{(n)}) > 1$ pro všechna $n \in \mathbb{N}$, taková, že

$$\lim_{n\to+\infty} |f^{(n)}| = +\infty \text{ a také } \mathrm{E}^*(\mathbf{u}) = \lim_{n\to+\infty} \mathrm{ind}_{\mathbf{u}}(f^{(n)}).$$

  Pro každé přirozené n nalezneme $w^{(n)}$ bispeciál takový, že

$$\mathrm{ind}_{\mathbf{u}}(f^{(n)}) \leq 1 + \frac{|w^{(n)}|}{|u^{(n)}|} \leq 1 + \frac{|w^{(n)}|}{|v^{(n)}|},$$

  kde $u^{(n)}$ je zřetězením návratových slov k $w^{(n)}$ a $v^{(n)}$ je nejkratší návratové slovo k $w^{(n)}$, a proto $|u^{(n)}| \geq |v^{(n)}|$.

  Proto

$$\mathrm{E}^*(\mathbf{u}) = \lim_{n\to+\infty} \mathrm{ind}_{\mathbf{u}}(f^{(n)}) \leq 1 + \limsup_{n\to+\infty} \frac{|w_n|}{|v_n|}$$

  a opačná nerovnost opět platí. □

Klíčové pro nás tedy bude dále prozkoumat délky bispeciálů a jejich návratových slov v aperiodických stejnoměrně rekurentních slovech.

Kapitola 3

Sturmovská slova

Sturmovská slova jsou nejznámější třída aperiodických slov nad binární abecedou. Existuje mnoho různých ekvivalentních definic v závislosti na tom, jakou vlastnost právě chceme zkoumat.

Definice 3.1. Faktorová komplexita *nekonečného slova* $\mathbf{u}$ *je funkce* $C_{\mathbf{u}} : \mathbb{N} \to \mathbb{N}$*, která přirozenému číslu* n *přiřadí počet různých faktorů délky* n*, které se ve slově* $\mathbf{u}$ *vyskytují, tzn.* $C_{\mathbf{u}}(n) = \#\{u \in \mathcal{L}(\mathbf{u}) : |u| = n\}$.

Můžeme vidět, že pro posléze periodická slova je faktorová komplexita omezená – pro každou délku existuje nejvýše tolik faktorů, kolik je součet délek předperiody a periody. Dle definice z [10] jsou sturmovská slova *aperiodická slova s nejmenší faktorovou komplexitou*, platí pro ni $C_{\mathbf{u}}(n) = n + 1$ pro každé n přirozené. Ze vzorce pro faktorovou komplexitu vyplývá, že pro každou délku n existuje právě jeden pravý, resp. levý speciál. ($n - 1$ faktorů se totiž jednoznačně prodlouží doprava, resp. doleva, a z toho posledního vzniknou 2 různé faktory přidáním jednoho písmena doprava, resp. doleva.)

Důležitá vlastnost sturmovských slov je uvedena v [10] – aperiodické slovo nad binární abecedou je balancované právě tehdy, když je sturmovské.

Další ekvivalentní definice uvedená v [18] říká, že nekonečné slovo $\mathbf{u}$ je sturmovské právě tehdy, když každý faktor $\mathbf{u}$ má právě dvě návratová slova.

Protože sturmovská slova jsou rekurentní, lze je napsat jako zřetězení návratových slov k libovolnému prefixu. Označíme-li vybraný prefix u, pak z předchozí definice $\mathcal{R}_{\mathbf{u}}(u) = \{r, s\}$, a slovo můžeme zapsat jako posloupnost nad dvoupísmennou abecedou $\mathcal{A} = \{r, s\}$. Takové posloupnosti říkáme *derivované slovo*, značíme $\mathrm{d}_{\mathbf{u}}(u)$. Není-li vybraný faktor prefixem, pak derivované slovo konstruujeme stejným postupem počínaje jeho prvním výskytem v $\mathbf{u}$ (prefix, který u neobsahuje, do $\mathrm{d}_{\mathbf{u}}(u)$ nezahrnujeme). Je dokázáno, že i derivované slovo je sturmovské.

Máme-li faktor u sturmovského slova $\mathbf{u}$, pak za prefixové návratové slovo k u budeme považovat faktor reprezentující první písmeno, které se v derivovaném slovu $\mathrm{d}_{\mathbf{u}}(u)$ vyskytne. Druhé návratové slovo budeme označovat jako neprefixové. Tato definice odpovídá tomu, že pokud je faktor u prefixem slova $\mathbf{u}$, pak i jeho prefixové návratové slovo bude prefixem.

Dále se budeme držet konvence, že na první místo v Parikhově vektoru budeme psát počet výskytů písmena $\mathtt{a}$, resp. $\mathtt{s}$ (obvykle to méně časté) a na druhé místo počet výskytů druhého písmena $\mathtt{b}$, resp. $\mathtt{r}$ (obvykle to častější).

Následující definice a věta budou pro nás důležité dále při práci s balancovanými slovy.

Definice 3.2. *Aperiodické slovo* $\mathbf{u}$ *nad abecedou* $\{\mathtt{a}, \mathtt{b}\}$ *má* dobře rozmístěné výskyty *(Well Distributed Occurrences, **WDO**), pokud pro každé* $m \in \mathbb{N}$ *a každé* $w \in \mathcal{L}(\mathbf{u})$

$$\left\{\vec{V}(p) \mod m : pw \text{ je prefixem } \mathbf{u}\right\} = \mathbb{Z}_m^2 = \left\{\begin{pmatrix}\ell\\ k\end{pmatrix} : k, \ell \in \{0, 1, \dots, m-1\}\right\}.$$

Věta 3.1. [3] *Sturmovská slova mají WDO.*

Sturmovské slovo $\mathbf{u}$ nazveme *standardní*, pokud každý prefix slova $\mathbf{u}$ je levý speciál. Bispeciály ve standardním sturmovském slově jsou jeho palindromické prefixy.

Věta 3.2. [14] *Ke každému sturmovskému slovu* $\mathbf{u}'$ *existuje standardní sturmovské slovo* $\mathbf{u}$ *takové, že jejich jazyky se rovnají, tzn.* $\mathcal{L}(\mathbf{u}') = \mathcal{L}(\mathbf{u})$.

Protože nás zajímá pouze jazyk nekonečných slov, dále budeme pracovat pouze se standardními sturmovskými slovy.

K popisu standardních sturmovských slov zavedeme dva morfismy:

$$G : \mathtt{a} \to \mathtt{a}, \quad \mathtt{b} \to \mathtt{ab}, \qquad D : \mathtt{a} \to \mathtt{ba}, \quad \mathtt{b} \to \mathtt{b}.$$

Je dokázáno [12], že pro $\mathbf{u}$ standardní sturmovské slovo existuje právě jedna posloupnost morfismů G a D, označíme $\Delta = \Delta_0\Delta_1\cdots$, a právě jedna posloupnost standardních sturmovských slov $(\mathbf{u}^{(n)})_{n\geq 1}$ taková, že pro každé n přirozené platí $\mathbf{u} = \Delta_0\Delta_1\cdots\Delta_{n-1}(\mathbf{u}^{(n)})$. Navíc Δ obsahuje D i G nekonečněkrát, tzn. $\Delta = G^{a_1}D^{a_2}G^{a_3}\cdots$ nebo $\Delta = D^{a_1}G^{a_2}D^{a_3}\cdots$, kde a_i je přirozené číslo pro každé $i \in \mathbb{N}$. Posloupnost Δ nazýváme *řídící posloupnost* slova $\mathbf{u}$.

K $\mathbf{u}$ s řídící posloupností $\Delta = G^{a_1}D^{a_2}G^{a_3}\cdots$ nebo $\Delta = D^{a_1}G^{a_2}D^{a_3}\cdots$ přiřazujeme iracionální číslo $\theta \in (0,1)$, jehož řetězový zlomek odpovídá posloupnosti exponentů v řídící posloupnosti

$$\theta = [0, a_1, a_2, a_3, a_4, \dots] = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \ddots}}}}.$$

K řetězovému zlomku $\theta = [0, a_1, a_2, a_3, a_4, \dots]$ definujeme posloupnosti (p_n) a (q_n), které budeme při výpočtech velmi často potřebovat.

Definice 3.3. *Pro* $\theta = [0, a_1, a_2, \dots]$ *definujeme posloupnosti* $(p_n)_{n\geq -1}$ *a* $(q_n)_{n\geq -1}$ *pomocí rekurence*

$$X_{n+1} = a_{n+1}X_n + X_{n-1} \text{ pro } n \geq 0$$

s počátečními podmínkami $p_{-1} = 1, p_0 = 0,\ q_{-1} = 0, q_0 = 1$.

Pak z vlastností řetězových zlomků platí

$$\theta = \lim_{n\to\infty} \frac{p_n}{q_n}.$$

Parametr θ bude hrát stěžejní roli při popisu bispeciálů a jejich návratových slov ve sturmovských a posléze i balancovaných slovech. Jde tedy o klíčový pojem pro odvození (asymptotického) kritického exponentu balancovaných slov.

3.1 Bispeciály a jejich návratová slova

Při výpočtu kritického exponentu budeme potřebovat umět generovat bispeciály sturmovských slov a dále s nimi počítat. Budeme tedy hojně využívat následující větu.

Věta 3.3. *Nechť* $\mathbf{u}$ *a* $\mathbf{u}'$ *jsou standardní sturmovská slova,* $\mathbf{u} = D(\mathbf{u}')$*. Pak platí následující tvrzení.*

1. *Je-li* z' *bispeciál v* $\mathbf{u}'$*, pak* $z = D(z')\mathtt{b}$ *je bispeciál v* $\mathbf{u}$*,*
2. *je-li* z *neprázdný bispeciál v* $\mathbf{u}$*, pak* $z = D(z')\mathtt{b}$ *pro jednoznačně daný bispeciál* z' *v* $\mathbf{u}'$*,*
3. r' *a* s' *jsou návratová slova k bispeciálu* z' *v* $\mathbf{u}'$ *právě tehdy, když* $r = D(r')$ *a* $s = D(s')$ *jsou návratová slova k bispeciálu* $z = D(z')\mathtt{b}$ *v* $\mathbf{u}$*,*
4. $\mathrm{d}_{\mathbf{u}}(z) = \mathrm{d}_{\mathbf{u}'}(z')$.

Důkaz. 1. Je-li z' bispeciál v $\mathbf{u}'$, pak $z'\mathtt{a}, z'\mathtt{b}, \mathtt{a}z', \mathtt{b}z' \in \mathcal{L}(\mathbf{u}')$, a tudíž $D(z'\mathtt{a}), D(z'\mathtt{b}), D(\mathtt{a}z'), D(\mathtt{b}z') \in \mathcal{L}(\mathbf{u})$. Napřed se podíváme na prodloužení doprava. $D(z'\mathtt{a}) = D(z')D(\mathtt{a}) = D(z')\mathtt{ba}$ a $D(z'\mathtt{b}) = D(z')\mathtt{b}$. Protože obrazem $\mathtt{a}$ při morfismu D je $\mathtt{ba}$ a $\mathtt{b}$ se přepíše na $\mathtt{b}$, pak za obrazem $z'\mathtt{b}$ při morfismu D vždy následuje opět $\mathtt{b}$. Proto $D(z')\mathtt{ba}$ a $D(z')\mathtt{bb}$ jsou faktory $\mathbf{u}$, a tudíž $D(z')\mathtt{b}$ je pravý speciál v $\mathbf{u}$.

   Dále $D(\mathtt{a}z') = \mathtt{ba}D(z')$ a $D(\mathtt{b}z') = \mathtt{b}D(z')$, a proto $D(z')$ je levý speciál v $\mathbf{u}$. Protože za $D(z')$ vždy následuje $\mathtt{b}$ (viz předchozí odstavec), pak i $D(z')\mathtt{b}$ je levý speciál v $\mathbf{u}$.

   Celkem je tedy $D(z')\mathtt{b}$ bispeciálem v $\mathbf{u}$.

2. Je-li z bispeciálem v $\mathcal{L}(\mathbf{u})$, pak $\mathtt{a}z, \mathtt{b}z, z\mathtt{a}, z\mathtt{b} \in \mathcal{L}(\mathbf{u})$. Pak z jistě začíná i končí na písmeno $\mathtt{b}$, protože z tvaru morfismu D nikdy nenastane, aby $\mathtt{aa}$ bylo faktorem $D(\mathbf{u}')$. Z toho vyplývá, že existuje $z' \in \mathcal{L}(\mathbf{u}')$ tak, že $z = D(z')\mathtt{b}$ (obrazy obou písmen při morfismu D začínají na $\mathtt{b}$). Navíc je z' dáno jednoznačně.

   Dále platí $\mathtt{ba}z = D(\mathtt{a})D(z')\mathtt{b} = D(\mathtt{a}z')\mathtt{b}$ a $\mathtt{b}z = D(\mathtt{b})D(z')\mathtt{b} = D(\mathtt{b}z')\mathtt{b}$. Tudíž z' je levým speciálem v $\mathbf{u}'$.
   $z\mathtt{a} = D(z')\mathtt{ba} = D(z'\mathtt{a})$ a $z\mathtt{b} = D(z')\mathtt{bb} = D(z'\mathtt{b})\mathtt{b}$, proto $z'\mathtt{a}, z'\mathtt{b} \in \mathcal{L}(\mathbf{u}')$, a z' je tudíž pravý speciál.

   Celkem tedy z' je bispeciálem v $\mathbf{u}'$ a platí $z = D(z')\mathtt{b}$.

3. r', s' jsou návratová slova k z' v $\mathbf{u}'$, proto $r'z', s'z' \in \mathcal{L}(\mathbf{u}')$. Dále si uvědomíme, že za $D(z')$ vždy následuje $\mathtt{b}$. Pak $D(r'z')\mathtt{b} = D(r')D(z')\mathtt{b}$ a také $D(s'z')\mathtt{b} = D(s')D(z')\mathtt{b} \in \mathcal{L}(\mathbf{u})$. Nakonec si stačí uvědomit, že pokud $r'z'$ má z' jako prefix, pak i $D(r')D(z')\mathtt{b}$ má $D(z')\mathtt{b}$ jako prefix a uprostřed se $D(z')\mathtt{b}$ nevyskytuje.

   Naopak pokud r, s jsou návratová slova k bispeciálu z, pak rz, $sz \in \mathcal{L}(\mathbf{u})$ a rz i sz mají z jako prefix. Existují tedy faktory $r', s' \in \mathcal{L}(\mathbf{u}')$ tak, že $rz = D(r'z')\mathtt{b}$ a $sz = D(s'z')\mathtt{b}$. Faktory $r'z', s'z'$ mají z' jako prefix, protože faktor z vznikl ze z' jednoznačně. Z jednoznačnosti také plyne, že z' se uprostřed $r'z'$ nevyskytuje.

4. Vyplývá z bodu 3 a z toho, že zobrazení D je morfismus, který nemění pořadí faktorů. □

Obdobná věta platí i pro morfismus G, pouze se místo $\mathtt{b}$ připojuje na konec písmeno $\mathtt{a}$. Protože víme, že v každém sturmovském slově je ε bispeciálem (mohou po něm následovat obě písmena i obě být před ním), můžeme takto při znalosti Δ generovat stále delší bispeciály. Bispeciály budeme číslovat podle počtu použitých morfismů nezápornými celými čísly n, kde prázdnému slovu přísluší $n = 0$. Toto číslování odpovídá dřívějšímu číslování dle délky. K tomuto číslování ještě přidáme následující definici.

Definice 3.4. *Nechť* $\mathbf{u}$ *je standardní sturmovské slovo a* u_n *je* n*-tý bispeciál v* $\mathbf{u}$. *K bispeciálu* u_n *jednoznačně přiřazujeme dvojici* (N, m) *splňující* $n = a_0 + a_1 + \cdots + a_N + m$, *kde* $0 \leq m < a_{N+1}$, $a_0 = 0$ *a* a_i *jsou pro všechna* $i \in \mathbb{N}$ *koeficienty řetězového zlomku* θ. *Pokud* $m = 0$, *říkáme, že* u_n *je* primární bispeciál. *Pro* $m > 0$ *jde o* sekundární bispeciál.

Dále si odvodíme vztahy pro Parikhovy vektory bispeciálů a jejich návratových slov. K tomu budeme potřebovat následující větu a lemmata.

Věta 3.4. *Nechť* b' *je* k*-tý bispeciál standardního sturmovského slova* $\mathbf{u}'$ *a* r' *návratové slovo k* b' *v* $\mathbf{u}'$.

1. *Je-li* $\mathbf{u} = D^h(\mathbf{u}')$, *pak* $(k+h)$*-tý bispeciál* b *v* $\mathbf{u}$ *a návratové slovo* r *k* b *splňují*

$$\vec{V}(b) = \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix} \vec{V}(b') + h \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$
$$\vec{V}(r) = \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix} \vec{V}(r').$$

2. *Je-li* $\mathbf{u} = G^h(\mathbf{u}')$, *pak* $(k+h)$*-tý bispeciál* b *v* $\mathbf{u}$ *a návratové slovo* r *k* b *splňují*

$$\vec{V}(b) = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \vec{V}(b') + h \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$
$$\vec{V}(r) = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \vec{V}(r').$$

Důkaz. 1. $(k+h)$-tý bispeciál v $\mathbf{u}$ získáme postupem popsaným ve větě 3.3, tzn. budeme postupně aplikovat morfismus D a připojovat na konec písmeno $\mathtt{b}$, tzn. $b = D(D(\cdots(D(b')\mathtt{b})\cdots)\mathtt{b})\mathtt{b}$. Protože $D(\mathtt{b}) = \mathtt{b}$, můžeme $\mathtt{b}$ ze závorek vytknout a získáme $b = D^h(b')\mathtt{b}^h$. Dále si stačí uvědomit, jak morfismus D mění počet písmenek ve slově – na $\mathtt{b}$ působí jako identita, za každé písmeno $\mathtt{a}$ přibude faktor $\mathtt{ba}$. Proto počet písmen $\mathtt{a}$ se zachová, počet písmen $\mathtt{b}$ v $D(b')$ je stejný jako součet počtu písmen $\mathtt{a}$ a $\mathtt{b}$ v b'. Proto maticově

$$\vec{V}(b) = \vec{V}(D^h(b')) + \vec{V}(\mathtt{b}^h) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^h \vec{V}(b') + h \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix} \vec{V}(b') + \begin{pmatrix} 0 \\ h \end{pmatrix}.$$

Návratové slovo r získáme podle věty 3.3 aplikací h morfismů D na r', proto $r = D^h(r')$ a dále

$$\vec{V}(r) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^h \vec{V}(r') = \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix} \vec{V}(r').$$

2. Pro morfismus G je postup podobný jako pro D. Zde se na konec připojuje písmeno $\mathtt{a}$, ale opět $G(\mathtt{a}) = \mathtt{a}$. Takže $b = G^h(b')\mathtt{a}^h$. Matici morfismu G získáme podobně, pouze nyní se nemění počet $\mathtt{b}$, počet $\mathtt{a}$ v $G(b')$ je stejný jako součet počtu $\mathtt{a}$ a $\mathtt{b}$ v původním faktoru. Proto

$$\vec{V}(b) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^h \vec{V}(b') + h \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \vec{V}(b') + \begin{pmatrix} h \\ 0 \end{pmatrix},$$
$$\vec{V}(r) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^h \vec{V}(r') = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \vec{V}(r').$$

□

Již víme, jak vyrábět bispeciály pomocí morfismů z řídící posloupnosti. Pro odvození vztahu pro jejich délky budeme potřebovat následující lemmata.

Lemma 3.1. *Pro všechna $N \in \mathbb{N}$ platí*

$$a_N \begin{pmatrix} p_{N-1} \\ q_{N-1} \end{pmatrix} + a_{N-1} \begin{pmatrix} p_{N-2} \\ q_{N-2} \end{pmatrix} + \cdots + a_2 \begin{pmatrix} p_1 \\ q_1 \end{pmatrix} + a_1 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = \begin{pmatrix} p_N \\ q_N \end{pmatrix} + \begin{pmatrix} p_{N-1} \\ q_{N-1} \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Důkaz. Důkaz provedeme matematickou indukcí.

- $N = 1$: Platí $\begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ a $\begin{pmatrix} p_{-1} \\ q_{-1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Z definice pak $p_1 = a_1 p_0 + p_{-1} = 1$, $q_1 = a_1 q_0 + q_{-1} = a_1$. Dokazovaný vztah se dá pak přepsat jako

$$a_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ a_1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

  a můžeme vidět, že platí.

- $N \to N + 1$:

$$a_{N+1} \begin{pmatrix} p_N \\ q_N \end{pmatrix} + a_N \begin{pmatrix} p_{N-1} \\ q_{N-1} \end{pmatrix} + \cdots + a_2 \begin{pmatrix} p_1 \\ q_1 \end{pmatrix} + a_1 \begin{pmatrix} p_0 \\ q_0 \end{pmatrix} = a_{N+1} \begin{pmatrix} p_N \\ q_N \end{pmatrix} + \begin{pmatrix} p_N \\ q_N \end{pmatrix} + \begin{pmatrix} p_{N-1} \\ q_{N-1} \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$a_{N+1} \begin{pmatrix} p_N \\ q_N \end{pmatrix} + \begin{pmatrix} p_{N-1} \\ q_{N-1} \end{pmatrix} = \begin{pmatrix} p_{N+1} \\ q_{N+1} \end{pmatrix},$$

$$a_{N+1} \begin{pmatrix} p_N \\ q_N \end{pmatrix} + \begin{pmatrix} p_N \\ q_N \end{pmatrix} + \begin{pmatrix} p_{N-1} \\ q_{N-1} \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} p_{N+1} \\ q_{N+1} \end{pmatrix} + \begin{pmatrix} p_N \\ q_N \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

  kde v první rovnosti jsme využili indukčního předpokladu a druhá rovnost vyplývá z definice p_N, q_N. Rovnost je tedy dokázána. □

Lemma 3.2. *Pro všechna $N \in \mathbb{N}$ platí*

$$\begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & a_2 \\ 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_{2N-2} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a_{2N-1} & 1 \end{pmatrix} = \begin{pmatrix} p_{2N-1} & p_{2N-2} \\ q_{2N-1} & q_{2N-2} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & a_2 \\ 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 \\ a_{2N-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & a_{2N} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} p_{2N-1} & p_{2N} \\ q_{2N-1} & q_{2N} \end{pmatrix}.$$

Důkaz. Důkaz opět provedeme matematickou indukcí.

- $N = 1$: Rozepíšeme si, co znamenají členy na každé straně a využijeme při tom počátečních podmínek a tvaru rekurence pro p_N, q_N.

$$\begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} = \begin{pmatrix} p_1 & p_0 \\ q_1 & q_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix}.$$

  První rovnost tedy platí. Ověříme ještě druhou.

$$\begin{pmatrix} 1 & 0 \\ a_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & a_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a_2 \\ a_1 & a_1 a_2 + 1 \end{pmatrix},$$

$$\begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix} = \begin{pmatrix} 1 & a_2 \cdot 1 + 0 \\ a_1 & a_2 \cdot a_1 + 1 \end{pmatrix}.$$

- $N \to N+1$: Předpokládejme, že vztahy platí pro $2N-1$ a $2N$. Pak

$$\begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & 0\\ a_{2N-1} & 1\end{pmatrix}\begin{pmatrix}1 & a_{2N}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ a_{2N+1} & 1\end{pmatrix} = \begin{pmatrix}p_{2N-1} & p_{2N}\\ q_{2N-1} & q_{2N}\end{pmatrix}\begin{pmatrix}1 & 0\\ a_{2N+1} & 1\end{pmatrix}$$
$$= \begin{pmatrix}a_{2N+1}p_{2N} + p_{2N-1} & p_{2N}\\ a_{2N+1}q_{2N} + q_{2N-1} & q_{2N}\end{pmatrix} = \begin{pmatrix}p_{2N+1} & p_{2N}\\ q_{2N+1} & q_{2N}\end{pmatrix}.$$

V prvním kroku jsme využili indukčního předpokladu platnosti pro posloupnost matic do $2N$ a poté vlastností p_N, q_N.

Dále

$$\begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & 0\\ a_{2N+1} & 1\end{pmatrix}\begin{pmatrix}1 & a_{2N+2}\\ 0 & 1\end{pmatrix} = \begin{pmatrix}p_{2N+1} & p_{2N}\\ q_{2N+1} & q_{2N}\end{pmatrix}\begin{pmatrix}1 & a_{2N+2}\\ 0 & 1\end{pmatrix}$$
$$= \begin{pmatrix}p_{2N+1} & a_{2N+2}p_{2N+1} + p_{2N}\\ g_{2N+1} & a_{2N+2}q_{2N+1} + q_{2N}\end{pmatrix} = \begin{pmatrix}p_{2N+1} & p_{2N+2}\\ q_{2N+1} & q_{2N+2}\end{pmatrix}.$$

V prvním kroku jsme využili platnost pro posloupnost matic do $2N+1$ dokázanou v předchozím bodě a poté vlastností p_N, q_N.

Vztahy proto platí i pro $2N+1$ a $2N+2$. □

Postupnou aplikací věty 3.3 a vlastností posloupností (p_n), (q_n) získáme následující tvrzení.

Věta 3.5. [8] *Nechť u_n je n-tý bispeciál standardního sturmovského slova s řídící posloupností ve tvaru $\Delta = D^{a_1}G^{a_2}D^{a_3}\ldots$. Označíme r, resp. s prefixové, resp. neprefixové návratové slovo k u_n, jemuž přísluší dvojice (N, m). Pak:*

1. $\vec{V}(r) = \begin{pmatrix}p_N\\ q_N\end{pmatrix}$, *tudíž* $|r| = p_N + q_N$,

2. $\vec{V}(s) = \begin{pmatrix}m\cdot p_N + p_{N-1}\\ m\cdot q_N + q_{N-1}\end{pmatrix}$, *tudíž* $|s| = m(p_N + q_N) + p_{N-1} + q_{N-1}$,

3. $\vec{V}(u_n) = \vec{V}(r) + \vec{V}(s) - \begin{pmatrix}1\\ 1\end{pmatrix}$, *tudíž* $|u_n| = |r| + |s| - 2$.

Začíná-li řídící posloupnost morfismem G, pak se složky vektorů vymění, např. $\vec{V}(r) = (q_N, p_N)^T$. To nastává pro případ, kdy první písmeno standardního sturmovského slova je a.

Důkaz. Důkaz rozdělíme na dvě části podle parity N.

N sudé: Zvolme $\mathbf{u}^{(N)}$ standardní sturmovské slovo s řídící posloupností $\Delta = D^{a_{N+1}}G^{a_{N+2}}D^{a_{N+3}}\cdots$. Potom $\mathbf{u} = D^{a_1}G^{a_2}\cdots D^{a_{N-1}}G^{a_N}(\mathbf{u}^{(N)})$. Dále m-tý bispeciál b' v $\mathbf{u}^{(N)}$ získáme z prázdného slova aplikací m morfismů D a postupným připojováním písmena b, protože $m < a_{N+1}$. Proto $b' = D^m(\varepsilon)\mathrm{b}^m = \mathrm{b}^m$. Obdobně návratová slova r'_m, resp. s'_m získáme aplikací m morfismů D na návratová slova k ε, což jsou b, resp. a. ($\mathbf{u}^{(N)}$ totiž vzniklo jako obraz nějakého standardního sturmovského slova při morfismu D, častější písmeno je v něm tedy b.) Proto $r'_m = D^m(\mathrm{b}) = \mathrm{b}$ a $s'_m = D^m(\mathrm{a}) = \mathrm{b}^m\mathrm{a}$.

Dále z věty 3.3 víme, že n-tý bispeciál u_n v $\mathbf{u}$, kde $n = a_1 + a_2 + \cdots + a_N + m$, získáme postupnou aplikací morfismů $D^{a_1}G^{a_2}\cdots D^{a_{N-1}}G^{a_N}$ na m-tý bispeciál v $\mathbf{u}^{(N)}$, kdy v každém kroku připojujeme písmeno b pro D, resp. a pro G. Návratová slova r, s k bispeciálu u_n v $\mathbf{u}$ získáme jako

$r = D^{a_1}G^{a_2}\cdots D^{a_{N-1}}G^{a_N}(\mathtt{b})$, $s = D^{a_1}G^{a_2}\cdots D^{a_{N-1}}G^{a_N}(\mathtt{b}^m\mathtt{a})$. Z toho získáme následující vztahy pro Parikhovy vektory.

$$\vec{V}(r) = \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & 0\\ a_{N-1} & 1\end{pmatrix}\begin{pmatrix}1 & a_N\\ 0 & 1\end{pmatrix}\vec{V}(r'_m) \overset{L.3.2}{=} \begin{pmatrix}p_{N-1} & p_N\\ q_{N-1} & q_N\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}p_N\\ q_N\end{pmatrix}.$$

$$\vec{V}(s) = \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & 0\\ a_{N-1} & 1\end{pmatrix}\begin{pmatrix}1 & a_N\\ 0 & 1\end{pmatrix}\vec{V}(s'_m) \overset{L.3.2}{=} \begin{pmatrix}p_{N-1} & p_N\\ q_{N-1} & q_N\end{pmatrix}\begin{pmatrix}1\\m\end{pmatrix} = \begin{pmatrix}mp_N + p_{N-1}\\ mq_N + q_{N-1}\end{pmatrix}.$$

$$\begin{aligned}\vec{V}(u_n) &= \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\left[\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\left[\cdots\left[\begin{pmatrix}1 & 0\\ a_{N-1} & 1\end{pmatrix}\left[\begin{pmatrix}1 & a_N\\ 0 & 1\end{pmatrix}\vec{V}(b') + \begin{pmatrix}a_N\\0\end{pmatrix}\right] + \begin{pmatrix}0\\a_{N-1}\end{pmatrix}\right] + \dots\right] + \begin{pmatrix}a_2\\0\end{pmatrix}\right] + \begin{pmatrix}0\\a_1\end{pmatrix}\\
&= \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & 0\\ a_{N-1} & 1\end{pmatrix}\begin{pmatrix}1 & a_N\\ 0 & 1\end{pmatrix}\begin{pmatrix}0\\m\end{pmatrix} + \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & 0\\ a_{N-1} & 1\end{pmatrix}\begin{pmatrix}a_N\\0\end{pmatrix}\\
&\quad + \cdots + \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}a_2\\0\end{pmatrix} + \begin{pmatrix}0\\a_1\end{pmatrix}\\
&\overset{L.3.2}{=} \begin{pmatrix}p_{N-1} & p_N\\ q_{N-1} & q_N\end{pmatrix}\begin{pmatrix}0\\m\end{pmatrix} + \begin{pmatrix}p_{N-1} & p_{N-2}\\ q_{N-1} & q_{N-2}\end{pmatrix}\begin{pmatrix}a_N\\0\end{pmatrix} + \cdots + \begin{pmatrix}p_1 & p_0\\ q_1 & q_0\end{pmatrix}\begin{pmatrix}a_2\\0\end{pmatrix} + \begin{pmatrix}0\\a_1\end{pmatrix}\\
&= \begin{pmatrix}mp_N\\ mq_N\end{pmatrix} + a_N\begin{pmatrix}p_{N-1}\\ q_{N-1}\end{pmatrix} + \cdots + a_2\begin{pmatrix}p_1\\ q_1\end{pmatrix} + a_1\begin{pmatrix}p_0\\ q_0\end{pmatrix} \overset{L.3.1}{=} \begin{pmatrix}mp_N\\ mq_N\end{pmatrix} + \begin{pmatrix}p_N\\ q_N\end{pmatrix} + \begin{pmatrix}p_{N-1}\\ q_{N-1}\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix}\\
&= \begin{pmatrix}p_N\\ q_N\end{pmatrix} + \begin{pmatrix}mp_N + p_{N-1}\\ mq_N + q_{N-1}\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix} = \vec{V}(r) + \vec{V}(s) - \begin{pmatrix}1\\1\end{pmatrix}.\end{aligned}$$

N liché: Opět zvolme $\mathbf{u}^{(N)}$ standardní sturmovské slovo s řídící posloupností $\Delta = G^{a_{N+1}}D^{a_{N+2}}G^{a_{N+3}}\cdots$. Pak $\mathbf{u} = D^{a_1}G^{a_2}\cdots G^{a_{N-1}}D^{a_N}(\mathbf{u}^{(N)})$. Nyní získáme m-tý bispeciál b' v $\mathbf{u}^{(N)}$ aplikací m morfismů G na ε a připojováním písmena $\mathtt{a}$. Proto $b' = G^m(\varepsilon)\mathtt{a}^m = \mathtt{a}^m$. Obdobně návratová slova r'_m, resp. s'_m získáme jako $r'_m = G^m(\mathtt{a}) = \mathtt{a}$ a $s'_m = G^m(\mathtt{b}) = \mathtt{a}^m\mathtt{b}$ (nyní je totiž častější písmeno $\mathtt{a}$).

n-tý bispeciál v $\mathbf{u}$ získáme stejným postupem jako v předchozím bodě, stejně tak návratová slova, pouze teď budeme mít lichý počet matic. Pro návratová slova platí $r = D^{a_1}G^{a_2}\cdots G^{a_{N-1}}D^{a_N}(\mathtt{a})$ a $s = D^{a_1}G^{a_2}\cdots G^{a_{N-1}}D^{a_N}(\mathtt{a}^m\mathtt{b})$. Z toho získáme následující vztahy pro Parikhovy vektory.

$$\vec{V}(r) = \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & a_{N-1}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ a_N & 1\end{pmatrix}\vec{V}(r'_m) \overset{L.3.2}{=} \begin{pmatrix}p_N & p_{N-1}\\ q_N & q_{N-1}\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}p_N\\ q_N\end{pmatrix}.$$

$$\vec{V}(s) = \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & a_{N-1}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ a_N & 1\end{pmatrix}\vec{V}(s'_m) \overset{L.3.2}{=} \begin{pmatrix}p_N & p_{N-1}\\ q_N & q_{N-1}\end{pmatrix}\begin{pmatrix}m\\1\end{pmatrix} = \begin{pmatrix}mp_N + p_{N-1}\\ mq_N + q_{N-1}\end{pmatrix}.$$

$$\begin{aligned}\vec{V}(u_n) &= \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\left[\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\left[\cdots\left[\begin{pmatrix}1 & a_{N-1}\\ 0 & 1\end{pmatrix}\left[\begin{pmatrix}1 & 0\\ a_N & 1\end{pmatrix}\vec{V}(b') + \begin{pmatrix}0\\a_N\end{pmatrix}\right] + \begin{pmatrix}a_{N-1}\\0\end{pmatrix}\right] + \dots\right] + \begin{pmatrix}a_2\\0\end{pmatrix}\right] + \begin{pmatrix}0\\a_1\end{pmatrix}\\
&= \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & a_{N-1}\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ a_N & 1\end{pmatrix}\begin{pmatrix}m\\0\end{pmatrix} + \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}1 & a_2\\ 0 & 1\end{pmatrix}\cdots\begin{pmatrix}1 & a_{N-1}\\ a_{N-1} & 1\end{pmatrix}\begin{pmatrix}0\\a_N\end{pmatrix}\\
&\quad + \cdots + \begin{pmatrix}1 & 0\\ a_1 & 1\end{pmatrix}\begin{pmatrix}a_2\\0\end{pmatrix} + \begin{pmatrix}0\\a_1\end{pmatrix}\\
&\overset{L.3.2}{=} \begin{pmatrix}p_N & p_{N-1}\\ q_N & q_{N-1}\end{pmatrix}\begin{pmatrix}m\\0\end{pmatrix} + \begin{pmatrix}p_{N-2} & p_{N-1}\\ q_{N-2} & q_{N-1}\end{pmatrix}\begin{pmatrix}0\\a_N\end{pmatrix} + \cdots + \begin{pmatrix}p_1 & p_0\\ q_1 & q_0\end{pmatrix}\begin{pmatrix}a_2\\0\end{pmatrix} + \begin{pmatrix}0\\a_1\end{pmatrix}\\
&= \begin{pmatrix}mp_N\\ mq_N\end{pmatrix} + a_N\begin{pmatrix}p_{N-1}\\ q_{N-1}\end{pmatrix} + \cdots + a_2\begin{pmatrix}p_1\\ q_1\end{pmatrix} + a_1\begin{pmatrix}p_0\\ q_0\end{pmatrix} \overset{L.3.1}{=} \begin{pmatrix}mp_N\\ mq_N\end{pmatrix} + \begin{pmatrix}p_N\\ q_N\end{pmatrix} + \begin{pmatrix}p_{N-1}\\ q_{N-1}\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix}\\
&= \begin{pmatrix}p_N\\ q_N\end{pmatrix} + \begin{pmatrix}mp_N + p_{N-1}\\ mq_N + q_{N-1}\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix} = \vec{V}(r) + \vec{V}(s) - \begin{pmatrix}1\\1\end{pmatrix}.\end{aligned}$$

Vztahy tedy platí pro všechna $N \in \mathbb{N}_0$ a $m \in \{0, 1, \ldots, a_{N+1} - 1\}$, platí proto pro všechna $n \in \mathbb{N}$. □

S pomocí těchto vzorců již dokážeme odvodit přesný vztah pro výpočet (asymptotického) kritického exponentu sturmovských slov [13].

Věta 3.6. *Nechť* $\mathbf{u}$ *je standardní sturmovské slovo s příslušným řetězovým zlomkem* $\theta = [0, a_1, a_2, \ldots]$. *Pak*

$$\begin{aligned} \mathrm{E}(\mathbf{u}) &= 1 + \sup\left\{a_{N+1} + 1 + \frac{Q_{N-1} - 2}{Q_N} : N \in \mathbb{N}_0\right\}, \\ \mathrm{E}^*(\mathbf{u}) &= 1 + \limsup_{N \to +\infty}\left(a_{N+1} + 1 + \frac{Q_{N-1}}{Q_N}\right), \end{aligned}$$

kde $Q_N = p_N + q_N$.

Důkaz. Nechť $u_n \in \mathcal{L}(\mathbf{u})$ je bispeciál s příslušnou dvojicí (N, m) a $v_n \in \mathcal{L}(\mathbf{u})$ je jeho nejkratší návratové slovo. S využitím odvozených vzorců z věty 3.5 získáme

$$\begin{aligned} \frac{|u_n|}{|v_n|} &= \frac{Q_N + Q_{N-1} - 2}{Q_{N-1}} = \frac{a_N Q_{N-1} + Q_{N-2} + Q_{N-1} - 2}{Q_{N-1}} && \text{pro } m = 0, \\ \frac{|u_n|}{|v_n|} &= \frac{Q_N + mQ_N + Q_{N-1} - 2}{Q_N} = m + 1 + \frac{Q_{N-1} - 2}{Q_N} && \text{pro } m \in \{1, \ldots, a_{N+1} - 1\}, \end{aligned}$$

kde jsme ve druhé rovnosti využili rekurentního vztahu pro člen posloupnosti Q_N pro případ $N \geq 1$. Navíc pro bispeciál u s příslušnou dvojicí $(N + 1, 0)$ a jeho nejkratší návratové slovo v platí

$$\frac{|u|}{|v|} = \frac{a_{N+1} Q_N + Q_{N-1} + Q_N - 2}{Q_N} = a_{N+1} + 1 + \frac{Q_{N-1} - 2}{Q_N} > m + 1 + \frac{Q_{N-1} - 2}{Q_N},$$

kde $m \in \{1, \ldots, a_{N+1} - 1\}$. Výsledná nerovnost zaručuje, že sekundární bispeciály při hledání (asymptotického) kritického exponentu nemusíme uvažovat. Dále si uvědomíme, že $\lim_{N \to +\infty} \frac{2}{Q_N} = 0$, proto ve vzorci pro asymptotický kritický exponent můžeme vypustit člen -2 v čitateli. Celkem získáme

$$\begin{aligned} \mathrm{E}(\mathbf{u}) &= 1 + \sup\left\{a_{N+1} + 1 + \frac{Q_{N-1} - 2}{Q_N} : N \in \mathbb{N}_0\right\}, \\ \mathrm{E}^*(\mathbf{u}) &= 1 + \limsup_{N \to +\infty}\left(a_{N+1} + 1 + \frac{Q_{N-1}}{Q_N}\right). \end{aligned}$$ □

Můžeme vidět, že pokud jsou koeficienty a_N řetězového zlomku θ omezené, bude (asymptotický) kritický exponent konečný. V opačném případě vyjde $+\infty$.

Lemma 3.3. *Nechť* $\mathbf{u}$ *je sturmovské slovo. Pak existují hustoty písmen* $\rho_{\mathtt{a}}$ *a* $\rho_{\mathtt{b}}$.

Důkaz. Bez újmy na obecnosti ukážeme, že existuje hustota písmene $\mathtt{a}$.

- Napřed ukážeme, že pro všechny neprázdné faktory $u, v \in \mathcal{L}(\mathbf{u})$ platí

$$\left|\frac{|u|_{\mathtt{a}}}{|u|} - \frac{|v|_{\mathtt{a}}}{|v|}\right| < \frac{1}{|u|} + \frac{1}{|v|}.$$

  - Pokud $|u| = |v|$, pak

$$\left|\frac{|u|_{\mathtt{a}}}{|u|} - \frac{|v|_{\mathtt{a}}}{|v|}\right| = \frac{\left||u|_{\mathtt{a}} - |v|_{\mathtt{a}}\right|}{|u|} \leq \frac{1}{|u|} < \frac{2}{|u|},$$

    kde jsme v první nerovnosti využili balancovanosti sturmovského slova $\mathbf{u}$.

– Bez újmy na obecnosti předpokládejme $|u| > |v|$, ozn. $u = zt$, kde $|z| = |v|$. Využijeme matematickou indukci na součet délek slov.

  * Indukční předpoklady

$$\left|\frac{|t|_{\mathtt{a}}}{|t|} - \frac{|v|_{\mathtt{a}}}{|v|}\right| < \frac{1}{|t|} + \frac{1}{|v|} \text{ a } \left|\frac{|z|_{\mathtt{a}}}{|z|} - \frac{|v|_{\mathtt{a}}}{|v|}\right| \leq \frac{1}{|v|}.$$

  * Pak

$$\begin{aligned}\left|\frac{|u|_{\mathtt{a}}}{|u|} - \frac{|v|_{\mathtt{a}}}{|v|}\right| &= \left|\frac{|z|}{|u|}\frac{|z|_{\mathtt{a}}}{|z|} + \frac{|t|}{|u|}\frac{|t|_{\mathtt{a}}}{|t|} - \frac{|z|+|t|}{|u|}\frac{|v|_{\mathtt{a}}}{|v|}\right| \\ &= \left|\frac{|z|}{|u|}\left(\frac{|z|_{\mathtt{a}}}{|z|} - \frac{|v|_{\mathtt{a}}}{|v|}\right) + \frac{|t|}{|u|}\left(\frac{|t|_{\mathtt{a}}}{|t|} - \frac{|v|_{\mathtt{a}}}{|v|}\right)\right| \\ &< \frac{|z|}{|u|}\frac{1}{|v|} + \frac{|t|}{|u|}\left(\frac{1}{|t|} + \frac{1}{|v|}\right) \\ &= \frac{1}{|v|} + \frac{1}{|u|}.\end{aligned}$$

Nerovnost je proto splněna.

- Nechť $(u_n)_{n\in\mathbb{N}}$ je posloupnost prefixů slova $\mathbf{u}$ délky n. Pak platí

$$\left|\frac{|u_n|_{\mathtt{a}}}{|u_n|} - \frac{|u_m|_{\mathtt{a}}}{|u_m|}\right| < \frac{1}{|u_n|} + \frac{1}{|u_m|} = \frac{1}{n} + \frac{1}{m}.$$

Pro libovolně malé δ dokážeme najít $n_0 \in \mathbb{N}$ tak, aby $\frac{2}{n_0} < \delta$. Z nerovnosti výše pak vyplývá, že pro $n, m \geq n_0$ je $\frac{1}{n} + \frac{1}{m} \leq \frac{2}{n_0} < \delta$. Posloupnost $\left(\frac{|u_n|_{\mathtt{a}}}{|u_n|}\right)_{n\in\mathbb{N}}$ je tedy cauchyovská a má limitu

$$\rho_{\mathtt{a}} = \lim_{n\to+\infty}\frac{|u_n|_{\mathtt{a}}}{|u_n|}.$$ □

Význam θ pro výpočet nám ukáže následující věta. První tvrzení věty je převzato z [14], druhé z [5].

Věta 3.7. *Nechť* $\mathtt{b}$ *je častější písmeno standardního sturmovského slova* $\mathbf{u}$, *ke kterému je přiřazeno iracionální číslo* θ. *Pak*

1. $\theta = \frac{\rho_{\mathtt{a}}}{\rho_{\mathtt{b}}}$, *kde* $\rho_{\mathtt{a}}$ *je hustota písmena* $\mathtt{a}$ *a* $\rho_{\mathtt{b}}$ *je hustota písmena* $\mathtt{b}$ *v* $\mathbf{u}$.

2. $\mathbf{u}$ *obsahuje faktor* u *takový, že* $|u|_{\mathtt{a}} = \ell$ *a* $|u|_{\mathtt{b}} = k$, *právě tehdy, když* $k, \ell \in \mathbb{N}_0$ *a platí*

$$(k-1)\theta - 1 < \ell < (k+1)\theta + 1.$$

Důkaz. Již jsme ukázali, že existují hustoty písmen $\rho_{\mathtt{b}} = \lim_{n\to\infty}\frac{|u_0u_1\ldots u_{n-1}|_{\mathtt{b}}}{n}$. Limita existuje, její hodnota se tedy musí rovnat hodnotě limity kterékoliv vybrané podposloupnosti. Uvažujme proto posloupnost prefixových návratových slov k n-tému bispeciálu, protože pro ně známe přesný tvar Parikhových vektorů. Pak

$$\frac{\rho_{\mathtt{a}}}{\rho_{\mathtt{b}}} = \lim_{n\to\infty}\frac{|r_n|_{\mathtt{a}}}{|r_n|_{\mathtt{b}}} = \lim_{n\to\infty}\frac{p_n}{q_n} = \theta.$$

Dále víme, že $\rho_{\mathtt{a}} + \rho_{\mathtt{b}} = 1$, proto $\rho_{\mathtt{b}} = \frac{1}{1+\theta}$, a tedy $\rho_{\mathtt{b}}$ je iracionální.

Pro každou délku $n > 0$ existují faktory $u^{(1)}$, $u^{(2)}$ délky n takové, že $|u^{(1)}|_{\mathtt{b}} > n\rho_{\mathtt{b}} > |u^{(2)}|_{\mathtt{b}}$. Z balancovanosti $|u^{(1)}|_{\mathtt{b}} - |u^{(2)}|_{\mathtt{b}} \leq 1$. A proto $|u^{(1)}|_{\mathtt{b}} = \lceil n\rho_{\mathtt{b}} \rceil$, $|u^{(2)}|_{\mathtt{b}} = \lfloor n\rho_{\mathtt{b}} \rfloor$. Navíc každý další faktor splňující $|v| = n$ splňuje $|v|_{\mathtt{b}} = |u^{(1)}|_{\mathtt{b}}$, nebo $|v|_{\mathtt{b}} = |u^{(2)}|_{\mathtt{b}}$.

Proto $\rho_{\mathtt{b}} n - 1 < |u|_{\mathtt{b}} < \rho_{\mathtt{b}} n + 1$.

$$\begin{aligned}
\rho_{\mathtt{b}}(k+\ell) - 1 &< k < \rho_{\mathtt{b}}(k+\ell) + 1 \\
\frac{1}{1+\theta}(k+\ell) - 1 &< k < \frac{1}{1+\theta}(k+\ell) + 1 \\
(k+\ell) - (1+\theta) &< k(1+\theta) < (k+\ell) + (1+\theta) \\
\theta(k-1) - 1 &< \ell < \theta(k+1) + 1
\end{aligned}$$

□

Při výpočtech budeme ještě potřebovat vztah pro θ derivovaného slova.

Věta 3.8. *[8] Nechť* $\mathbf{u}$ *je standardní sturmovské slovo s* $\theta = [0, a_1, a_2, a_3, a_4, \ldots]$. *Dále nechť* u_n *je dle délky v pořadí n-tý bispeciál slova* $\mathbf{u}$ *a přísluší mu dvojice* (N, m). *Pak* θ' *příslušná derivovanému slovu* $\mathrm{d}_{\mathbf{u}}(u_n)$ *splňuje* $\theta' = [0, (a_{N+1} - m), a_{N+2}, a_{N+3}, \ldots]$.

Důkaz. Již víme, že n-tý bispeciál můžeme generovat z prázdného slova aplikací prvních n morfismů z řídící posloupnosti a připojením písmena a, resp. b v každém kroku. Obdobně můžeme generovat návratová slova k n-tému bispeciálu prostou aplikací morfismů na návratová slova k prázdnému slovu, což jsou jednotlivá písmenka. Protože morfismus nemění pořadí písmen a faktorů z nich vytvořených, pak návratová slova k n-tému bispeciálu v $\mathbf{u}$ budou ve stejném pořadí, jako písmena a a b v $\mathbf{u}^{(n)}$, kde častější písmeno odpovídá častějšímu návratovému slovu. Proto $\mathbf{u}^{(n)}$ odpovídá $\mathrm{d}_{\mathbf{u}}(u_n)$, pouze nad jinak značenou abecedou.

Nakonec si uvědomíme, že řetězový zlomek θ odpovídá mocninám morfismů G a D, které se vyskytují v řídící posloupnosti $\Delta = G^{a_1} D^{a_2} \cdots$ nebo $\Delta = D^{a_1} G^{a_2} \cdots$. Chceme-li pak získat řídící posloupnost příslušnou k $\mathbf{u}^{(n)}$, stačí z Δ odebrat právě prvních n morfismů. Takže pro $n = a_0 + a_1 + \cdots + a_N + m$, kde $a_0 = 0$ a $0 \leq m < a_{N+1}$, získáme $\Delta' = G^{a_{N+1}-m} D^{a_{N+2}} \cdots$ nebo $\Delta' = D^{a_{N+1}-m} G^{a_{N+2}} \cdots$ podle toho, na který morfismus začínala Δ a zda je N sudé nebo liché. Proto řídící posloupnost příslušná $\mathbf{u}^{(n)}$ je Δ' a příslušná $\theta' = [0, (a_{N+1} - m), a_{N+2}, a_{N+3}, \ldots]$. □

Při výpočtu (asymptotického) kritického exponentu budeme pracovat s posléze periodickým řetězovým zlomkem θ.

Věta 3.9. *(Lagrange) Nechť* $\xi \in \mathbb{R} \setminus \mathbb{Q}$. *Pak řetězový zlomek* ξ *je posléze periodický, právě když* ξ *je algebraické číslo stupně 2.*

Proto si zavedeme potřebné značení pro θ' derivované posloupnosti $d_{\mathbf{u}}(u_n)$.

Definice 3.5. *Nechť* θ *má posléze periodický řetězový zlomek ve tvaru*

$$\theta = [0, d_1, d_2, \ldots, d_h, (z_0, z_1, \ldots, z_{M-1})^\omega]$$

a nechť u_n *je n-tý bispeciál v* $\mathbf{u}$ *s příslušnou dvojicí* (N, m). *Pak řetězový zlomek* θ' *příslušný k jeho derivovanému slovu* $\mathrm{d}_{\mathbf{u}}(u_n)$ *označíme*

- $\hat{\theta}_{(N,m)} = [0, d_{N+1} - m, d_{N+2}, \ldots, d_h, (z_0, \ldots, z_{M-1})^\omega]$, *pokud* $N < h$,
- $\theta_{(i,m)} = [0, z_i - m, (z_{i+1}, \ldots, z_{M-1}, z_0, \ldots, z_i)^\omega]$ *pro* $N \geq h$ *a* $i = (N - h) \mod M$.

Příklad 3.1. Vrátíme se opět k Fibonacciho slovu, které je pevným bodem morfismu

$$\begin{aligned}\varphi : \mathtt{a} &\to \mathtt{ab},\\ \mathtt{b} &\to \mathtt{a}.\end{aligned}$$

Morfismus φ na první pohled neodpovídá ani jednomu z morfismů G a D, ale φ^2 již z těchto morfismů složit dokážeme

$$\begin{aligned}\varphi^2 : \mathtt{a} &\to \mathtt{aba}, \qquad & (GD) : \mathtt{a} &\to G(\mathtt{ba}) \to \mathtt{aba},\\ \mathtt{b} &\to \mathtt{ab}. & \mathtt{b} &\to G(\mathtt{b}) \to \mathtt{ab}.\end{aligned}$$

Tudíž $\Delta = (GD)^\omega$, $\theta = [0,(1)^\omega]$, $p_n = p_{n-1} + p_{n-2}$ s počátečními podmínkami $p_0 = 0, p_1 = 1$ a $q_n = q_{n-1} + q_{n-2}$ s počátečními podmínkami $q_0 = 1, q_1 = 1$. Tudíž $p_n = F_n$ a $q_n = F_{n+1}$, kde F_n jsou členy Fibonacciho posloupnosti.

Protože $0 \le m < a_n = 1$ pro všechny $n \in \mathbb{N}$, je $m = 0$ pro všechny bispeciály. Pak n-tému bispeciálu přísluší dvojice $(N,m) = (n,0)$ a pro jeho návratová slova platí:

$$\begin{aligned}\vec{V}(r) &= \begin{pmatrix} q_N \\ p_N \end{pmatrix} = \begin{pmatrix} F_{N+1} \\ F_N \end{pmatrix},\\ \vec{V}(s) &= \begin{pmatrix} q_{N-1} \\ p_{N-1} \end{pmatrix} = \begin{pmatrix} F_N \\ F_{N-1} \end{pmatrix},\\ \vec{V}(u_N) &= \begin{pmatrix} F_{N+1} + F_N - 1 \\ F_N + F_{N-1} - 1 \end{pmatrix} = \begin{pmatrix} F_{N+2} - 1 \\ F_{N+1} - 1 \end{pmatrix}.\end{aligned}$$

Například třetí bispeciál získáme jako $u_3 = G(D(G(\varepsilon)\mathtt{a})\mathtt{b})\mathtt{a} = G(D(\mathtt{a})\mathtt{b})\mathtt{a} = G(\mathtt{bab})\mathtt{a} = \mathtt{abaaba}$ a jeho délka je opravdu $F_5 + F_4 - 2 = 5 + 3 - 2 = 6$.

Kapitola 4

Balancovaná slova

Již víme, že aperiodická balancovaná slova nad binární abecedou splývají s těmi sturmovskými. Pomocí sturmovských slov lze dokonce generovat balancovaná slova i nad vícepísmennými abecedami, k tomu budeme potřebovat následující definici.

Definice 4.1. *Nekonečné slovo* $\mathbf{y}$ *nad abecedou* $\mathcal{A}$ *nazveme* slovem s konstantními mezerami*, pokud pro každé písmeno* $c \in \mathcal{A}$ *existuje konstanta* $K > 0$ *taková, že vzdálenost mezi každými dvěma po sobě jdoucími výskyty* c *ve slově* $\mathbf{y}$ *je* K*.*

Takové slovo je vždy periodické. Minimální periodu $\mathbf{y}$ značíme $\mathrm{Per}(\mathbf{y})$.

Máme-li y faktor slova s konstantními mezerami $\mathbf{y}$, pak jako $\mathrm{gap}_{\mathbf{y}}(y)$ značíme vzdálenost mezi dvěma po sobě jdoucími výskyty y v $\mathbf{y}$. Je to také délka návratového slova k y. Triviálně $\mathrm{gap}_{\mathbf{y}}(\varepsilon) = 1$. Dále pro každé přirozené číslo $n \in \mathbb{N}_0$ definujeme množinu všech mezer, které najdeme mezi faktory délky n, $\mathrm{gap}(\mathbf{y}, n) = \{i \in \mathbb{N} : y \in \mathcal{L}(\mathbf{y}) \wedge |y| = n \wedge \mathrm{gap}_{\mathbf{y}}(y) = i\}$. Již víme, že v periodickém slově $\mathbf{y}$ je délka nejdelšího bispeciálu omezená, jeho délku si označíme $\beta(\mathbf{y})$. Můžeme vidět, že $\beta(\mathbf{y}) < \mathrm{Per}(\mathbf{y})$.

Již máme všechny potřebné definice a můžeme si ukázat, jak lze balancovaná slova generovat.

Definice 4.2. *Nechť* $\mathbf{y}$*, resp.* $\mathbf{y}'$ *je slovo s konstantními mezerami nad abecedou* $\mathcal{A}$*, resp.* $\mathcal{B}$*, kde* $\mathcal{A}$ *a* $\mathcal{B}$ *jsou disjunktní. Označme* $\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$ *slovo, které získáme ze slov* $\mathbf{u}$*,* $\mathbf{y}$ *a* $\mathbf{y}'$ *následujícím postupem.*

Postupně čteme písmena ve slově $\mathbf{u}$ *a nahrazujeme:*

$$\begin{aligned} \mathtt{a} &\rightarrow \textit{další písmeno v pořadí ze slova } \mathbf{y}, \\ \mathtt{b} &\rightarrow \textit{další písmeno v pořadí ze slova } \mathbf{y}'. \end{aligned}$$

Říkáme, že $\mathbf{v}$ *je* obarvení $\mathbf{u}$ *slovy* $\mathbf{y}$ *a* $\mathbf{y}'$*. Dále užíváme projekci* $\pi : \mathcal{L}(\mathbf{v}) \rightarrow \mathcal{L}(\mathbf{u})$*, která funguje jako morfismus a přiřazuje*

$$\begin{aligned} a &\mapsto \mathtt{a} \textit{ pro všechna } a \in \mathcal{A}, \\ b &\mapsto \mathtt{b} \textit{ pro všechna } b \in \mathcal{B}. \end{aligned}$$

Řekneme, že $\mathbf{u} = \pi(\mathbf{v})$*, resp.* $u = \pi(v)$*, je* projekce $\mathbf{v}$*, resp.* v*.*

Věta 4.1. [11] *Rekurentní aperiodické slovo* $\mathbf{v}$ *je balancované právě tehdy, když existuje sturmovské slovo* $\mathbf{u}$ *a slova s konstantními mezerami* $\mathbf{y}$ *a* $\mathbf{y}'$ *nad disjunktními abecedami tak, že* $\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$*.*

Tento způsob generování rekurentních aperiodických slov budeme hojně využívat. Například mnoho vlastností balancovaných slov se dá odvodit z analogických vlastností slov sturmovských.

Lemma 4.1. *Nechť* $\mathbf{u}$ *je sturmovské slovo,* $\mathbf{y}$ *a* $\mathbf{y}'$ *dvě slova s konstantními mezerami nad disjunktními abecedami a* $\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$*. Dále nechť* $u \in \mathcal{L}(\mathbf{u})$*. Pak pro všechna* $i, j \in \mathbb{N}_0$ *je slovo* w *vzniklé z* u *obarvením slovy* $\mathbf{y}^{(i)} = y_i y_{i+1} y_{i+1} \dots$ *a* $\mathbf{y}'^{(j)} = y'_j y'_{j+1} y'_{j+2} \dots$ *faktorem slova* $\mathbf{v}$*.*

Důkaz. Díky periodicitě slov $\mathbf{y}$ a $\mathbf{y}'$ stačí uvažovat $i < \mathrm{Per}(\mathbf{y})$, $j < \mathrm{Per}(\mathbf{y}')$. S využitím WDO pro sturmovské slovo $\mathbf{u}$ nalezneme prefix p slova $\mathbf{u}$ takový, že za ním následuje u a navíc

$$\vec{V}(p) \equiv \begin{pmatrix} i \mod \mathrm{Per}(\mathbf{y}) \cdot \mathrm{Per}(\mathbf{y}') \\ j \mod \mathrm{Per}(\mathbf{y}) \cdot \mathrm{Per}(\mathbf{y}') \end{pmatrix}.$$

Pak po obarvení prefixu p slova $\mathbf{u}$ budeme ve slovech $\mathbf{y}$ a $\mathbf{y}'$ právě na pozicích $y_i y_{i+1} \dots$ a $y'_j y'_{j+1} \dots$ a obarvení tohoto výskytu faktoru u je hledaný faktor w. □

Věta 4.2. *Pokud máme dvě sturmovská slova* $\mathbf{u}, \mathbf{u}'$*, která mají stejný jazyk, a dvě slova s konstantními mezerami* $\mathbf{y}$ *a* $\mathbf{y}'$ *nad disjunktními abecedami, pak i* $\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$ *a* $\mathbf{v}' = \mathrm{colour}(\mathbf{u}', \mathbf{y}, \mathbf{y}')$ *mají stejný jazyk.*

Důkaz. Předchozí lemma zaručuje, že jazyk slova $\mathbf{v}$ nezávisí na přesném výskytu faktorů ve slově $\mathbf{u}$, pouze na jeho jazyku. Jazyk slova $\mathbf{v}$ je totiž tvořen všemi možnými obarveními faktorů $u \in \mathcal{L}(\mathbf{u})$ slovy $y_i y_{i+1} y_{i+1} \dots$ a $y'_j y'_{j+1} y'_{j+2} \dots$ pro $i, j \in \mathbb{N}_0$. □

Navíc z balancovanosti lze ukázat (v důkazu lemmatu 3.3 jsme využili pouze balancovanosti sturmovských slov), že pro každé písmeno z abecedy existuje jeho hustota v rekurentním balancovaném slově. Pro periodická slova hustota všech písmen existuje vždy, pro ta rekurentní aperiodická lze využít vlastnosti barvení. Pak hustota písmena $a \in \mathcal{A}$ ve slově $\mathbf{v}$ závisí na hustotě $\rho_{\mathtt{a}}$ písmena $\mathtt{a}$ ve sturmovském slově $\mathbf{u}$ a mezeře $\mathrm{gap}_{\mathbf{y}}(a)$ písmena a ve slově s konstantními mezerami $\mathbf{y}$ předpisem $\rho_a = \frac{\rho_{\mathtt{a}}}{\mathrm{gap}_{\mathbf{y}}(a)}$.

Příklad 4.1. Slovo $(01)^\omega$ je balancované, ale například $\mathbf{v} = (0110)^\omega$ již ne, protože $00, 11 \in \mathcal{L}(\mathbf{v})$ mají stejnou délku, ale počet jedniček i nul se liší o 2.

Slovo $(010203)^\omega$ je slovo s konstantními mezerami nad 4-písmennou abecedou. Pro 0 je mezera 2, pro ostatní písmena 6. Faktor 0 je tudíž nejdelším bispeciálem a $\beta((010203)^\omega) = 1$.

Slovo $(101)^\omega$ není slovo s konstantními mezerami.

Máme-li standardní sturmovské slovo $\mathbf{u}$ s řídící posloupností $\Delta = (GD^2)^\omega$, pak jeho prefix (a zároveň šestý bispeciál) je

$$\begin{aligned} G(D(D(G(D(D(\varepsilon)\mathtt{b})\mathtt{b})\mathtt{a})\mathtt{b})\mathtt{b})\mathtt{a} &= G(D(D(G(\mathtt{bb})\mathtt{a})\mathtt{b})\mathtt{b})\mathtt{a} = G(D(D(\mathtt{ababa})\mathtt{b})\mathtt{b})\mathtt{a} \\ &= G(\mathtt{bbabbbabbbabb})\mathtt{a} = \mathtt{ababaababababaababababaababa}. \end{aligned}$$

Nyní prefix obarvíme slovy s konstantními mezerami $\mathbf{y} = (0102)^\omega$ a $\mathbf{y}' = (34)^\omega$. Získáme tak prefix balancovaného slova $\mathbf{v} = \mathrm{colour}(\mathbf{u}, (0102)^\omega, (34)^\omega)$ nad pětipísmennou abecedou $\mathcal{A} \dot\cup \mathcal{B} = \{0, 1, 2, 3, 4\}$.

u =	a	b	a	b	a	a	b	a	b	a	b	a	a	b	a	b	a	b	a	a	b	a	b	a	…
obarvení a:	0		1		0	2		0		1		0	2		0		1		0	2		0		1	…
obarvení b:		3		4			3		4		3			4		3		4			3		4		…
v =	0	3	1	4	0	2	3	0	4	1	3	0	2	4	0	3	1	4	0	2	3	0	4	1	…

Můžeme vidět, že $\pi(031) = \mathtt{aba}$, $\pi(3140) = \mathtt{baba}$ a například $\pi^{-1}(\mathtt{ab}) = \{03, 14, 23, 04, 13, 24\}$.

4.1 Bispeciály balancovaného slova

Již známe vztahy pro délky bispeciálů a jejich návratových slov ve sturmovských slovech, nyní potřebujeme prozkoumat stejné vztahy pro slova balancovaná. Většina poznatků ze sturmovských slov však půjde využít, jak nám ukáže následující věta.

Věta 4.3. *Nechť* $\mathbf{u}$ *je standardní sturmovské slovo,* $\mathbf{y}$ *a* $\mathbf{y}'$ *jsou slova s konstantními mezerami nad disjunktními abecedami a nechť* $\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$*. Mějme faktory* $w \in \mathcal{L}(\mathbf{v}), u \in \mathcal{L}(\mathbf{u})$ *takové, že* $\pi(w) = u$*.*

1. *Pokud* u *je bispeciál v* $\mathbf{u}$*, pak* w *je bispeciál ve* $\mathbf{v}$*.*

2. *Pokud* w *je bispeciál ve* $\mathbf{v}$ *a* u *je navíc dostatečně dlouhé ve smyslu* $|u|_{\mathsf{a}} > \beta(\mathbf{y})$ *a* $|u|_{\mathsf{b}} > \beta(\mathbf{y}')$*, pak* u *je bispeciál v* $\mathbf{u}$*.*

Důkaz. 1. Pokud u je bispeciál v $\mathbf{u}$, pak $\mathsf{a}u, \mathsf{b}u, u\mathsf{a}, u\mathsf{b} \in \mathcal{L}(\mathbf{u})$. Při obarvování faktorů $\mathsf{a}u$, resp. $\mathsf{b}u$ budeme brát první písmeno z abecedy $\mathcal{A}$, resp. $\mathcal{B}$. Z disjunktnosti abeced a vlastnosti WDO pak vyplývá, že existují písmena a, b tak, že $a \in \mathcal{A}, b \in \mathcal{B}$ a $a \neq b$ a zároveň $aw, bw \in \mathcal{L}(\mathbf{v})$. Obdobně to platí i pro pravé prodloužení, a tudíž w je bispeciálem ve $\mathbf{v}$.

2. Napřed si uvědomíme skutečnost, že nekonečné slovo $\mathbf{y}$, resp $\mathbf{y}'$ je posléze periodické. Proto pokud sledujeme faktory o délce větší, než je $\beta(\mathbf{y})$, resp. $\beta(\mathbf{y}')$, pak tyto faktory mají právě jedno možné prodloužení doprava a právě jedno možné prodloužení doleva.

   Zbytek důkazu provedeme sporem, budeme tedy předpokládat, že u je dostatečně dlouhé, ale není bispeciálem v $\mathbf{u}$. Pak barvíme-li napřed písmena a faktoru u slovem y, přečteme faktor délky $|u|_a > \beta(\mathbf{y})$. Z předchozího pozorování ale vyplývá, že následuje-li za u v $\mathbf{u}$ opět písmeno a, pak existuje právě jedno písmeno a z abecedy $\mathcal{A}$ takové, že $wa \in \mathcal{L}(\mathbf{v})$. Pak ale w není bispeciálem ve $\mathbf{v}$. Stejná úvaha lze udělat i pro barvení písmen b, popř. pro prodloužení doleva. □

Kromě bispeciálů v $\mathbf{u}$ tedy budeme muset ještě vyšetřit krátké faktory, které ačkoliv nejsou v $\mathbf{u}$ bispeciály, mohou se do balancovaného slova $\mathbf{v}$ jako bispeciály přepsat.

K bispeciálům $w \in \mathcal{L}(\mathbf{v})$ takovým, že $\pi(w)$ je bispeciál v $\mathbf{u}$, přiřadíme stejnou dvojici (N, m), která přísluší k $\pi(w)$ ve slově $\mathbf{u}$.

Příklad 4.2. Vezmeme si balancované slovo z příkladu 4.1, $\mathbf{v} = 031402\underline{30}41\underline{30}240314023041\ldots$, které vzniklo obarvením $\mathbf{u} = \mathsf{ababaababababaabababaababa}\ldots$ slovy $\mathbf{y} = (0102)^{\omega}$ a $\mathbf{y}' = (34)^{\omega}$.

Zaměříme-li se na faktor 30, porovnáním podtržených výskytů můžeme vidět, že je to bispeciál ve $\mathbf{v}$. Projekce tohoto bispeciálu $\pi(30) = \mathsf{ba}$ však sama bispeciálem není, protože před písmenem b je v $\mathbf{u}$ vždy a. Prozkoumáme-li slova $\mathbf{y}$ a $\mathbf{y}'$, zjistíme, že $\beta(\mathbf{y}) = 1 = |\mathsf{ba}|_{\mathsf{a}}$, a tudíž nebyly splněny předpoklady 2. tvrzení předchozí věty.

4.2 Nejkratší návratová slova v balancovaných slovech

V aperiodických balancovaných slovech může mít jeden faktor několik návratových slov, nás však budou zajímat pouze ta nejkratší. K odvození délky nejkratšího návratového slova k faktorům balancovaných slov využijeme následující dvě lemmata.

Lemma 4.2. *Nechť* $\mathbf{u}$ *je standardní sturmovské slovo,* $\mathbf{y}$ *a* $\mathbf{y}'$ *dvě slova s konstantními mezerami nad disjunktními abecedami a* $\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$*. Nechť* z *je faktor* $\mathbf{u}$ *a* i, j *jsou jeho výskyty ve slově* $\mathbf{u}$*.*

Označme w, resp. w', faktory délky $|z|$, které se vyskytují ve $\mathbf{v}$ na pozicích i, resp. j, (tzn. $\pi(w) = \pi(w') = z$). Pokud $w = w'$, pak existují $n \in \mathrm{gap}(\mathbf{y}, |z|_a)$ a $n' \in \mathrm{gap}(\mathbf{y}', |z|_b)$ takové, že

$$\vec{V}(u_i u_{i+1} \dots u_{j-1}) = \begin{pmatrix} 0 \mod n \\ 0 \mod n' \end{pmatrix}.$$

Důkaz. Po přečtení faktoru $v = v_i v_{i+1} \dots v_{j-1}$ ve $\mathbf{v}$ musíme číst opět stejný faktor w. To z definice barvení nastane, když po $n_{\mathsf{a}} := |u_i u_{i+1} \dots u_{j-1}|_{\mathsf{a}}$ písmenech budeme číst ve slově $\mathbf{y}$ stejný faktor délky $|z|_{\mathsf{a}}$ a ve slově $\mathbf{y}'$ po $n_{\mathsf{b}} := |u_i u_{i+1} \dots u_{j-1}|_{\mathsf{b}}$ písmenech přečteme stejný faktor délky $|z|_{\mathsf{b}}$. To ale může nastat pouze v případě, že existuje faktor y slova $\mathbf{y}$ a faktor y' slova $\mathbf{y}'$ tak, že

$$|y| = |z|_{\mathsf{a}} \wedge n_{\mathsf{a}} \equiv 0 \mod \mathrm{gap}_{\mathbf{y}}(y) \wedge |y'| = |z|_{\mathsf{b}} \wedge n_{\mathsf{b}} \equiv 0 \mod \mathrm{gap}_{\mathbf{y}}(y').$$

Z toho pak ale vyplývá tvrzení věty, protože stačí položit $n := \mathrm{gap}_{\mathbf{y}}(y) \in \mathrm{gap}(\mathbf{y}, |z|_{\mathsf{a}})$ a $n' := \mathrm{gap}_{\mathbf{y}}(y') \in \mathrm{gap}(\mathbf{y}', |z|_{\mathsf{b}})$. □

Lemma 4.3. *Nechť $\mathbf{u}$, $\mathbf{y}$, $\mathbf{y}'$ a $\mathbf{v}$ jsou jako v lemmatu 4.2. A nechť $z \in \mathcal{L}(\mathbf{u})$, $n \in \mathrm{gap}(\mathbf{y}, |z|_a)$ a $n' \in \mathrm{gap}(\mathbf{y}', |z|_b)$. Pak existují výskyty i, j, $i < j$, faktoru z v $\mathbf{u}$ takové, že*

$$\vec{V}(u_i u_{i+1} \dots u_{j-1}) = \begin{pmatrix} 0 \mod n \\ 0 \mod n' \end{pmatrix}$$

a existuje faktor $w \in \mathcal{L}(\mathbf{v})$, $\pi(w) = z$, tak, že i a j jsou jeho výskyty ve $\mathbf{v}$.

Faktor $u_i u_{i+1} \dots u_{j-1} z$ má z jako prefix i jako sufix. Bude tedy tvořen zřetězením návratových slov k z.

Důkaz. Protože $n \in \mathrm{gap}(\mathbf{y}, |z|_a)$, existuje faktor y slova $\mathbf{y}$ takový, že $|y| = |z|_a$ a $\mathrm{gap}_{\mathbf{y}}(y) = n$. Obdobně pro n' získáme faktor y' slova $\mathbf{y}'$ tak, že $|y'| = |z|_b$ a $\mathrm{gap}_{\mathbf{y}}(y') = n'$. Označme k, resp. ℓ, první výskyty faktorů y v $\mathbf{y}$, resp. y' v $\mathbf{y}'$.

Z WDO získáme, že existují indexy $i, j \in \mathbb{N}, i < j$, takové, že

$$\vec{V}(u_0 u_1 \dots u_{i-1}) = \vec{V}(u_0 u_1 \dots u_{j-1}) = \begin{pmatrix} k \mod n \cdot n' \\ \ell \mod n \cdot n' \end{pmatrix}$$

a po $u_0 u_1 \dots u_{i-1}$ i $u_0 u_1 \dots u_{j-1}$ v $\mathbf{u}$ následuje faktor z. Proto faktor $u_i u_{i+1} \dots u_{j-1}$ má z jako prefix a po něm následuje z.

Dále si uvědomíme, že po přečtení $k + b \cdot n, b \in \mathbb{N}_0$, písmen z $\mathbf{y}$ bude v $\mathbf{y}$ následovat opět faktor y, obdobně pro $\mathbf{y}'$ a faktor y'. Proto se oba výskyty i, j faktoru z v $\mathbf{u}$ budou barvit stejnými faktory z $\mathbf{y}$ a $\mathbf{y}'$. V balancovaném slově $\mathbf{v}$ budeme tedy na pozicích i, j číst stejné faktory délky $|z|$.

Nakonec si stačí uvědomit, že

$$\vec{V}(u_i u_{i+1} \dots u_{j-1}) = \begin{pmatrix} 0 \mod n \cdot n' \\ 0 \mod n \cdot n' \end{pmatrix} \Rightarrow \vec{V}(u_i u_{i+1} \dots u_{j-1}) = \begin{pmatrix} 0 \mod n \\ 0 \mod n' \end{pmatrix}$$ □

Ještě doplníme, že pro $|z|_{\mathsf{a}} > \beta(\mathbf{y})$ a pro $|z|_{\mathsf{b}} > \beta(\mathbf{y}')$ jsou $n = \mathrm{Per}(\mathbf{y})$ a $n' = \mathrm{Per}(\mathbf{y}')$.

Pozorování. Je-li r prefixovým a s neprefixovým návratovým slovem k faktoru $z \in \mathcal{L}(\mathbf{u})$ a i, j výskyty faktoru z v $\mathbf{u}$, pak $\vec{V}(u_i u_{i+1} \dots u_{j-1}) = k\vec{V}(r) + \ell\vec{V}(s)$, kde $\begin{pmatrix} \ell \\ k \end{pmatrix}$ je Parikhův vektor faktoru derivovaného slova $\mathrm{d}_{\mathbf{u}}(z)$.

Tato skutečnost vyplývá z toho, že daný faktor je zřetězením návratových slov k z.

Z lemmat rovnou vyplývá vztah mezi návratovými slovy ve sturmovském slově a v tom odvozeném balancovaném.

Věta 4.4. *Nechť r, resp. s je prefixové, resp. neprefixové návratové slovo k faktoru $u \in \mathcal{L}(\mathbf{u})$. Definujeme množinu $\mathcal{S}(u) = \mathcal{S}_1(u) \cap \mathcal{S}_2(u) \cap \mathcal{S}_3$, kde*

$$\mathcal{S}_1(u) = \left\{ \begin{pmatrix} \ell \\ k \end{pmatrix} : \begin{pmatrix} \ell \\ k \end{pmatrix} \text{ je Parikhův vektor faktoru ve slově } \mathrm{d}_{\mathbf{u}}(u) \right\},$$

$$\mathcal{S}_2(u) = \bigcup_{n \in \mathrm{gap}(\mathbf{y}, |u|_{\mathtt{a}})} \bigcup_{n' \in \mathrm{gap}(\mathbf{y}', |u|_{\mathtt{b}})} \left\{ \begin{pmatrix} \ell \\ k \end{pmatrix} : k\vec{V}(r) + \ell\vec{V}(s) = \begin{pmatrix} 0 \mod n \\ 0 \mod n' \end{pmatrix} \right\},$$

$$\mathcal{S}_3 = \left\{ \begin{pmatrix} \ell \\ k \end{pmatrix} : 1 \le k + \ell \le \mathrm{Per}(\mathbf{y})\mathrm{Per}(\mathbf{y}') \right\}.$$

Pak nejkratší návratové slovo $v \in \mathcal{L}(\mathbf{v})$ k faktoru $w \in \mathcal{L}(\mathbf{v})$, kde $\pi(w) = u$, má délku

$$|v| \ge \min \left\{ k|r| + \ell|s| : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(u) \right\}$$

a existuje $\hat{w} \in \mathcal{L}(\mathbf{v}), \pi(\hat{w}) = u$ tak, že nastává rovnost.

Důkaz. To, že uvažovat můžeme pouze takové $\begin{pmatrix} \ell \\ k \end{pmatrix}$, které náleží do množin $\mathcal{S}_1(u)$ a $\mathcal{S}_2(u)$, vyplývá z lemmat 4.2 a 4.3. Z důkazu lemmatu 4.2 dokonce vyplývá, že takový bispeciál, aby ve vzorci pro nejkratší návratové slovo nastala rovnost, nalezneme (vzniká barvením u pomocí nalezených faktorů y a y'). Jediný bod, který zbývá dokázat, tedy je, že se při zkoumání $\begin{pmatrix} \ell \\ k \end{pmatrix}$ stačí omezit na dvojice z množiny $\mathcal{S}_3$.

Předpokládejme pro spor, že existují k_0, ℓ_0 tak, že

$$k_0|r| + \ell_0|s| = \min \left\{ k|r| + \ell|s| : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}_1(u) \cap \mathcal{S}_2(u) \right\},$$

ale $k_0 + \ell_0 > \mathrm{Per}(\mathbf{y})\mathrm{Per}(\mathbf{y}')$. Pak existuje faktor $d = d_1 d_2 \dots d_p \in \mathcal{L}(d_{\mathbf{u}}(u))$, kde $p = k_0 + \ell_0$, takový, že $\vec{V}(d) = \begin{pmatrix} \ell_0 \\ k_0 \end{pmatrix}$.

Označme dále $\vec{V}(d_1 d_2 \dots d_i) =: \begin{pmatrix} \ell_i \\ k_i \end{pmatrix}$ a $\vec{X}_i := k_i\vec{V}(r) + \ell_i\vec{V}(s)$. Protože počet tříd ekvivalence pro mod $\begin{pmatrix} \mathrm{Per}(\mathbf{y}) \\ \mathrm{Per}(\mathbf{y}') \end{pmatrix}$ je právě $\mathrm{Per}(\mathbf{y})\mathrm{Per}(\mathbf{y}') < p$, musí existovat indexy $1 \le i < j \le p$ tak, že X_i je ekvivalentní X_j. Nechť $\begin{pmatrix} \ell' \\ k' \end{pmatrix} := \vec{V}(d_{i+1} d_2 \dots d_j)$, pak platí $k'\vec{V}(r) + \ell'\vec{V}(s) = X_j - X_i = \begin{pmatrix} 0 \mod \mathrm{Per}(\mathbf{y}) \\ 0 \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix}$, a proto $\begin{pmatrix} \ell' \\ k' \end{pmatrix} \in \mathcal{S}_1(u) \cap \mathcal{S}_2(u)$. Navíc $k' \le k_0$, $\ell' \le \ell_0$ a $k' + \ell' < k_0 + \ell_0$. Výsledná nerovnost je však ve sporu s minimalitou $k_0|r| + \ell_0|s|$.

Díky minimalitě v se nikdy nestane, že by se faktor w vyskytl i někde uprostřed vw, tudíž v je opravdu návratovým slovem. □

Abychom to shrnuli, $\mathcal{S}_1$ zajišťuje, že existuje faktor $\pi(v) = z$ slova $\mathbf{u}$ skládající se z k faktorů r a ℓ faktorů s. $\mathcal{S}_2$ garantuje, že po přečtení u budeme ve slovech $\mathbf{y}$ a $\mathbf{y}'$ na „správných“ pozicích, a $\mathcal{S}_3$ pouze omezuje počet k a ℓ, která musíme procházet.

Příklad 4.3. Opět využijeme vygenerovaný prefix z příkladu 4.1, $\mathbf{v} = 0314\underline{0}23\underline{0}3130240314023041\dots$, který vznikl obarvením $\mathbf{u} = \mathtt{ababaabababaabababaababa}\dots$ slovy $\mathbf{y} = (0102)^\omega$ a $\mathbf{y}' = (34)^\omega$. Příslušný řetězový zlomek je $\theta = [0, (1,2)^\omega]$.

Budeme sledovat faktor 0, kde $u = \pi(0) = \mathtt{a}$. Pak v prefixu $\mathbf{u}$ vidíme, že $r = \mathtt{ab}$, $s = \mathtt{a}$. Protože $|u|_{\mathtt{a}} = 1$, $|u|_{\mathtt{b}} = 0$, potřebujeme najít množiny $\text{gap}(\mathbf{y}, 1)$ a $\text{gap}(\mathbf{y}', 0)$. V $\mathbf{y}$ hledáme mezery příslušné písmenům, mezera písmena 0 je rovná 2, pro písmena $1, 2$ je rovna periodě. Tudíž $\text{gap}(\mathbf{y}, 1) = \{2, 4\}$. Mezera pro prázdné slovo je v jakémkoliv slově s konstantními mezerami vždy 1, tudíž $\text{gap}(\mathbf{y}', 0) = \{1\}$.

Množinu $\mathcal{S}_2(\mathtt{a})$ pak získáme jako

$$\mathcal{S}_2(\mathtt{a}) = \bigcup_{n \in \{2,4\}} \left\{ \begin{pmatrix} \ell \\ k \end{pmatrix} : k \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \ell \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & \mod n \\ 0 & \mod 1 \end{pmatrix} \right\}.$$

Vidíme, že řešením je například $k = 0, \ell = 2$, nebo $k = 2, l = 0$, nebo $k = 1, \ell = 1$.

Písmeno $\mathtt{a}$ je bispeciálem i ve sturmovském slově, přísluší mu dvojice $(N, m) = (1, 0)$ s řetězovým zlomkem $\theta_{(1,0)} = [0, (2, 1)^{\omega}]$. Přepíšeme-li pak $\mathbf{u}$ počínaje prvním výskytem u pomocí návratových slov r a s, získáme prefix

$$\mathrm{d}_{\mathbf{u}}(\mathtt{a}) = \mathtt{rrsrrrsrrrsrr}...$$

Lze vidět, že derivované slovo faktor $\mathtt{ss}$ neobsahuje, ale $\mathtt{rr}$ i $\mathtt{sr}$ již ano (pokud by obsahoval i $\mathtt{ss}$, nastal by spor s balancovaností). Proto $k = 0, l = 2$ není prvkem $\mathcal{S}_1(u)$, ale $k = 2, \ell = 0$ i $k = 1, \ell = 1$ leží v množině $\mathcal{S}_1(u)$. Dále platí $2 + 0 < 4 \cdot 2$ i $1 + 1 < 4 \cdot 2$, tudíž $k = 2, \ell = 0$ i $k = 1, l = 1$ náleží i do $\mathcal{S}_3$. Dále $|rr| = 4$, zatímco $|sr| = 3$, proto budeme dále uvažovat pouze $k = 1, \ell = 1$. Nakonec se dá ověřit, že opravdu $|v| = |sr| = 3$ má nejmenší délku mezi návratovými slovy k 0.

Pokud bychom vybrali faktor 1, pak ačkoliv má stejnou délku i stejnou projekci $\pi(1) = \mathtt{a}$, jeho návratová slova jsou delší než 3: mezera pro 1 je 4, a tudíž po 2 písmenech $\mathtt{a}$, které se ve faktoru sr vyskytují, nenásleduje v $\mathbf{y}$, a tudíž ani v balancovaném slově, písmeno 1, ale 2.

Kapitola 5

Výpočet asymptotického kritického exponentu

Asymptotický kritický exponent odráží chování exponentů pro délku faktoru jdoucí do nekonečna. Nemusíme se tudíž zabývat faktory, které jsou příliš krátké a mohly by při výpočtu „dělat problémy“.

V této práci se zaměříme na výpočet asymptotického kritického exponentu pro stejnoměrně rekurentní aperiodická balancovaná slova, u kterých ještě budeme předpokládat speciální tvar θ – chceme, aby θ byla kvadratická iracionalita, tj. s posléze periodickým řetězovým zlomkem

$$\theta = [0, d_1, d_2, \dots, d_h, (z_0, z_1, \dots, z_{M-1})^\omega].$$

Délku předperiody θ budeme i dále značit h a délku periody M.

Pro výpočet využijeme vzorec z věty 2.2

$$\mathrm{E}^*(\mathbf{v}) = 1 + \limsup_{n\to\infty} \frac{|w_n|}{|v_n|},$$

kde $(w_n)_{n\in\mathbb{N}}$ je posloupnost všech bispeciálů ve $\mathbf{v}$ seřazených dle délky a v_n je nejkratší návratové slovo k w_n ve $\mathbf{v}$ pro všechna $n \in \mathbb{N}$.

Protože limita nezávisí na konečném počtu prvních členů, stačí se zaměřit na bispeciály w „příslušné periodě“, tzn. takové, kterým přísluší dvojice (N, m) pro $N \geq h$. Dále stačí zkoumat pouze bispeciály w, které splňují $|\pi(w)|_a > \beta(\mathbf{y})$ a $|\pi(w)|_b > \beta(\mathbf{y}')$. V případě potřeby proto prodloužíme předperiodu h tak, aby toto bylo splněno pro všechny bispeciály v periodě. Pro takové bispeciály se množina $\mathcal{S}_2$ z věty 4.4 redukuje na

$$\mathcal{S}_2(u) = \left\{ \begin{pmatrix} \ell \\ k \end{pmatrix} : k\vec{V}(r) + \ell\vec{V}(s) = \begin{pmatrix} 0 & \mod \mathrm{Per}(\mathbf{y}) \\ 0 & \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix} \right\}. \tag{5.1}$$

Můžeme si tedy všimnout, že pro výpočet vůbec nepotřebujeme znát tvar slov $\mathbf{y}$ a $\mathbf{y}'$, zajímá nás pouze jejich perioda a skutečnost, že to jsou dvě slova s konstantními mezerami nad disjunktními abecedami.

Výpočet asymptotického kritického exponentu nám umožní následující ekvivalence a její důsledky.

Definice 5.1. *Nechť $N_1, N_2 \in \mathbb{N}$ takové, že $N_1, N_2 \geq h$. Pak řekneme, že $N_1 \sim N_2$ právě tehdy, když*

1. $N_1 \equiv N_2 \mod M$,
2. $\begin{pmatrix} p_{N_1-1} & \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_1-1} & \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix} = \begin{pmatrix} p_{N_2-1} & \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_2-1} & \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix}$,

3. $\begin{pmatrix} p_{N_1} \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_1} \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix} = \begin{pmatrix} p_{N_2} \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_2} \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix}$.

Tříd této ekvivalence je nejvýše $M \cdot (\mathrm{Per}(\mathbf{y})\mathrm{Per}(\mathbf{y}'))^2$, jejich skutečný počet budeme dále v textu značit H. Nejdůležitější motivací pro zavedení ekvivalence právě takto je následující věta.

Věta 5.1. *Nechť $b^{(1)}, b^{(2)}$ jsou bispeciály v $\mathbf{u}$ a $(N_1, m_1), (N_2, m_2)$, kde $N_1, N_2 \geq h$, příslušné dvojice. Pokud $N_1 \sim N_2$ a $m_1 = m_2$, pak i $\mathcal{S}(b^{(1)}) = \mathcal{S}(b^{(2)})$.*

Důkaz.

- Označme $i = N_1 - h \mod M$. Pak protože $N_1 \equiv N_2 \mod M$, získáme $i = N_2 - h \mod M$. Dále $m_1 = m_2 =: m$. Z věty 3.8 získáme $\theta'_1 = \theta'_2 = \theta' = [0, (z_i - m), (z_{i+1}, \ldots, z_{M-1}, z_0, z_1, \ldots, z_i)^\omega]$, kde θ'_j je řetězový zlomek příslušný k derivovanému slovu bispeciálu s dvojicí (N_j, m_j), $j \in \{1, 2\}$. Protože řetězové zlomky příslušných derivovaných slov jsou shodné, budou obě derivovaná slova shodná, a proto i množiny $\mathcal{S}_1$ se pro obě slova budou shodovat.

- Využijeme vztahu pro Parikhovy vektory návratových slov k bispeciálu. Pak
$$\mathcal{S}_2(b^{(1)}) = \left\{ \begin{pmatrix} \ell \\ k \end{pmatrix} : k \begin{pmatrix} p_{N_1} \\ q_{N_1} \end{pmatrix} + \ell \begin{pmatrix} m \cdot p_{N_1} + p_{N_1 - 1} \\ m \cdot q_{N_1} + q_{N_1 - 1} \end{pmatrix} = \begin{pmatrix} 0 \mod \mathrm{Per}(\mathbf{y}) \\ 0 \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix} \right\}.$$
  Protože
$$\begin{pmatrix} p_{N_1-1} \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_1-1} \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix} = \begin{pmatrix} p_{N_2-1} \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_2-1} \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix} \text{ a } \begin{pmatrix} p_{N_1} \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_1} \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix} = \begin{pmatrix} p_{N_2} \mod \mathrm{Per}(\mathbf{y}) \\ q_{N_2} \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix},$$
  množina se nezmění, pokud budeme místo N_1 uvažovat N_2. Proto i množiny $\mathcal{S}_2$ budou shodné.

- Množina $\mathcal{S}_3$ nezávisí na zadaném bispeciálu, je stejná pro všechny bispeciály.

Ukázali jsme, že se shodují jednotlivé množiny v průniku, proto jsou shodné i množiny $\mathcal{S}(b^{(1)})$ a $\mathcal{S}(b^{(2)})$. □

Nyní trochu upravíme vztah $1 + \frac{|w|}{|v|}$ za pomoci již vyslovených vět 3.5 a 4.4. Označíme $Q_N := p_N + q_N$. Posloupnost $(Q_N)_{N \geq -1}$ pak splňuje rekurenci $Q_{N+1} = a_{N+1} Q_N + Q_{N-1}$, $N \geq 0$, s počátečními podmínkami $Q_{-1} = 1 = Q_0$.

Pozorování. Nechť u_n je n-tý bispeciál ve sturmovském slově $\mathbf{u}$ a $w \in \mathcal{L}(\mathbf{v})$ je příslušný bispeciál v balancovaném slově, tzn. $\pi(w) = u_n$. Dále nechť k n přísluší dvojice (N, m), $N \geq h$. Pak platí

1. délka prefixového návratového slova r k u_n je $|r| = Q_N$,
2. délka neprefixového návratového slova s k u_n je $|s| = m \cdot Q_N + Q_{N-1}$,
3. délka bispeciálů u_n i w je $|u_n| = |w| = |r| + |s| - 2 = (m + 1)Q_N + Q_{N-1} - 2$,
4. délka nejkratšího návratového slova v k bispeciálu w ve $\mathbf{v}$ splňuje
$$|v| = \min \left\{ k|r| + \ell|s| : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(b) \right\} = \min \left\{ (k + \ell m)Q_N + \ell Q_{N-1} : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(b) \right\}.$$
5. označíme výraz $\frac{|w|}{|v|}$ jako $\mathrm{I}(N, m)$ pro dané (N, m) a získáme
$$\begin{aligned} 1 + \frac{|w|}{|v|} = 1 + \mathrm{I}(N, m) &= 1 + \max \left\{ \frac{(m + 1)Q_N + Q_{N-1} - 2}{(k + \ell m)Q_N + \ell Q_{N-1}} : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(b) \right\} \\ &= 1 + \max \left\{ \frac{(m + 1) + \frac{Q_{N-1}}{Q_N} - \frac{2}{Q_N}}{(k + \ell m) + \ell \frac{Q_{N-1}}{Q_N}} : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(b) \right\}. \end{aligned} \tag{5.2}$$

Nyní si rozdělíme dlouhé bispeciály slova $\mathbf{u}$ do tříd $\mathrm{C}(i, m)$ pro všechna $i \in \{0, 1, \ldots, H-1\}$ a všechna $m \in \{0, 1, \ldots, z_{i \mod M} - 1\}$ tak, že bispeciál, kterému přísluší (N_1, m_1), patří do třídy $C(i, m)$ právě tehdy, když $N_1 \geq h$, $N_1 \sim (i + h)$. Již víme, že všechny bispeciály v jedné třídě mají stejné množiny $\mathcal{S}(b)$, tuto skutečnost budeme dále značit $\mathcal{S}(i, m) := \mathcal{S}(b)$ pro $b \in \mathrm{C}(i, m)$. Z vlastnosti ekvivalence $\sim$ lze ukázat, že bispeciály ve třídě $\mathrm{C}(i, m)$ jsou právě ty, kterým přísluší dvojice $(h + i + tH, m)$ pro $t \in \mathbb{N}_0$.

Označíme dále

$$\mathrm{E}^*_{\mathbf{v}}(i, m) = 1 + \limsup_{t \to +\infty} \mathrm{I}(h + i + tH, m).$$

Věta 5.2. *Nechť* $\theta = [0, d_1, d_2, \ldots, d_h, (z_0, z_1, \ldots, z_{M-1})^\omega]$. *Pak*

$$L_i := \lim_{t \to +\infty} \frac{Q_{i+h+tH-1}}{Q_{i+h+tH}} \tag{5.3}$$

existuje a je konečná pro všechna $i \in \{0, 1, \ldots, H - 1\}$.

Důkaz. Definujeme-li matici $\mathbb{A}^{(j)}$ pro $j \in \{0, 1, \ldots, M - 1\}$ pomocí koeficientů periody θ následovně

$$\mathbb{A}^{(j)} = \begin{pmatrix} 0 & 1 \\ 1 & z_j \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & z_{j+1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & z_{M-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & z_0 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & z_{j-1} \end{pmatrix}, \tag{5.4}$$

pak řádkové vektory $(Q_{h+i+tH}, Q_{h+i+tH-1})$ splňují pro všechna $i \in \{0, 1, \ldots, H - 1\}$ rekurenci

$$(Q_{h+i+tH-1}, Q_{h+i+tH}) = (Q_{h+i-1}, Q_{h+i}) \cdot \left(\mathbb{A}^{(i \mod M)}\right)^{t\frac{H}{M}}.$$

Je-li λ vlastní číslo matice $\mathbb{A}^{(j)}$, kde $j = (i \mod M)$ takové, že $|\lambda| < 1$ (a takové vlastní číslo vždy existuje), a vlastní vektor k tomuto vlastnímu číslu má složky $\vec{x}^{(j)} = \begin{pmatrix} x_j \\ y_j \end{pmatrix}$, pak vynásobením předchozí rovnosti zprava vlastním vektorem $\vec{x}^{(j)}$ získáme

$$x_j Q_{h+i+tH-1} + y_j Q_{h+i+tH} = \lambda^{t\frac{H}{M}} (x_j Q_{h+i-1} + y_j Q_{h+i}).$$

Nakonec úpravou a provedením limity $t \to +\infty$ (pravá strana jde limitně k 0, protože $|\lambda| < 1$) dostaneme

$$L_i = \lim_{t \to +\infty} \frac{Q_{h+i+tH-1}}{Q_{h+i+tH}} = -\frac{y_j}{x_j}. \tag{5.5}$$

Ze vzorce můžeme vidět, že posloupnost L_i je periodická s periodou M. □

Růst posloupnosti $Q_N \to +\infty$ znamená, že $\lim_{N \to +\infty} \frac{2}{Q_N} = 0$. Výraz pro $\mathrm{E}^*_{\mathbf{v}}(i, m)$ tudíž můžeme napsat v řeči limity L_i (provedli jsme záměnu limity a maxima z konečného počtu hodnot) a tento výraz bude konečný

$$\mathrm{E}^*_{\mathbf{v}}(i, m) = 1 + \max \left\{ \frac{(m+1) + L_i - 0}{(k + \ell m) + \ell \cdot L_i} : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(i, m) \right\}. \tag{5.6}$$

Nyní již stačí vybrat maximum ze všech $\mathrm{E}^*_{\mathbf{v}}(i, m)$ a získáme tak konečný výraz pro asymptotický kritický exponent

$$\mathrm{E}^*(\mathbf{v}) = \max \left\{ \mathrm{E}^*_{\mathbf{v}}(i, m) : i \in \{0, 1, \ldots, H - 1\}, m \in \{0, 1, \ldots, z_{i \mod M} - 1\} \right\}.$$

Příklad 5.1. Nalezneme asymptotický kritický exponent pro balancované slovo s následujícími parametry

$$\theta = [0, 1, 1, 1, (3, 1)^\omega], \quad \mathbf{y} = (01)^\omega, \quad \mathbf{y}' = (234235)^\omega.$$

Vidíme, že $h = 3$, $M = 2$, $\mathrm{Per}(\mathbf{y}) = 2$ a $\mathrm{Per}(\mathbf{y}') = 6$. Dále nalezneme H. Budeme procházet matice tvaru

$$\begin{pmatrix} p_{N-1} \mod \mathrm{Per}(\mathbf{y}) & p_N \mod \mathrm{Per}(\mathbf{y}) \\ q_{N-1} \mod \mathrm{Per}(\mathbf{y}') & q_N \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix}$$

a sledovat hodnotu $j = N - h \mod M$, která je indexem v periodě řetězového zlomku θ. Ve chvíli, kdy se nám matice zopakuje i s příslušným indexem j, víme, že jsme našli všechny třídy ekvivalence $\sim$. K tomu nám pomůže následující tabulka.

Tabulka 5.1: Členy posloupnosti (p_n) a (q_n) s příslušnými koeficienty řetězového zlomku θ.

n	2	3	4	5	6	7	8	9
a_n	1	1	3	1	3	1	3	1
p_n	1	2	7	9	34	43	163	206
q_n	2	3	11	14	53	67	254	321

$$\begin{pmatrix} p_{h-1} & p_h \\ q_{h-1} & q_h \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \backsim \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, N - h = 3 - 3 = 0,$$

$$\begin{pmatrix} p_h & p_{h+1} \\ q_h & q_{h+1} \end{pmatrix} \backsim \begin{pmatrix} 0 & 1 \\ 3 & 5 \end{pmatrix}, N - h = 4 - 3 \equiv 1 \mod 2,$$

$$\begin{pmatrix} p_{h+1} & p_{h+2} \\ q_{h+1} & q_{h+2} \end{pmatrix} \backsim \begin{pmatrix} 1 & 1 \\ 5 & 2 \end{pmatrix}, N - h = 5 - 3 \equiv 0 \mod 2,$$

$$\begin{pmatrix} p_{h+2} & p_{h+3} \\ q_{h+2} & q_{h+3} \end{pmatrix} \backsim \begin{pmatrix} 1 & 0 \\ 2 & 5 \end{pmatrix}, N - h = 6 - 3 \equiv 1 \mod 2,$$

$$\begin{pmatrix} p_{h+3} & p_{h+4} \\ q_{h+3} & q_{h+4} \end{pmatrix} \backsim \begin{pmatrix} 0 & 1 \\ 5 & 1 \end{pmatrix}, N - h = 7 - 3 \equiv 0 \mod 2,$$

$$\begin{pmatrix} p_{h+4} & p_{h+5} \\ q_{h+4} & q_{h+5} \end{pmatrix} \backsim \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, N - h = 8 - 3 \equiv 1 \mod 2,$$

$$\begin{pmatrix} p_{h+5} & p_{h+6} \\ q_{h+5} & q_{h+6} \end{pmatrix} \backsim \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, N - h = 9 - 3 \equiv 0 \mod 2.$$

V poslední matici jsme v periodě řetězového zlomku θ na stejné pozici a začínáme se stejnou maticí. H je tedy rovno 6 a budeme muset projít třídy ekvivalence $\mathrm{C}(i, m)$ pro

$$(i, m) \in \{(0, 0), (0, 1), (0, 2), (1, 0), (2, 0), (2, 1), (2, 2), (3, 0), (4, 0), (4, 1), (4, 2), (5, 0)\}.$$

Dále vypočítáme hodnoty L_j pro $j = 0, 1$ podle vzorce (5.5).

$$\mathbb{A}^{(0)} = \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} \to (1 - \lambda)(4 - \lambda) - 3 = 0 \wedge |\lambda| < 1 \to \lambda = \frac{5 - \sqrt{21}}{2},$$

$$\vec{x} = \begin{pmatrix} 1 \\ \frac{3 - \sqrt{21}}{2} \end{pmatrix} \to L_0 = \frac{\sqrt{21} - 3}{2}.$$

$$\mathbb{A}^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} \to (1 - \lambda)(4 - \lambda) - 3 = 0 \wedge |\lambda| < 1 \to \lambda = \frac{5 - \sqrt{21}}{2},$$

$$\vec{x} = \begin{pmatrix} 3 \\ \frac{3 - \sqrt{21}}{2} \end{pmatrix} \to L_1 = \frac{\sqrt{21} - 3}{6}.$$

Nakonec z periodicity L_i získáme $L_i = \begin{cases} \frac{\sqrt{21}-3}{2} & \text{pro } i \text{ sudé}, \\ \frac{\sqrt{21}-3}{6} & \text{pro } i \text{ liché}. \end{cases}$

C(0,0): Máme $\mathcal{S}_2(0,0) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}0\\3\end{pmatrix} + \ell\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší dvojice $k = 2, \ell = 0$ a faktor `rr` je jistě faktorem derivovaného slova s $\theta_{(0,0)} = [0, (3,1)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(0,0) = 1 + \frac{(0+1) + \frac{\sqrt{21}-3}{2}}{2 + 0 + 0 \cdot \frac{\sqrt{21}-3}{2}} = 1 + \frac{\sqrt{21}-1}{4} \doteq 1{,}8956.$$

C(0,1): Máme $\mathcal{S}_2(0,1) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}0\\3\end{pmatrix} + \ell\begin{pmatrix}0+1\\3+2\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší nejmenší dvojice $k = 2$, $\ell = 0$ a faktor `rr` je jistě faktorem derivovaného slova s $\theta_{(0,1)} = [0, 2, (1,3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(0,1) = 1 + \frac{(1+1) + \frac{\sqrt{21}-3}{2}}{2 + 0 \cdot 1 + 0} = 1 + \frac{\sqrt{21}+1}{4} \doteq 2{,}3956.$$

C(0,2): Máme $\mathcal{S}_2(0,2) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}0\\3\end{pmatrix} + \ell\begin{pmatrix}2\cdot 0+1\\2\cdot 3+2\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší nejmenší dvojice $k = 2, \ell = 0$ a faktor `rr` je jistě faktorem derivovaného slova s $\theta_{(0,2)} = [0, 1, (1,3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(0,2) = 1 + \frac{(2+1) + \frac{\sqrt{21}-3}{2}}{2 + 0 \cdot 2 + 0} = 1 + \frac{\sqrt{21}+3}{4} \doteq 2{,}8956.$$

C(1,0): Máme $\mathcal{S}_2(1,0) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\5\end{pmatrix} + \ell\begin{pmatrix}0\\3\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší nejmenší dvojice $k = 0$, $\ell = 2$, ale taková dvojice neleží v $\mathcal{S}_1(1,0)$. Další dvojicí je $k = 6, \ell = 4$. Prefix slova s řetězovým zlomkem $\theta_{(1,0)} = [0, (1,3)^\omega]$ vypadá následovně

$$DG^3DG^2(\mathtt{b}) = \mathtt{bababa}\underline{\mathtt{bbabababab}}\mathtt{babababa}\ldots$$

Můžeme se proto přesvědčit, že například faktor `rrsrsrsrsr` je faktorem derivovaného slova a má příslušný Parikhův vektor. Proto

$$\mathrm{E}^*_{\mathbf{v}}(1,0) = 1 + \frac{(0+1) + \frac{\sqrt{21}-3}{6}}{6 + 4 \cdot 0 + 4 \cdot \frac{\sqrt{21}-3}{6}} = 1 + \frac{\sqrt{21}+3}{4(\sqrt{21}+6)} \doteq 1{,}1791.$$

C(2,0): Máme $\mathcal{S}_2(2,0) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\2\end{pmatrix} + \ell\begin{pmatrix}1\\5\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší dvojice $k = 4, \ell = 2$ a faktor `srrrrs` je faktorem derivovaného slova s $\theta_{(2,0)} = [0, (3,1)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(2,0) = 1 + \frac{(0+1) + \frac{\sqrt{21}-3}{2}}{4 + 2 \cdot 0 + 2 \cdot \frac{\sqrt{21}-3}{2}} = 1 + \frac{\sqrt{21}-1}{2(\sqrt{21}+1)} \doteq 1{,}3209.$$

C(2,1): Máme $\mathcal{S}_2(2,1) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\2\end{pmatrix} + \ell\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší dvojice $k = 2, \ell = 2$ a faktor `srrs` je jistě faktorem derivovaného slova s $\theta_{(2,1)} = [0, 2, (1,3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(2,1) = 1 + \frac{(1+1) + \frac{\sqrt{21}-3}{2}}{2 + 2 \cdot 1 + 2 \cdot \frac{\sqrt{21}-3}{2}} = 1 + \frac{1}{2} = 1{,}5.$$

C(2,2): Máme $\mathcal{S}_2(2,2) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\2\end{pmatrix} + \ell\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší dvojice $k = 6, \ell = 2$ a faktor `rrsrrsrr` je faktorem derivovaného slova s $\theta_{(2,2)} = [0, 1, (1, 3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(2,2) = 1 + \frac{(2+1) + \frac{\sqrt{21}-3}{2}}{6 + 2\cdot 2 + 2\cdot\frac{\sqrt{21}-3}{2}} = 1 + \frac{\sqrt{21}+3}{2(\sqrt{21}+7)} \doteq 1{,}3273.$$

C(3,0): Máme $\mathcal{S}_2(3,0) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}0\\5\end{pmatrix} + \ell\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší nejmenší dvojice $k = 4$, $\ell = 2$ a faktor `rrsrsr` je faktorem derivovaného slova s $\theta_{(3,0)} = [0, (1, 3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(3,0) = 1 + \frac{(0+1) + \frac{\sqrt{21}-3}{6}}{4 + 2\cdot 0 + 2\cdot\frac{\sqrt{21}-3}{6}} = 1 + \frac{\sqrt{21}+3}{2(\sqrt{21}+9)} \doteq 1{,}2791.$$

C(4,0): Máme $\mathcal{S}_2(4,0) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\1\end{pmatrix} + \ell\begin{pmatrix}0\\5\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší dvojice $k = 8, \ell = 2$ a faktor `rrrrsrrrrs` je faktorem derivovaného slova s $\theta_{(4,0)} = [0, (3, 1)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(4,0) = 1 + \frac{(0+1) + \frac{\sqrt{21}-3}{2}}{8 + 2\cdot 0 + 2\cdot\frac{\sqrt{21}-3}{2}} = 1 + \frac{\sqrt{21}-1}{2(\sqrt{21}+5)} \doteq 1{,}1869.$$

C(4,1): Máme $\mathcal{S}_2(4,1) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\1\end{pmatrix} + \ell\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší dvojice $k = 6, \ell = 2$ (dvojice $k = 0$, $\ell = 2$ neleží v $\mathcal{S}_1(4,1)$) a faktor `rrrsrrsr` je faktorem derivovaného slova s $\theta_{(4,1)} = [0, 2, (1, 3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(4,1) = 1 + \frac{(1+1) + \frac{\sqrt{21}-3}{2}}{6 + 2\cdot 1 + 2\cdot\frac{\sqrt{21}-3}{2}} = 1 + \frac{\sqrt{21}+1}{2(\sqrt{21}+5)} = 1{,}2913.$$

C(4,2): Máme $\mathcal{S}_2(4,2) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\1\end{pmatrix} + \ell\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší dvojice $k = 4, \ell = 2$ a faktor `rrsrsr` je faktorem derivovaného slova s $\theta_{(4,2)} = [0, 1, (1, 3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(4,2) = 1 + \frac{(2+1) + \frac{\sqrt{21}-3}{2}}{4 + 2\cdot 2 + 2\cdot\frac{\sqrt{21}-3}{2}} = 1 + \frac{\sqrt{21}+3}{2(\sqrt{21}+5)} \doteq 1{,}3956.$$

C(5,0): Máme $\mathcal{S}_2(5,0) = \left\{\begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\2\end{pmatrix} + \ell\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 6\end{pmatrix}\right\}$. Tuto rovnici řeší nejmenší dvojice $k = 2$, $\ell = 2$ a faktor `srrs` je faktorem derivovaného slova s $\theta_{(5,0)} = [0, (1, 3)^\omega]$. Proto

$$\mathrm{E}^*_{\mathbf{v}}(5,0) = 1 + \frac{(0+1) + \frac{\sqrt{21}-3}{6}}{2 + 2\cdot 0 + 2\cdot\frac{\sqrt{21}-3}{6}} = 1 + \frac{1}{2} = 1{,}5.$$

Asymptotický kritický exponent je tudíž

$$\mathrm{E}^*(\mathbf{v}) = \max\left\{\mathrm{E}^*_{\mathbf{v}}(i,m) : 0 \le i \le 5 \wedge 0 \le m \le z_{i \mod 2} - 1\right\} = \mathrm{E}^*_{\mathbf{v}}(0,2) = 1 + \frac{\sqrt{21}+3}{4} \doteq 2{,}8956.$$

Čtenář je vítán, aby sám ověřil, zda faktory v mezivýpočtech jsou opravdu faktory derivovaných slov, k vygenerování prefixu mu pomůže věta 3.3 a příklad 4.1.

Kapitola 6

Výpočet kritického exponentu

I v této kapitole budeme předpokládat (pokud nebude uvedeno jinak), že $\mathbf{u}$ je standardní sturmovské slovo, $\mathbf{y}$ a $\mathbf{y}'$ jsou dvě slova s konstantními mezerami nad disjunktními abecedami a $\mathbf{v} = \mathrm{colour}(\mathbf{u}, \mathbf{y}, \mathbf{y}')$ je balancované slovo. I nyní výpočet popíšeme pouze pro stejnoměrně rekurentní aperiodická balancovaná slova, pro které má θ posléze periodický řetězový zlomek, tzn. $\theta = [0, d_1, d_2, \dots, d_h, (z_0, z_1, \dots, z_{M-1})^\omega]$. Délku předperiody θ budeme i nadále značit h a délku periody M.

Pro výpočty využijeme vzorec z věty 2.2

$$\mathrm{E}(\mathbf{v}) = 1 + \sup_{n \in \mathbb{N}} \left\{ \frac{|w_n|}{|v_n|} \right\},$$

kde $(w_n)_{n\in\mathbb{N}}$ je posloupnost všech bispeciálů ve $\mathbf{v}$ seřazených dle délky a $(v_n)_{n\in\mathbb{N}}$ je posloupnost nejkratších návratových slov k w_n.

Samotný výpočet si rozdělíme do tří částí, a to na krátké, středně dlouhé a dlouhé bispeciály.

6.1 Krátké bispeciály

Definice 6.1. *Za krátké bispeciály považujeme bispeciály $w \in \mathcal{L}(\mathbf{v})$ takové, že $|\pi(w)|_{\mathrm{a}} \leq \beta(\mathbf{y})$ nebo $|\pi(w)|_{\mathrm{b}} \leq \beta(\mathbf{y}')$. Množinu všech krátkých bispeciálů označíme*

$$\mathcal{W}^{short} := \{w \in \mathcal{L}(\mathbf{v}) : w \text{ bispeciál a } |\pi(w)|_a \leq \beta(\mathbf{y}) \text{ nebo } |\pi(w)|_b \leq \beta(\mathbf{y}')\}.$$

Tyto bispeciály budeme vyšetřovat zvlášť, protože projekce těchto faktorů nemusí být v $\mathbf{u}$ vůbec bispeciály, ale ve $\mathbf{v}$ se na bispeciály mohou přepsat. K výpočtu délek nejkratších návratových slov krátkých bispeciálů potřebujeme znát Parikhovy vektory návratových slov k libovolnému faktoru $\mathbf{u}$, nejen k bispeciálům. K tomu nám poslouží následující pozorování.

Pozorování. Je-li u faktor aperiodického rekurentního slova $\mathbf{u}$, pak ho lze jednoznačně natáhnout na nejkratší bispeciál b, který tento faktor obsahuje. Bispeciál b je tedy tvaru $b = xuy$ pro faktory $x, y \in \mathcal{L}(\mathbf{u})$ (mohou být i prázdné). Dále v je návratové slovo k u právě tehdy, když $z = xvx^{-1}$ je návratové slovo k bispeciálu b. Exponent -1 zde značí inverzní operaci ke zřetězení a znamená to, že z vznikne z faktoru xv „useknutím“ sufixu x (naopak $v = x^{-1}zx$). Speciálně Parikhovy vektory návratových slov k b a k u se rovnají. Dále také $\mathrm{d}_{\mathbf{u}}(b) = \mathrm{d}_{\mathbf{u}}(u)$.

Tudíž máme-li faktor u standardního sturmovského slova, pak ho stačí natáhnout na nejkratší bispeciál b, který tento faktor obsahuje, a najít návratová slova k němu. Pro to již budeme moci použít větu 3.5. Následně budeme muset najít množinu $\mathcal{S}(u)$ podle vzorce z věty 4.4, kde pro krátké bispeciály musíme počítat s množinami $\mathrm{gap}(\mathbf{y}, |u|_{\mathrm{a}})$ a $\mathrm{gap}(\mathbf{y}', |u|_{\mathrm{b}})$.

Výpočet je shrnut v následující větě.

Věta 6.1. *Nechť* $w \in \mathcal{W}^{short}$, $\pi(w) = u$. *Nejkratší návratové slovo k* w *ve* $\mathbf{v}$ *označíme* v. *Nechť dále* $u_n \in \mathcal{L}(\mathbf{u})$ *je nejkratší bispeciál obsahující faktor* u *a k bispeciálu* u_n *přísluší dvojice* (N, m)*, pak*

$$1 + \frac{|w|}{|v|} = 1 + \frac{|u|}{|v|} \leq 1 + \max\left\{ \frac{|u|}{(k + \ell m)Q_N + \ell Q_{N-1}} : \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(u) \right\} \tag{6.1}$$

a existuje faktor $w' \in \mathcal{W}^{short}$ *splňující* $\pi(w') = u$ *takový, že ve vzorci* (6.1) *nastává rovnost.*

Doplníme ještě jedno pozorování ospravedlňující to, že dále ve výpočtech nebudeme kontrolovat, zda $w \in \mathcal{L}(\mathbf{v})$ takový, že jeho projekce je krátká, je bispeciálem.

Pozorování. Natáhneme-li faktor w na nejkratší bispeciál w', který w obsahuje, tak délka jejich nejkratšího návratového slova $|v| = |v'|$ je stejná, proto $1 + \frac{|w|}{|v|} \leq 1 + \frac{|w'|}{|v|}$, tudíž ve větě 2.2 lze místo posloupnosti bispeciálů uvažovat libovolnou posloupnost faktorů, která má bispeciály jako svou podposloupnost.

Dále označíme

$$\mathrm{E}^{short}(\mathbf{v}) = 1 + \max\left\{ \frac{|w|}{(k + \ell m)Q_N + \ell Q_{N-1}} : w \in \mathcal{W}^{short} \text{ a } \begin{pmatrix} \ell \\ k \end{pmatrix} \in \mathcal{S}(\pi(w)) \right\}.$$

6.2 Středně dlouhé bispeciály

Do této skupiny patří bispeciály, jejichž projekce je bispeciálem v $\mathbf{u}$, ale ve dvojici (N, m) je $N < h$. Množinu všech těchto bispeciálů označíme $\mathcal{W}^{med}$. Tato množina není disjunktní s množinou $\mathcal{W}^{short}$, nám stačí vyšetřit délku nejkratších návratových slov pouze pro ty, které do $\mathcal{W}^{short}$ nepatří.

Tyto bispeciály budeme vyšetřovat pomocí vzorce (5.2), výpočet $\mathcal{S}_2(u)$ se opět zjednoduší na tvar (5.1). Označíme

$$\mathrm{E}^{med}(\mathbf{v}) = 1 + \max\{\mathrm{I}(N, m) : N \in \mathbb{N}_0, N < h, m \in \{0, 1, \ldots, d_{N+1} - 1\}\}.$$

6.3 Dlouhé bispeciály

Nyní již nám zbývá vyšetřit bispeciály, jejichž obraz je také bispeciál a přísluší mu $N \geq h$, budeme tedy hojně využívat vzorce (5.2). Protože bispeciálů v této množině je nekonečně mnoho, budeme potřebovat ještě jednu větu, která nám umožní využít výsledku pro asymptotický kritický exponent a dále vyšetřit pouze konečnou množinu bispeciálů. Následující věta nám dává nutnou podmínku pro to, aby podíl $1 + \frac{|w|}{|v|}$ mohl nabývat větší hodnoty než asymptotický kritický exponent. Počet tříd ekvivalence $\sim$ z definice 5.1 budeme i nadále v textu značit H.

Věta 6.2. *Nechť* L_i *splňuje* (5.3) *pro každé* $i \in \{0, 1, \ldots H - 1\}$ *a* $\mathrm{I}(N, m)$ *je definováno vzorcem* (5.2). *Pak pro každé* $i \in \{0, 1, \ldots H - 1\}$*, každé* $m \in \{0, 1, \ldots, z_{i \bmod M} - 1\}$ *a každé* $t \in \mathbb{N}_0$ *platí*

$$1 + \mathrm{I}(h + i + tH, m) > \mathrm{E}^*_{\mathbf{v}}(i, m) \implies |Q_{h+i+tH-1} - L_i Q_{h+i+tH}| > 2L_i.$$

Důkaz. Budeme upravovat nerovnost $1 + \mathrm{I}(h + i + tH, m) > \mathrm{E}^*_{\mathbf{v}}(i, m)$. Pro zjednodušení zápisu označíme $S := Q_{h+i+tH-1}$ a $T := Q_{h+i+tH}$. Dále nechť $\begin{pmatrix} \ell \\ k \end{pmatrix}$ je vektor, pro který je nabýváno maximum v definici $\mathrm{I}(h + i + tH, m)$, viz vzorec (5.2).

Pak

$$\begin{aligned}
1 + \mathrm{I}(h+i+tH,m) = 1 + \frac{(1+m)T + S - 2}{(k+\ell m)T + \ell S} > \mathrm{E}^*_{\mathbf{v}}(i,m) > 1 + \frac{1+m+L_i}{k+\ell m+\ell L_i} \\
((1+m)T + S - 2)(k+\ell m+\ell L_i) > ((k+\ell m)T + \ell S)(1+m+L_i) \\
\ell(1+m)L_iT + (k+\ell m)S - 2(k+\ell m) - 2\ell L_i > (k+\ell m)TL_i + (1+m)\ell S \\
\ell L_iT + kS - 2(k+\ell m) - 2\ell L_i > kTL_i + \ell S \\
(k-\ell)(S - L_iT) > 2(k+\ell m+\ell L_i).
\end{aligned}$$

Dále využijeme toho, že $0 < L_i < 1$, proto $2(k+\ell m+\ell L_i) > 2(k+\ell m+\ell)L_i$. Nakonec si příklad rozdělíme na 2 případy (pro případ $k = \ell$ nemůže být splněna první nerovnost).

1. $k > \ell$

$$\begin{aligned}
(k-\ell)(S - L_iT) > 2(k+\ell m+\ell)L_i > 2(k-\ell)L_i \\
S - L_iT > 2L_i
\end{aligned}$$

2. $k < \ell$

$$\begin{aligned}
(\ell-k)(L_iT - S) > 2(k+\ell m+\ell)L_i > 2(\ell-k)L_i \\
L_iT - S > 2L_i
\end{aligned}$$

Celkem tedy $|S - L_iT| > 2L_i$, což jsme měli dokázat. □

Z důkazu věty 5.2 v kapitole 5 můžeme vidět další důležitý vztah, a to ten, že pro $|\lambda| < 1$ vlastní číslo matice $\mathbb{A}^{(j)}$ platí

$$|Q_{h+i+tH-1} - L_iQ_{h+i+tH}| = |\lambda|^{t\frac{H}{M}}|Q_{h+i-1} - L_iQ_{h+i}|.$$

Tato posloupnost je tedy v t klesající, z toho vyplývá následující věta.

Věta 6.3. *Nechť $i \in \{0, 1, \ldots, H-1\}$, L_i je limita definována vzorcem* (5.3) *a λ je nedominantní vlastní číslo matice $\mathbb{A}^{(i \mod M)}$ definované vzorcem* (5.4). *Pokud pro $N^{(i)} \in \mathbb{N}$ platí*

$$|\lambda|^{N^{(i)}\frac{H}{M}}|Q_{h+i-1} - L_iQ_{h+i}| \le 2L_i, \ pak\ 1 + \mathrm{I}(h+i+tH,m) \le \mathrm{E}^*_{\mathbf{v}}(i,m) \le \mathrm{E}^*(\mathbf{v}) \tag{6.2}$$

pro všechna $t \ge N^{(i)}$.

Pro výpočet tedy stačí pro každé $i \in \{0, 1, \ldots, H-1\}$ najít nejmenší $N^{(i)} \in \mathbb{N}_0$ takové, že splňuje nerovnost (6.2), a následně projít bispeciály odpovídající $(h+i+tH, m)$ pro $t \in \{0, 1, \ldots, N^{(i)}-1\}$ a $m \in \{0, 1, \ldots, z_{i \mod M} - 1\}$.

Výpočet tedy dále rozdělíme do tří kroků:

1. Vypočítáme asymptotický kritický exponent, při výpočtu si zapamatujeme H značící počet tříd ekvivalence $\sim$, limity L_i, popř. vypočítané $\theta_{(i,m)}$.

2. Vyšetříme konečnou množinu bispeciálů příslušejících periodě, pro které by mohla nastat nerovnost $1 + \frac{|w|}{|v|} > \mathrm{E}^*(\mathbf{v})$ dle věty 6.3. Pro výpočet můžeme využít vzorec (5.2) a výpočet $\mathcal{S}_2$ se opět zjednoduší na (5.1).

3. Vezmeme maximum z bodů 1 a 2 a označíme

$$\mathrm{E}^{long}(\mathbf{v}) = \max\left\{\mathrm{E}^*(\mathbf{v}), 1 + \mathrm{I}(h+i+tH,m) : i \in \mathbb{Z}_{H-1}, t \in \mathbb{Z}_{N^{(i)}-1} \text{ a } m \in \mathbb{Z}_{z_{i \mod M}-1}\right\}.$$

Kritický exponent pak získáme jako maximum ze všech mezivýpočtů, tzn.

$$\mathrm{E}(\mathbf{v}) = \max\left\{\mathrm{E}^{short}(\mathbf{v}), \mathrm{E}^{med}(\mathbf{v}), \mathrm{E}^{long}(\mathbf{v})\right\}.$$

Příklad 6.1. Vypočítáme kritický exponent stejného slova jako v příkladu 5.1

$$\theta = [0, 1, 1, 1, (3, 1)^{\omega}], \quad \mathbf{y} = (01)^{\omega}, \quad \mathbf{y}' = (234235)^{\omega}.$$

Již víme, že $h = 3$, $M = 2$, $\mathrm{Per}(\mathbf{y}) = 2$ a $\mathrm{Per}(\mathbf{y}') = 6$, $H = 6$ a

$$L_i = \begin{cases} \frac{\sqrt{21}-3}{2} & \text{pro } i \text{ sudé},\\ \frac{\sqrt{21}-3}{6} & \text{pro } i \text{ liché}. \end{cases}$$

Dále se podíváme na slova s konstantními mezerami. Vidíme, že $\beta(\mathbf{y}) = 0$ a $\beta(\mathbf{y}') = 2$, protože faktor 23 je bispeciálem v $\mathbf{y}'$.

Krátké faktory

V této části budeme muset projít faktory balancovaného slova, jejichž projekce u splňuje $|u|_{\mathtt{a}} \leq 0$ nebo $|u|_{\mathtt{b}} \leq 2$. Při výpočtu dále budeme potřebovat znát konkrétní tvar prefixu, ten vypadá takto

$$\mathtt{babbababbababbababbabbababbababbababbababbabbab}.$$

Z tvaru prefixu a ze skutečnosti, že sturmovská slova jsou balancovaná, můžeme určit, jak vypadá množina W^{short}.

$$\mathcal{W}^{short} = \{w \in \mathcal{L}(\mathbf{v}) : w \text{ bispeciál a } \pi(w) \in \{\mathtt{b}, \mathtt{ba}, \mathtt{bab}, \mathtt{a}, \mathtt{ab}, \mathtt{abb}, \mathtt{bb}, \mathtt{bba}, \mathtt{abba}, \mathtt{baba}, \mathtt{aba}, \mathtt{abab}\}\}.$$

Při výpočtu nebudeme kontrolovat, zda $w \in \mathcal{L}(\mathbf{v})$ splňující $|\pi(w)|_{\mathtt{a}} \leq \beta(\mathbf{y})$ nebo $|\pi(w)|_{\mathtt{b}} \leq \beta(\mathbf{y}')$ je bispeciál, budeme pracovat pouze s jejich projekcemi do $\mathbf{u}$. K tomuto kroku nás ospravedlňuje věta 4.4 a skutečnost, že maximum $I(N, m)$ je nabýváno právě na bispeciálech.

K výpočtu budeme dále potřebovat několik počátečních členů posloupností (p_n) a (q_n).

Tabulka 6.1: Prvních osm členů posloupností (p_n), (q_n) a (Q_n) a příslušné koeficienty řetězového zlomku θ.

n	-1	0	1	2	3	4	5	6
a_n			1	1	1	3	1	3
p_n	1	0	1	1	2	7	9	34
q_n	0	1	1	2	3	11	14	53
Q_n	1	1	2	3	5	18	23	87

Dále budeme potřebovat množiny možných mezer slov $\mathbf{y}$ a $\mathbf{y}'$.

Tabulka 6.2: Množiny $\mathrm{gap}(\mathbf{y}, n)$ a $\mathrm{gap}(\mathbf{y}', n)$ pro $n \leq 6$.

n	0	1	2	3	4	5	6
$\mathrm{gap}((01)^{\omega}, n)$	{1}	{2}	{2}	{2}	{2}	{2}	{2}
$\mathrm{gap}((234235)^{\omega}, n)$	{1}	{3,6}	{3,6}	{6}	{6}	{6}	{6}

`b:` Je to bispeciál ve slově **u** s příslušnou dvojicí $(1,0)$, návratová slova tudíž splňují $\vec{V}(r) = \begin{pmatrix}1\\1\end{pmatrix}$ a $\vec{V}(s) = \begin{pmatrix}0\\1\end{pmatrix}$. Pohledem do prefixu můžeme zkontrolovat, že to jsou faktory $r = \mathtt{ba}$, $s = \mathtt{b}$. Pomocí tabulky 6.2 zjistíme, že $n \in \{1\}$ a $n' \in \{3,6\}$, prozkoumáme tedy množinu

$$\mathcal{S}_2(\mathtt{b}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\1\end{pmatrix} + \ell\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0 \mod 1\\0 \mod 3 \text{ nebo } 6\end{pmatrix} \right\}.$$

Nejmenším přípustným řešením je dvojice $k = 1, \ell = 2$ a slovo $\mathtt{srs}$ je faktorem derivovaného slova s $\hat{\theta}_{(1,0)} = [0,1,1,(3,1)^\omega]$. Proto $|v| = k|r| + \ell|s| = 1\cdot 2 + 2\cdot 1 = 4$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{1}{4}.$$

`bab:` Je to bispeciál v **u** s příslušnou dvojicí $(2,0)$, návratová slova tudíž splňují $\vec{V}(r) = \begin{pmatrix}1\\2\end{pmatrix}$ a $\vec{V}(s) = \begin{pmatrix}1\\1\end{pmatrix}$. Pomocí tabulky 6.2 zjistíme, že $n \in \{2\}$ a $n' \in \{3,6\}$,

$$\mathcal{S}_2(\mathtt{bab}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\2\end{pmatrix} + \ell\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 3 \text{ nebo } 6\end{pmatrix} \right\}.$$

Nejmenším přípustným řešením je dvojice $k = 1, \ell = 1$ a $|v| = 1\cdot 3 + 1\cdot 2 = 5$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{3}{5}.$$

`ba:` Není bispeciálem v **u**, nejkratším bispeciálem, který ho obsahuje, je $\mathtt{bab}$, který jsme vyšetřili v předchozím kroku. Množiny mezer budou také stejné, protože $\operatorname{gap}(\mathbf{y}', 2) = \operatorname{gap}(\mathbf{y}', 1)$. Vyšetříme tedy množinu

$$\mathcal{S}_2(\mathtt{ba}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\2\end{pmatrix} + \ell\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 3 \text{ nebo } 6\end{pmatrix} \right\}.$$

se stejným řešením $k = 1, \ell = 1$ a $|v| = 5$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{2}{5}.$$

`a:` Není bispeciálem v **u**, nejkratší bispeciál, který ho obsahuje, je opět $\mathtt{bab}$. Nyní $n \in \{2\}$ a $n' \in \{1\}$, a tudíž vyšetříme množinu

$$\mathcal{S}_2(\mathtt{a}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}1\\2\end{pmatrix} + \ell\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 1\end{pmatrix} \right\}.$$

Nejmenším přípustným řešením je $k = 1, \ell = 1$ s $|v| = 1\cdot 3 + 1\cdot 2 = 5$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{1}{5}.$$

`ab:` I tento faktor roztáhneme na $\mathtt{bab}$ a opět vyšetřujeme stejnou množinu jako v případě $\mathtt{ba}$. Proto

$$1 + \frac{|w|}{|v|} = 1 + \frac{2}{5}.$$

`abba`: Není to bispeciál v **u**, dá se natáhnout na bispeciál `babbab` s příslušnou dvojicí $(3,0)$. Parikhovy vektory návratových slov jsou $\vec{V}(r) = \begin{pmatrix}2\\3\end{pmatrix}$ a $\vec{V}(s) = \begin{pmatrix}1\\2\end{pmatrix}$. Vyšetříme proto množinu

$$\mathcal{S}_2(\mathtt{abba}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}2\\3\end{pmatrix} + \ell\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 3 \text{ nebo } 6\end{pmatrix} \right\}.$$

Nejmenším přípustným řešením je $k = 1, \ell = 0$ s $|v| = 5 + 0$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{4}{5}.$$

`abb`: Není to bispeciál, natáhneme ho na nejkratší možný a získáme tak opět `babbab`. Vyšetříme proto množinu

$$\mathcal{S}_2(\mathtt{abb}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}2\\3\end{pmatrix} + \ell\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 3 \text{ nebo } 6\end{pmatrix} \right\}.$$

Nejmenším přípustným řešením je $k = 1, \ell = 0$ s $|v| = 5 + 0$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{3}{5}.$$

`bb`: I tento faktor lze natáhnout na `babbab`. Vyšetříme tudíž množinu

$$\mathcal{S}_2(\mathtt{bb}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}2\\3\end{pmatrix} + \ell\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}0 \mod 1\\0 \mod 3 \text{ nebo } 6\end{pmatrix} \right\}.$$

Nejmenším přípustným řešením je $k = 1, \ell = 0$ s $|v| = 5 + 0$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{2}{5}.$$

`bba`: Další faktor, který natáhneme na `babbab`, budeme vyšetřovat stejnou množinu jako pro `abb`, tudíž

$$1 + \frac{|w|}{|v|} = 1 + \frac{3}{5}.$$

`baba`: Toto není bispeciál v **u**, lze ho natáhnout na `babbababbab`, kterému přísluší dvojice $(3,1)$ a Parikhovy vektory příslušných návratových slov pak jsou $\vec{V}(r) = \begin{pmatrix}2\\3\end{pmatrix}$ a $\vec{V}(s) = \begin{pmatrix}3\\5\end{pmatrix}$. Vyšetříme tedy množinu

$$\mathcal{S}_2(\mathtt{baba}) = \left\{ \begin{pmatrix}\ell\\k\end{pmatrix} : k\begin{pmatrix}2\\3\end{pmatrix} + \ell\begin{pmatrix}3\\5\end{pmatrix} = \begin{pmatrix}0 \mod 2\\0 \mod 3 \text{ nebo } 6\end{pmatrix} \right\}.$$

Řešením je opět dvojice $k = 1, \ell = 0$ s $|v| = 5$.

$$1 + \frac{|w|}{|v|} = 1 + \frac{4}{5}.$$

`aba`: Tento faktor lze natáhnout opět na bispeciál `babbababbab`, vyšetřovali bychom stejnou množinu jako v předchozím příkladu.

$$1 + \frac{|w|}{|v|} = 1 + \frac{3}{5}.$$

`abab`: I tento faktor lze natáhnout na `babbababbab`, vyšetřovaná množina $\mathcal{S}_2$ zůstává stejná. Tudíž

$$1 + \frac{|w|}{|v|} = 1 + \frac{4}{5}.$$

Dále je vhodné čtenáře upozornit, že v textu nebylo kontrolováno, zda $\binom{\ell}{k} \in \mathcal{S}_1(u) \cap \mathcal{S}_3$, k chybě však nedošlo, protože kromě prvního bodu vycházely faktory `r` nebo `rs`, které jsou vždy faktory derivovaného slova k libovolnému faktoru sturmovského slova.

Nyní stačí vzít maximum z nalezených hodnot

$$\mathrm{E}_{\mathbf{v}}^{short} = \max\left\{1 + \frac{1}{4}, 1 + \frac{3}{5}, 1 + \frac{2}{5}, 1 + \frac{1}{5}, 1 + \frac{4}{5}\right\} = 1 + \frac{4}{5} = 1{,}8.$$

Středně dlouhé bispeciály
V rámci krátkých faktorů jsme vyšetřili rovnou i všechny bispeciály příslušné předperiodě, tato množina je tudíž celá podmnožinou $\mathcal{W}^{short}$, a není tedy nutné ji znovu vyšetřovat.

Dlouhé bispeciály
Z příkladu 5.1 již víme, že $\mathrm{E}^*(\mathbf{v}) = 1 + \frac{\sqrt{21}+3}{4}$. Dále potřebujeme spočítat menší vlastní číslo matice $\mathbb{A}^{(0)}$, které jsme použili již pro výpočet L_i. Máme $\mathbb{A}^{(0)} = \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}$, $\lambda = \frac{5-\sqrt{21}}{2} \doteq 0{,}2087$. K výpočtu dále budeme potřebovat rozšířit naši tabulku $(p_n), (q_n), (Q_n)$.

Tabulka 6.3: Dalších několik členů posloupností (p_n), (q_n) a (Q_n) s příslušnými koeficienty řetězového zlomku θ.

n	2	3	4	5	6	7	8	9	10
a_n	1	1	3	1	3	1	3	1	3
p_n	1	2	7	9	34	43	163	206	781
q_n	2	3	11	14	53	67	254	321	1217
Q_n	3	5	18	23	87	110	417	527	1998

Zopakujeme ještě vztah pro hodnoty L_i, abychom ho měli po ruce. $L_i = \begin{cases} \frac{\sqrt{21}-3}{2} & \text{pro } i \text{ sudé}, \\ \frac{\sqrt{21}-3}{6} & \text{pro } i \text{ liché}. \end{cases}$

$(i = 0)$

$$|Q_{3+0-1} - Q_{3+0}L_0| = |3 - 5L_0| \doteq 0{,}956 < 2L_0 \doteq 1{,}5826.$$

Nerovnost je splněna již pro $N^{(0)} = 0$, žádné další bispeciály vyšetřovat nemusíme.

$(i = 1)$

$$|Q_{3+1-1} - L_1 Q_{3+1}| = |5 - 18 \cdot L_1| \doteq 0{,}2523 < 2L_1 \doteq 0{,}5275.$$

Nerovnost je tedy splněna již pro $N^{(1)} = 0$.

$(i = 2)$

$$|Q_{3+2-1} - L_0 Q_{3+2}| = |18 - 23 \cdot L_0| \doteq 0{,}1996 < 2L_0 \doteq 1{,}5826.$$

Tudíž $N^{(2)} = 0$.

($i = 3$)

$$|Q_{3+3-1} - L_1 Q_{3+3}| = |23 - 87 \cdot L_1| \doteq 0{,}0527 < 2L_1 \doteq 0{,}5275.$$

Tudíž $N^{(3)} = 0$.

($i = 4$)

$$|Q_{3+4-1} - L_0 Q_{3+4}| = |87 - 110 \cdot L_0| \doteq 0{,}0417 < 2L_0 \doteq 1{,}5826.$$

Tudíž $N^{(4)} = 0$.

($i = 5$)

$$|Q_{3+5-1} - L_1 Q_{3+5}| = |110 - 417 \cdot L_1| \doteq 0{,}011 < 2L_1 \doteq 0{,}5275.$$

Tudíž $N^{(5)} = 0$.

Pro žádný dlouhý faktor nenastává $1 + \frac{|w|}{|v|} > \mathrm{E}^*(\mathbf{v})$. Kritický exponent proto je

$$\mathrm{E}(\mathbf{v}) = \max\left\{1 + \frac{4}{5}, 1 + \frac{\sqrt{21}+3}{4}\right\} = 1 + \frac{\sqrt{21}+3}{4} = \mathrm{E}^*(\mathbf{v}).$$

Kapitola 7

Popis programu

Programy v průběhu práce vznikly dva a byly postupně doplňovány. První program počítá asymptotický kritický exponent. Kromě výpočtu asymptotického kritického exponentu pro dané balancované slovo (zadá se předperioda a perioda theta θ a délky period $\mathbf{y}$ a $\mathbf{y}'$) umí program projít také všechny možné předperiody θ pro zadané ostatní parametry a nalézt minimální asymptotický kritický exponent pro tuto skupinu balancovaných slov. Druhý program umožňuje výpočet kritického exponentu, kam uživatel musí zadat jak tvar předperiody, tak periody řetězového zlomku θ a slova $\mathbf{y}$ a $\mathbf{y}'$, a program pak nalezne kritický exponent daného slova.

Oba programy byly psány objektově v jazyce C++, nejnovější verze je doplněna o uživatelské rozhraní s použitím prostředí a knihoven Qt [16]. Existují dvě jazykové verze – anglická a česká, obě je možno nalézt na webové stránce [15]:
`https://github.com/Opocedan/Kriticky_exponent_balancovanych_slov`.

7.1 Program pro výpočet asymptotického kritického exponentu

7.1.1 Hlavní okno

Program umožňuje zvolit, zda uživatel chce zadat pouze tvar periody řetězového zlomku θ a program má projít všechny možné předperiody, nebo zda uživatel zadá i předperiodu řetězového zlomku θ. Dále je potřeba zadat délky nejmenších period slov s konstantními mezerami $\mathbf{y}$ a $\mathbf{y}'$. Nakonec jsou k dispozici 4 vrstvy výpisů podle toho, jak podrobný výpis uživatel chce a které části výpočtu uživatele zajímají:

- `Vypsat seznam matic předperiod` vypíše seznam matic předperiod ve tvaru
$$\begin{pmatrix} p_{h-1} & p_h \\ q_{h-1} & q_h \end{pmatrix} \mod \begin{pmatrix} \mathrm{Per}(\mathbf{y}) \\ \mathrm{Per}(\mathbf{y}') \end{pmatrix},$$
které jsou následně procházeny.

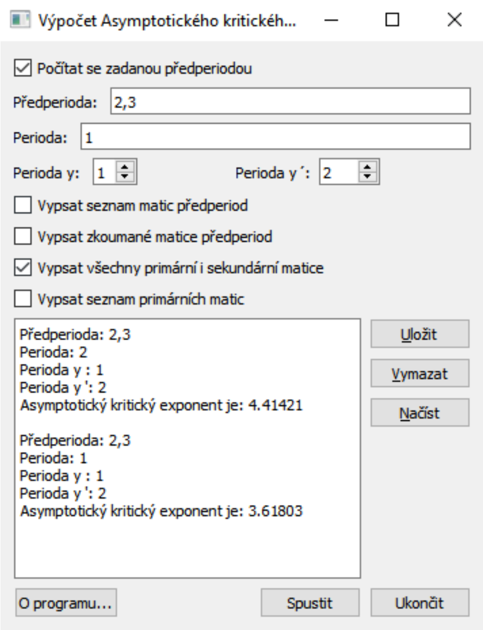
- `Vypsat zkoumané matice předperiod` vypíše v průběhu výpočtu právě zkoumanou matici předperiody a k ní vypočítanou hodnotu asymptotického kritického exponentu E^* pro danou předperiodu, a to jako zlomek s odmocninou v čitateli i jako desetinné číslo.

- `Vypsat primární i sekundární matice` vypíše v průběhu výpočtu primární i sekundární matice, kde primární, resp. sekundární, matice jsou ty, které obsahují modulované složky Parikhových vektorů návratových slov bispeciálů s příslušnou dvojicí $(h+i, 0)$, resp. $(h+i, m)$ pro $0 < m < z_i$. Dále jsou pod každou matici vypsány příslušné mezivýsledky potřebné pro pro výpočet $\mathrm{E}^*(i, m)$, viz (5.6).

- `Vypsat seznam primárních matic` vypíše matici předperiody a poté seznam prošlých primárních matic, které by se průběžně vypisovaly v předchozím bodě.

Ve všech hladinách výpisu, a to i v případě, že uživatel žádnou z možností nezaškrtne, se vypíše perioda řetězového zlomku θ, délky period slov $\mathbf{y}$ a $\mathbf{y}'$, (minimální) asymptotický kritický exponent a matice předperiody řetězového zlomku θ, pro které je nabýváno minimum (pokud je předperioda zadána, pak je to matice příslušná právě této předperiodě).

Při stisku tlačítka `Spustit` se otevře nové okno, které bude obsahovat výpočet s dříve zvolenou hladinou výpisu. Výpis lze uložit do textového souboru. V hlavním okně přibude pouze výpis zadaných parametrů s výsledkem, i tento výpis lze uložit do souboru. Pokud uživatel chce pokračovat v předešlém zkoumání, může také výpis ze souboru načíst, nebo vymazat.



Obrázek 7.1: Hlavní okno programu pro výpočet asymptotického kritického exponentu

7.1.2 Vlastní výpočet

Zavedeme značení, že pro matici $\mathbb{A} \in \mathbb{N}^{2,2}$ platí

$$\mathbb{A} \mod \begin{pmatrix} \mathrm{Per}(\mathbf{y}) \\ \mathrm{Per}(\mathbf{y}') \end{pmatrix} := \begin{pmatrix} a_{1,1} \mod \mathrm{Per}(\mathbf{y}) & a_{1,2} \mod \mathrm{Per}(\mathbf{y}) \\ a_{2,1} \mod \mathrm{Per}(\mathbf{y}') & a_{2,2} \mod \mathrm{Per}(\mathbf{y}') \end{pmatrix}.$$

Jako první krok se vytvoří třída Computation, které se předají navolené hodnoty z hlavního okna. Při předávání se kontroluje, zda je perioda, příp. předperioda ve správném tvaru (musí to být přirozená čísla oddělená čárkami).

Následně se zavolá funkce `void initialize ()`, která předpočítá hodnoty $\theta_{(i,m)}$ definované v 3.5 pro všechna $i \in \{0, 1, \ldots, M-1\}$ a $m \in \{0, 1, \ldots, z_i - 1\}$. Vzorec pro výpočet nám dává následující věta.

Věta 7.1. *Uvažujme* $\theta_{(0,0)} = [0, (z_0, z_1, \ldots, z_{M-1})^\omega]$. *Nechť dále*

$$\mathbb{A}^{(0)} = \begin{pmatrix} 0 & 1 \\ 1 & z_0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & z_1 \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & z_{M-1} \end{pmatrix},$$

Λ *je dominantní vlastní číslo matice* $\mathbb{A}^{(0)}$ *a* $\vec{\hat{x}} = \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}$ *příslušný vlastní vektor, pak*

$$\theta_{(0,0)} = \frac{\hat{x}}{\hat{y}}.$$

Důkaz. Řádkové vektory (p_{NM-1}, p_{NM}) i (q_{NM-1}, q_{NM}) splňují vztah

$$(X_{NM-1}, X_{NM}) = (X_{-1}, X_0)\left(\mathbb{A}^{(0)}\right)^N.$$

Je-li λ vlastní číslo matice $\mathbb{A}^{(0)}$ takové, že $|\lambda| < 1$ a $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ příslušný vlastní vektor, pak platí

$$\lim_{N\to+\infty} \frac{p_{NM-1}}{p_{NM}} = \lim_{N\to+\infty} \frac{q_{NM-1}}{q_{NM}} = -\frac{y}{x}.$$

Přenásobením rekurentního vztahu vlastním vektorem $\vec{\hat{x}}$ příslušným k Λ získáme

$$\hat{x}X_{NM-1} + \hat{y}X_{NM} = \Lambda^N(\hat{x}X_{-1} + \hat{y}X_0).$$

Úpravou získáme

$$\frac{\hat{x}p_{NM-1} + \hat{y}p_{NM}}{\hat{x}q_{NM-1} + \hat{y}q_{NM}} = \frac{p_{NM}}{q_{NM}} \frac{\hat{x}\frac{p_{NM-1}}{p_{NM}} + \hat{y}}{\hat{x}\frac{q_{NM-1}}{q_{NM}} + \hat{y}} = \frac{\hat{x}p_{-1} + \hat{y}p_0}{\hat{x}q_{-1} + \hat{y}q_0} = \frac{\hat{x}\cdot 1 + \hat{y}\cdot 0}{\hat{x}\cdot 0 + \hat{y}\cdot 1} = \frac{\hat{x}}{\hat{y}}.$$

Následnou limitou $N \to +\infty$ a s využitím skutečnosti, že limita $\lim_{n\to+\infty} \frac{p_n}{q_n}$ existuje a je rovna $\theta_{(0,0)}$, získáme

$$\theta_{(0,0)} = [0, (z_0, z_1, \ldots, z_{M-1})^\omega] = \lim_{N\to+\infty} \frac{p_{NM}}{q_{NM}} = \frac{\hat{x}}{\hat{y}}. \tag{7.1}$$

□

Další hodnoty již vypočítáme jednoduše rekurentně pomocí vzorců

$$\begin{aligned} \theta_{(i,0)} &= [0, (z_i, z_{i+1}, \ldots, z_{M-1}, z_0, \ldots, z_{i-1})^\omega] = \frac{1}{\theta_{(i-1,0)}} - z_{i-1}, \\ \theta_{(i,m)} &= [0, z_i - m, (z_{i+1}, \ldots, z_{M-1}, z_0, \ldots, z_i)^\omega] = \frac{1}{z_i - m + \theta_{(i+1 \mod M,0)}}. \end{aligned} \tag{7.2}$$

Tyto hodnoty se uloží do pole uvnitř třídy, protože je budeme používat opakovaně ve více funkcích.

Dále se vytvoří matice seznamů $\mathcal{S}_1(i,m) \cap \mathcal{S}_3$ pomocí věty 3.7, kde za θ bereme vždy již napočítané hodnoty $\theta_{(i,m)}$.

Dalším krokem je vytvoření seznamu možných matic předperiod listm. Pokud je předperioda zadána, do seznamu se vloží pouze matice tvaru $\begin{pmatrix} p_{h-1} & p_h \\ q_{h-1} & q_h \end{pmatrix} \mod \begin{pmatrix} \mathrm{Per}(\mathbf{y}) \\ \mathrm{Per}(\mathbf{y}') \end{pmatrix}$, jinak se zavolá třída MComputation, která seznam naplní, postup je popsán v části 7.1.2.1.

Následuje vlastní výpočet asymptotického kritického exponentu. Ještě je dobré poukázat na to, že indexování v C++ je od 0. Pseudokód vypadá následovně:

- Vytvoř si proměnnou Min a ulož do ní maximální možnou hodnotu.
- Dokud seznam listm není prázdný, opakuj následující kroky:

- vyber ze seznamu listm matici C a vytvoř si následující proměnné:
  1. seznam prošlých případů examined,
  2. dosud dosažené maximum Max a vlož do něj 0,
  3. index i = 0,
  4. indexy pro dosažené maximum AMaxi, AMaxm, Maxk, Maxl.
- Dokud matice C s indexem i není v seznamu examined, opakuj:
  * pro všechna m = $0, 1, \ldots,$ period[i]−1 proveď tyto kroky:
    · Vytvoř si matici D obsahující Parikhovy vektory návratových slov s a r tak, že první sloupec matice D = první sloupec matice C + m-krát druhý sloupec C a druhý sloupec D = druhý sloupec matice C.
    · Vytvoř si v proměnné Sim kopii seznamu $\mathcal{S}_1(i,m) \cap \mathcal{S}_3$.
    · Dokud je seznam Sim neprázdný, vyjmi ze něj dvojici k,l a zkus, jestli (l∗D[0][0] + k∗D[0][1]) $\mod \mathrm{Per}(\mathbf{y}) = 0$ a (l∗D[1][0] + k∗D[1][1]) $\mod \mathrm{Per}(\mathbf{y}') = 0$. Tento krok je totožný s kontrolou, zda $\begin{pmatrix}\ell\\k\end{pmatrix} \in \mathcal{S}_2(i,m)$.
    Pokud ano, do proměnné z ulož hodnotu $1 + \lim_{t\to+\infty} \mathrm{I}(h+i+tH, m)$, kterou získáš ze vzorce (5.6).
    Pokud je z > Max, do Max ulož z, do AMaxi, resp. AMaxm ulož i, resp. m a nakonec do Maxk, resp. Maxl ulož k, resp. l.
  * Vlož matici C s indexem i do seznamu examined.
  * Do matice C si ulož novou matici E $\mod \begin{pmatrix}\mathrm{Per}(\mathbf{y})\\ \mathrm{Per}(\mathbf{y}')\end{pmatrix}$, kde první sloupec E = druhý sloupec C a druhý sloupec E = period[i]*(druhý sloupec C) + první sloupec C.
  Tímto postupem budou sloupce v nové matici tvořit Parikhovy vektory návratových slov k bispeciálu s N o jedna vyšším a $m = 0$.
  * Do i ulož (i+1) $\mod M$.

  Na konci tohoto cyklu jsme prošli všechny třídy ekvivalence $\sim$ a v Max je hodnota asymptotického kritického exponentu pro třídu slov mající předperiodu, které odpovídá počáteční matice C.
- Pokud je Max < Min, pak do Min ulož Max.

Na konci tohoto cyklu jsme prošli všechny matice předperiod ze seznamu a v Min je uložena hodnota nejmenšího asymptotického kritického exponentu, která byla nalezena.

7.1.2.1 MComputation

Tato třída byla doplněna později a umožňuje projít všechny možné matice předperiod, tzn. matice obsahující Parikhovy vektory návratových slov s a r k bispeciálu s příslušnou dvojicí $(h, 0)$. Tyto matice mají vždy rozměr 2×2. Dále protože jsou sloupce matice využívány k určení množiny $\mathcal{S}_2$, stačí je počítat $\mod \begin{pmatrix}\mathrm{Per}(\mathbf{y})\\ \mathrm{Per}(\mathbf{y}')\end{pmatrix}$, a je jich tudíž vždy maximálně $(\mathrm{Per}(\mathbf{y})\mathrm{Per}(\mathbf{y}'))^2$. Reálně je jich však ještě méně, protože tyto matice musí mít před modulací determinant roven ± 1, což vyplývá z vlastností řetězových zlomků.

V textu i nadále budeme předpokládat, že častější písmeno ve zkoumaném sturmovském slově je $\mathtt{b}$ a řídící posloupnost tudíž začíná morfismem D. Pokud by bylo častější písmeno $\mathtt{a}$, pak by řídící posloupnost začínala morfismem G.

Dále si připomeneme, že pokud $\mathbf{u}$ je standardní sturmovské slovo a $\mathbf{u} = D(\mathbf{u}')$, pak častější písmeno je $\mathtt{b}$, a tudíž prefixové návratové slovo k prázdnému slovu v $\mathbf{u}$ je $r = \mathtt{b}$ a neprefixové návratové slovo je pak $s = \mathtt{a}$. Naopak pro standardní sturmovské slovo $\mathbf{u}'' = G(\mathbf{u}')$ je prefixové návratové slovo k prázdnému slovu v $\mathbf{u}''$ $r'' = \mathtt{a}$ a neprefixové $s'' = \mathtt{b}$. Z tohoto pozorování a věty 3.3 získáme následující vzorec pro generování návratových slov k bispeciálu s příslušnou dvojicí $(N, 0)$.

- N liché $r = D^{a_1}G^{a_2}D^{a_3}\cdots D^{a_N}(\mathtt{a})$,
  $s = D^{a_1}G^{a_2}D^{a_3}\cdots D^{a_N}(\mathtt{b})$,
- N sudé $r = D^{a_1}G^{a_2}D^{a_3}\cdots G^{a_N}(\mathtt{b})$,
  $s = D^{a_1}G^{a_2}D^{a_3}\cdots G^{a_N}(\mathtt{a})$,

kde a_k jsou pro všechna k členy rozvoje řetězového zlomku θ, a tudíž přirozená čísla. Z toho vyplývá následující věta, kterou využijeme v programu.

Věta 7.2. *Nechť* $\mathtt{b}$ *je častější písmeno ve standardním sturmovské slově. Označíme-li* $\mathbb{M}_D = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ *matici morfismu D,* $\mathbb{M}_G = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ *matici morfismu G, r prefixové a s neprefixové návratové slovo k bispeciálu s příslušnou dvojicí* $(N, 0)$*, pak všechny možné matice* $\begin{pmatrix} \vec{V}(s) & \vec{V}(r) \end{pmatrix}$ *jsou ve tvaru*

1. $\mathbb{M}_D^{a_1}\mathbb{M}_G^{a_2}\cdots\mathbb{M}_D^{a_N}\cdot\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *pro N liché a*
2. $\mathbb{M}_D^{a_1}\mathbb{M}_G^{a_2}\cdots\mathbb{M}_G^{a_N}\cdot\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *pro N sudé,*

kde a_k jsou přirozená čísla pro všechna $k \in \{1, 2, \ldots, N\}$.

Nám tedy bude stačit generovat matice předperiod pomocí předchozí věty pro $N = 0, 1, \ldots,$ při každém kroku je počítat $\mod \begin{pmatrix} \mathrm{Per}(\mathbf{y}) \\ \mathrm{Per}(\mathbf{y}') \end{pmatrix}$ a sledovat, zda stále vznikají nové matice. Skutečný postup v programu je popsán v dalším textu.

Popis programu

Na začátku jsou vytvořeny dva prázdné seznamy. Pomocný seznam L1, ve kterém se všechny prvky matice počítají modulo $\mathrm{nsn}(\mathrm{Per}(\mathbf{y}), \mathrm{Per}(\mathbf{y}'))$ a obsahuje mezivýsledky pro součiny matic, a pak seznam výsledných matic res, do kterého se matice ukládají již $\mod \begin{pmatrix} \mathrm{Per}(\mathbf{y}) \\ \mathrm{Per}(\mathbf{y}') \end{pmatrix}$.

Vkládání do seznamů probíhá tak, že se vkládaná matice vypočítá příslušným modulem, projde se seznam, zda v něm matice není, a teprve pokud ne, tak se matice vloží na konec seznamu. Jinak se matice „zahodí“ a dále se neuvažuje.

Program začne s maticí $\mathbb{M}_D$, tu vloží do seznamu L1, do seznamu res je vložena až po vynásobení zprava maticí $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Iterátor seznamu se nastaví na začátek seznamu L1. Následuje postup popsaný tímto pseudokódem:

Dokud iterátor neukazuje na konec seznamu L1, opakuj:

- Zkopíruj si do B matici z pozice, na kterou ukazuje iterátor.
- Vytvoř z ní dvě matice:
  1. B$\cdot\mathbb{M}_G$. Vlož tuto matici do obou seznamů. (Násobení jednotkovou maticí totiž nemá žádný vliv.)
  2. B$\cdot\mathbb{M}_D$. Vlož tuto matici do seznamu L1, prohoď její sloupce a vlož modifikovanou matici do res.

Násobení stejnou maticí (a tudíž aplikace stejného morfismu) jako v předchozím kroku odpovídá zvětšení aktuálního koeficientu řetězového zlomku o 1, násobení tou druhou znamená přidání nového členu na konec předperiody.

- Posuň iterátor v seznamu L1 na další matici.

Vlož do res jednotkovou matici, ta odpovídá prázdné předperiodě.

Na konci tohoto cyklu obsahuje seznam res všechny možné matice předperiod, které mohou v programu pro asymptotický kritický exponent pro zadané Per($\mathbf{y}$) a Per($\mathbf{y}'$) vzniknout.

Program není těžké upravit, aby si kromě matice předperiody pamatoval i tvar předperiody. Každá matice má v tomto případě k sobě připojen seznam obsahující prvky předperiody, tyto prvky se vytváří při generování matice zvětšováním posledního prvku o jedna nebo přidáním 1 jako nového prvku na konec. Z důvodu paměťové náročnosti to ale není ve výsledném programu implementováno.

7.1.3 Výstup programu pro výpočet asymptotického kritického exponentu

Výstup programu probíhá tak, jak bylo popsáno výše. V jednom případě však může být vyvolána výjimka a program tudíž nemůže vypsat hodnotu asymptotického kritického exponentu. Místo toho je vypsána chybová hláška „Přetečení při výpočtu theta“ a do hlavního okna se pod parametry připíše Overflow in computation of Li_theta. Nastává to v případě, kdy uživatel zadá velmi dlouho periodu, a proto při výpočtu θ program musí počítat s velkými čísly. Při výpočtu se počítá determinant s prvky p_n a q_n, které již samy o sobě velmi rychle rostou, natož když se mezi sebou násobí. V budoucnosti se pokusíme program buď doplnit o třídu pracující s velkými celými čísly, nebo optimalizovat výpočet θ jiným způsobem. Pro námi zkoumaná slova však tato chyba nenastává.

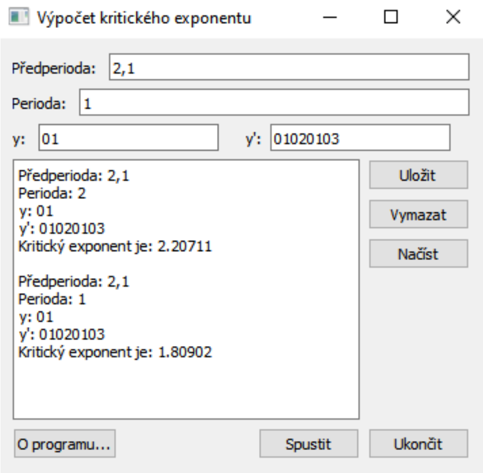
Příklad 7.1. Výstup programu pro výpočet asymptotického kritického exponentu stejného slova jako v příkladu 5.1 (tudíž byla zadána i předperioda). Jeho parametry jsou

$$\theta = [0, 1, 1, 1, (3, 1)^\omega], \quad \mathbf{y} = (01)^\omega, \quad \mathbf{y}' = (234235)^\omega$$

```
Předperioda:
1 0
2 3

(i,m) = (0 , 0)
1 0
2 3
z (k,l) = 1.89564 (2,0)
z (k,l) = 1.44782 (4,0)

(i,m) = (0 , 1)
1 0
5 3
z (k,l) = 2.39564 (2,0)

(i,m) = (0 , 2)
1 0
2 3
z (k,l) = 2.89564 (2,0)

(i,m) = (1 , 0)
0 1
3 5
z (k,l) = 1.17913 (6,4)
z (k,l) = 1.16667 (6,6)

(i,m) = (2 , 0)
1 1
5 2
z (k,l) = 1.32087 (4,2)
z (k,l) = 1.15465 (10,2)

(i,m) = (2 , 1)
0 1
1 2
z (k,l) = 1.5 (2,2)
z (k,l) = 1.24099 (8,2)

(i,m) = (2 , 2)
1 1
3 2
z (k,l) = 1.32733 (6,2)
z (k,l) = 1.22087 (6,4)

(i,m) = (3 , 0)
1 0
2 5
z (k,l) = 1.27913 (4,2)
z (k,l) = 1.16667 (6,6)
(i,m) = (4 , 0)
0 1
5 1
z (k,l) = 1.18693 (8,2)

(i,m) = (4 , 1)
1 1
0 1
z (k,l) = 1.29129 (6,2)

(i,m) = (4 , 2)
0 1
1 1
z (k,l) = 1.39564 (4,2)
z (k,l) = 1.19782 (8,4)

(i,m) = (5 , 0)
1 1
1 2
z (k,l) = 1.5 (2,2)
z (k,l) = 1.25 (4,4)
z (k,l) = 1.16667 (6,6)
z = [7 + 1 sqrt(21)]/4
Maximum: 2.89564, (0, 2)

Prozkoumané případy:
1, 0
2, 3   i = 0
0, 1
3, 5   i = 1
1, 1
5, 2   i = 0
1, 0
2, 5   i = 1
0, 1
5, 1   i = 0
1, 1
1, 2   i = 1


Pro periodu 3,1 :
Per(y) = 2
Per(y') = 6
Asymptotický kritický exponent je: 2.89564
Pro předperiodu:
1 0
2 3
```

7.2 Program pro výpočet kritického exponentu

Program pro výpočet kritického exponentu umožňuje zadat předperiodu a periodu řetězového zlomku θ, opět jako přirozená čísla oddělená mezerami, a slova s konstantními mezerami $\mathbf{y}$ a $\mathbf{y}'$ jako posloupnosti znaků (písmena anglické abecedy a číslice) bez mezer.



Obrázek 7.2: Hlavní okno programu pro výpočet kritického exponentu

Po stisknutí tlačítka Spustit je vytvořena třída Computation, která vytvoří vlastní okno a kroky mezivýpočtu jsou vypisovány do tohoto okna. Do hlavního okna jsou pak vypsány pouze parametry a výsledek. Opět je umožněno výpisy uložit, výpis v hlavním okně načíst ze souboru, nebo smazat.

Vlastní výpočet je podobně jako v kapitole 6 rozdělen do tří hlavních částí, a to generování a procházení prefixu, výpočet asymptotického kritického exponentu a procházení dlouhých bispeciálů.

7.2.1 Prefix

Pro výpočet kritického exponentu je důležité projít množinu krátkých faktorů, kterou jsme značili $\mathcal{W}^{short}$. Podíváme-li se na postup výpočtu popsaný v části 6.1, můžeme vidět, že nepotřebujeme pracovat s balancovaným slovem, ale pouze s faktory standardního sturmovského slova. Abychom našli správný tvar všech faktorů a hlavně abychom dokázali určit nejkratší bispeciál, který je obsahuje, potřebujeme však vygenerovat dostatečně dlouhý prefix sturmovského slova.

Prvním krokem je vygenerování seznamu mezer pro slova s konstantními mezerami $\mathrm{gap}(\mathbf{y}, n)$ a $\mathrm{gap}(\mathbf{y}', n')$ pro $n \in \{0, 1, \ldots, \mathrm{Per}(\mathbf{y})\}$, $n' \in \{0, 1, \ldots, \mathrm{Per}(\mathbf{y}')\}$ a nastavení $\beta(\mathbf{y})$ a $\beta(\mathbf{y}')$. Do seznamu je ukládána vždy dvojice (nalezená mezera, délka faktoru).

Následujícím krokem výpočtu je určit, jak dlouhý bispeciál musíme vygenerovat (bispeciály jsou totiž prefixy standardního sturmovského slova a generovat je umíme dle věty 3.3), abychom v něm nalezli všechny krátké faktory. K tomu nám pomůže následující věta.

Věta 7.3. [10] *Nechť $\ell \in [Q_K, Q_{K+1}) \cap \mathbb{N}$, pak prefix slova $\mathbf{u}$ délky $\mathrm{R}_{\mathbf{u}}(\ell) = Q_{K+1} + Q_K + \ell - 1$ obsahuje všechny faktory délky ℓ.*

Lemma 7.1. *Nechť $N \in \mathbb{N}$ je nejmenší přirozené číslo takové, že $p_N > \beta(\mathbf{y}) + 1$ a $q_N > \beta(\mathbf{y}') + 1$. Dále nechť $n = Q_N = p_N + q_N$. Pak každý faktor $u \in \mathcal{L}(\mathbf{u})$ takový, že $|u| \geq n$, splňuje $|u|_{\mathsf{a}} > \beta(\mathbf{y})$ a $|u|_{\mathsf{b}} > \beta(\mathbf{y}')$.*

Důkaz. Ve slově se jistě vyskytuje faktor r s Parikhovým vektorem $\vec{V}(r) = \begin{pmatrix} p_N \\ q_N \end{pmatrix}$, je to například častější návratové slovo k bispeciálu s dvojicí $(N, 0)$, viz věta 3.5. Z balancovanosti pak vyplývá, že pro jakýkoliv faktor $v \in \mathcal{L}(\mathbf{u})$, $|v| = n$ platí

$$\begin{aligned} ||r|_\mathsf{a} - |v|_\mathsf{a}| &= |p_N - |v|_\mathsf{a}| \le 1 \Rightarrow |v|_\mathsf{a} \ge p_N - 1 > \beta(\mathbf{y})\,, \\ ||r|_\mathsf{b} - |v|_\mathsf{b}| &= |q_N - |v|_\mathsf{b}| \le 1 \Rightarrow |v|_\mathsf{b} \ge q_N - 1 > \beta(\mathbf{y}')\,. \end{aligned}$$

Nakonec si uvědomíme, že každý faktor délky větší než n vznikl prodloužením faktoru délky n, a proto se počet výskytů písmena a ani b nemohl snížit. □

Chceme-li tedy generovat prefix, který obsahuje všechny krátké faktory, tj. projekce množiny W^{short}, pak podle lemmatu 7.1 stačí takový, který obsahuje všechny faktory kratší než Q_N, tudíž podle věty 7.3 délky

$$R_\mathbf{u}(Q_N - 1) = Q_N + Q_{N-1} + Q_N - 1 - 1 = 2Q_N + Q_{N-1} - 2.$$

Tento prefix je jistě obsažen v bispeciálu délky $Q_{N+1} + Q_N - 2$. Tudíž budeme generovat bispeciál, kterému přísluší dvojice $(N + 1, 0)$.

Při hledání N rovnou najdeme nejmenší i takové, že $p_i + p_{i-1} > \beta(\mathbf{y}) + 1$ a $q_i + q_{i-1} > \beta(\mathbf{y}') + 1$. Takto nalezneme nejkratší bispeciál s příslušnou dvojicí $(i, 0)$ tak, že obsahuje „dost“ písmen a a b, viz věta 3.5. Pokud $i > h$, přenastavíme předperiodu tak, aby $h = i$ (předperiodu doplníme $i - h$ prvky z periody a periodu cyklicky posuneme o příslušný index). Tím budeme mít zajištěno, že všechny krátké bispeciály přísluší předperiodě a pro dlouhé bispeciály vždy bude platit vzorec (5.1) pro $\mathcal{S}_2$.

Dále krok, který je zatím v programu doplněn, ale není optimální, je, že pokud $N + 1 < h$, pak se $N + 1$ položí rovno h. Tato úprava umožní stejným kódem projít všechny bispeciály příslušné předperiodě najednou a nepotřebujeme dělení na $\mathcal{W}^{med}$, viz část 6.2, pro dlouhé předperiody přibude ale prohledávání prefixu navíc.

Nyní již máme správně dlouhou předperiodu, známe index $N + 1$, a máme napočítané členy posloupnosti p_n a q_n pro n od -1 do N+1. Můžeme proto již napočítat hodnoty $\theta_{(i,0)}$ příslušné periodě pro $i \in \{0, 1, \ldots, M - 1\}$ dle vzorců (7.1) a (7.2) a $\hat{\theta}_{(k,0)}$ příslušné předperiodě pro $k \in \{1, 2, \ldots, h\}$ dle vzorců

$$\begin{aligned} \hat{\theta}_{(h,0)} &= \frac{1}{d_h + \theta_{(0,0)}}, \\ \hat{\theta}_{(j,0)} &= \frac{1}{d_j + \hat{\theta}_{(j+1,0)}}, \ j < h. \end{aligned}$$

Nyní již nepočítáme dopředu hodnoty $\theta_{(i,m)}$ pro $m > 0$ tak, jak tomu bylo v programu pro asymptotický kritický exponent, protože ty jsou v programu potřeba pouze jednou a napočítají se přímo při vytváření seznamu Sim.

Nyní již máme všechno připraveno pro to, abychom vygenerovali příslušný bispeciál pomocí věty 3.3 a příkladu 4.1. Zde vždy předpokládáme, že *častější písmeno ve sturmovském slově* je b. Pokud by uživatel chtěl vypočítat kritický exponent v případě, že častější písmeno je a, stačí vyměnit $\mathbf{y}$ a $\mathbf{y}'$.

Dalším krokem je procházení prefixu. Program ho prochází následovně:

Polož z = 0. Do proměnné PrefixLengthIndex si ulož hodnotu N+1 vypočítanou v předchozím kroku.

- Pro $N = 0, 1, \ldots,$ PrefixLengthIndex opakuj následující kroky:
  - Pokud $N = 0$, pak $n = 1$, jinak $n = 0$. Pak pro $m = n, \ldots, a_{N+1} - 1$, kde a_N je N-tý prvek řetězového zlomku θ, opakuj:

- Vypočítej si délku bispeciálu, kterému přísluší (N, m) dle vzorce 3.5 a ulož to do proměnné end. Dále vypočítej délku předchozího bispeciálu a tu ulož do start .
- Připrav si seznam $\mathcal{S}_1(N, m) \cap \mathcal{S}_3$ příslušný bispeciálu (N, m) pomocí věty 3.7 a předpočítaného $\hat{\theta}_{(N,m)}$, pokud $N < h$, resp. $\theta_{(N-h \mod M, m)}$ pro bispeciál příslušný periodě, a ulož ho do proměnné kl. To můžeme udělat, protože derivované slovo je stejné pro bispeciál i pro všechny faktory, které se na tento bispeciál natáhnou. Lišit se budou množinami $\mathcal{S}_2$.
- Projdi všechny faktory u, které začínají na pozici i $\leq$ end a končí na pozici j$\in$ (start , end]. Pokud se faktor vyskytl v prefixu poprvé a $|u|_{\mathsf{a}} \leq \beta(\mathbf{y})$ nebo $|u|_{\mathsf{b}} \leq \beta(\mathbf{y}')$ nebo je to bispeciál pokračuj, jinak vyber další faktor. Pokud faktory došly, vyber další m, resp. N.
- Zkopíruj si do kl1 seznam kl a vytvoř si seznam všech možných dvojic (n, n'), kde $n \in \text{gap}(\mathbf{y}, |u|_{\mathsf{a}})$ a $n' \in \text{gap}(\mathbf{y}', |u|_{\mathsf{b}})$.
- Do proměnné Min ulož maximální možnou hodnotu daného datového typu.
- Dokud není seznam kl1 prázdný, opakuj:
  - Vyjmi dvojici k, ℓ ze seznamu a zkus, zda pro nějakou dvojici (n, n') platí

    $$((k + \ell m)p_N + \ell p_{N-1}) \mod n = 0 \text{ a } ((k + \ell m)q_N + \ell q_{N-1}) \mod n' = 0,$$

    pokud ano, pak $\begin{pmatrix}\ell \\ k\end{pmatrix} \in \mathcal{S}_2(u)$ a vypočítej délku faktoru (v případě minimality to bude návratové slovo) $|v| = (k + \ell m)Q_N + \ell Q_{N-1}$, viz věta 4.4.
  - Zkus, zda $|v| <$ Min. Pokud ano, do Min ulož $|v|$.
  - Zkus, zda z $< \frac{|u|}{\text{Min}}$. Pokud ano, do z ulož $\frac{|u|}{\text{Min}}$.

Na konci tohoto cyklu proměnná z obsahuje $\mathrm{E}^{short}(\mathbf{v})$.

Jak můžeme vidět, v kódu máme podmínku, že se vyšetří všechny bispeciály obsažené v prefixu, nejen ty krátké a středně dlouhé. Toto je v kódu přidáno, protože už máme předpočítané seznamy kl a návratová slova kvůli faktorům, které se na tento bispeciál prodlouží. Tyto bispeciály se pak nebudou vyšetřovat podruhé v rámci dlouhých bispeciálů ani v případě, že do této skupiny patří.

7.2.2 Asymptotický kritický exponent

Výpočet probíhá stejným způsobem jako v programu pro výpočet asymptotického kritického exponentu. Program pouze nenapočítává znovu p_h a q_h a místo toho využije již předpočítaných hodnot. Program také nastaví H na počet tříd ekvivalence $\sim$.

7.2.3 Dlouhé bispeciály

Program si vytvoří matici $\mathbb{A}^{(0)} = \begin{pmatrix}0 & 1 \\ 1 & z_0\end{pmatrix} \cdots \begin{pmatrix}0 & 1 \\ 1 & z_{M-1}\end{pmatrix}$ a vypočítá její dominantní vlastní číslo. Stačí nám vlastní číslo napočítat pouze pro tuto matici, protože všechny matice $\mathbb{A}^{(i)}$ jsou si podobné, a mají tudíž stejná vlastní čísla. Dále prochází $i \in \{0, 1, \ldots, H-1\}$ a pro každé i napočítá nejmenší možné $N^{(i)}$ dle vzorce (6.2).

Pak již stačí projít všechny bispeciály s příslušnými (N, m), kde $N = h + i + tH$, $i \in \{0, 1, \ldots, H-1\}$, $t \in \{0, 1, \ldots, N^{(i)} - 1\}$ a $m \in \{0, 1, \ldots, z_{i \mod M} - 1\}$ standardním postupem popsaným následujícím pseudokódem.

1. Načti již předpočítaný seznam kl pro dané (i, m) (seznam je stejný jako pro výpočet asymptotického kritického exponentu) do proměnné kl1.

2. Vytvoř proměnnou Max a ulož do ní 0.
3. Nalezni Parikhovy vektory návratových slov dle věty 3.5.
4. Dokud není seznam kl1 prázdný, opakuj
   - vyjmi ze seznamu dvojici $\binom{\ell}{k}$ a zkus, zda $\binom{\ell}{k} \in \mathcal{S}_2(i, m)$ dle vzorce (5.1),
   - pokud ano, pomocí vzorce (5.2) vypočítej hodnotu z $= 1 + \frac{|w|}{|v|}$ a zkus, zda z $>$ Max. Pokud ano, do Max ulož z.

Nakonec je nutno vybrat maximum z hodnot $\mathrm{E}^{short}(\mathbf{v})$, $\mathrm{E}^{*}(\mathbf{v})$ a $\mathrm{E}^{long}(\mathbf{v})$, čímž dostaneme kritický exponent balancovaného slova **v**.

7.2.4 Výstup programu pro výpočet kritického exponentu

Výstup programu byl podrobně popsán výše. I zde se však počítají hodnoty θ, může proto dojít ke stejné chybě jako v programu pro asymptotický kritický exponent. V tom případě vypadá výpis stejným způsobem. Dalším místem, kde může dojít k přetečení a výpočet proto nemůže být dokončen, je výpočet dlouhých bispeciálů. Ve chvíli, kdy H je vysoké číslo, mohou členy posloupnosti Q_N přesáhnout maximální hodnotu pro daný datový typ. I v tomto případě je uživatel upozorněn chybovou hláškou a v hlavním okně je napsán důvod nedokončení výpočtu.

Příklad 7.2. Výstup programu pro výpočet kritického exponentu stejného slova jako v příkladu 5.1.

```
pn : 1,0,1,1,2,7,9
qn : 0,1,1,2,3,11,14
Parametry:
předperioda : 1,1,1
perioda : 3,1
y: 01
y': 234235
Prefix:
babbababbababbababbabbababbababbababbab
Délka prefixu = 39
Prefix index = 5
Gaps y:
(1,0)
(2,1)
(2,2)
Délka nejdelšího bispeciálu: 0
Gaps y':
(1,0)
(3,1)
(6,1)
(3,2)
(6,2)
(6,3)
(6,4)
(6,5)
(6,6)
Délka nejdelšího bispeciálu: 2

Krátké faktory a bispeciály, které je obsahují:
```

```
(N,m) = 1,0
b
Seznam mezer:
(1,3)
(1,6)
|v| = 4, (k,l) = (1,2)
|v| = 5, (k,l) = (2,1)
|v| = 9, (k,l) = (3,3)
|v| = 10, (k,l) = (4,2)
|v| = 14, (k,l) = (5,4)
|v| = 15, (k,l) = (6,3)
|v| = 19, (k,l) = (7,5)
|v| = 20, (k,l) = (8,4)
Minimální délka návratového slova = 4;
(k,l) = (1, 2)
1+(|w|/|v|) = 1.25

(N,m) = 2,0
ba
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,1)
|v| = 10, (k,l) = (2,2)
|v| = 15, (k,l) = (3,3)
|v| = 20, (k,l) = (4,4)
|v| = 25, (k,l) = (5,5)
|v| = 30, (k,l) = (6,6)
Minimální délka návratového slova = 5;
(k,l) = (1, 1)
1+(|w|/|v|) = 1.4
```

```
(N,m) = 2,0
bab
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,1)
|v| = 10, (k,l) = (2,2)
|v| = 15, (k,l) = (3,3)
|v| = 20, (k,l) = (4,4)
|v| = 25, (k,l) = (5,5)
|v| = 30, (k,l) = (6,6)
Minimální délka návratového slova = 5;
(k,l) = (1, 1)
1+(|w|/|v|) = 1.6

(N,m) = 2,0
a
Seznam mezer:
(2,1)
|v| = 5, (k,l) = (1,1)
|v| = 6, (k,l) = (2,0)
|v| = 10, (k,l) = (2,2)
|v| = 11, (k,l) = (3,1)
|v| = 15, (k,l) = (3,3)
|v| = 16, (k,l) = (4,2)
|v| = 20, (k,l) = (4,4)
|v| = 21, (k,l) = (5,3)
|v| = 25, (k,l) = (5,5)
|v| = 26, (k,l) = (6,4)
|v| = 30, (k,l) = (6,6)
|v| = 31, (k,l) = (7,5)
Minimální délka návratového slova = 5;
(k,l) = (1, 1)
1+(|w|/|v|) = 1.2

(N,m) = 2,0
ab
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,1)
|v| = 10, (k,l) = (2,2)
|v| = 15, (k,l) = (3,3)
|v| = 20, (k,l) = (4,4)
|v| = 25, (k,l) = (5,5)
|v| = 30, (k,l) = (6,6)
Minimální délka návratového slova = 5;
(k,l) = (1, 1)
1+(|w|/|v|) = 1.4

(N,m) = 3,0
babbab
Seznam mezer:
(2,6)
|v| = 10, (k,l) = (2,0)
|v| = 20, (k,l) = (4,0)
Minimální délka návratového slova = 10;
(k,l) = (2, 0)
1+(|w|/|v|) = 1.6

(N,m) = 3,0
abb
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,0)
|v| = 10, (k,l) = (2,0)
|v| = 15, (k,l) = (3,0)
|v| = 20, (k,l) = (4,0)
Minimální délka návratového slova = 5;
(k,l) = (1, 0)
1+(|w|/|v|) = 1.6

(N,m) = 3,0
abba
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,0)
|v| = 10, (k,l) = (2,0)
|v| = 15, (k,l) = (3,0)
|v| = 20, (k,l) = (4,0)
Minimální délka návratového slova = 5;
(k,l) = (1, 0)
1+(|w|/|v|) = 1.8

(N,m) = 3,0
bb
Seznam mezer:
(1,3)
(1,6)
|v| = 5, (k,l) = (1,0)
|v| = 10, (k,l) = (2,0)
|v| = 15, (k,l) = (3,0)
|v| = 20, (k,l) = (4,0)
|v| = 44, (k,l) = (7,3)
|v| = 49, (k,l) = (8,3)
|v| = 54, (k,l) = (9,3)
Minimální délka návratového slova = 5;
(k,l) = (1, 0)
1+(|w|/|v|) = 1.4

(N,m) = 3,0
bba
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,0)
|v| = 10, (k,l) = (2,0)
|v| = 15, (k,l) = (3,0)
|v| = 20, (k,l) = (4,0)
Minimální délka návratového slova = 5;
(k,l) = (1, 0)
1+(|w|/|v|) = 1.6
```

```
(N,m) = 3,1
babbababbab
Seznam mezer:
(2,6)
|v| = 10, (k,l) = (2,0)
Minimální délka návratového slova = 10;
(k,l) = (2, 0)
1+(|w|/|v|) = 2.1

(N,m) = 3,1
baba
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,0)
|v| = 10, (k,l) = (2,0)
|v| = 15, (k,l) = (3,0)
Minimální délka návratového slova = 5;
(k,l) = (1, 0)
1+(|w|/|v|) = 1.8

(N,m) = 3,1
aba
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,0)
|v| = 10, (k,l) = (2,0)
|v| = 15, (k,l) = (3,0)
Minimální délka návratového slova = 5;
(k,l) = (1, 0)
1+(|w|/|v|) = 1.6

(N,m) = 3,1
abab
Seznam mezer:
(2,3)
(2,6)
|v| = 5, (k,l) = (1,0)
|v| = 10, (k,l) = (2,0)
|v| = 15, (k,l) = (3,0)
Minimální délka návratového slova = 5;
(k,l) = (1, 0)
1+(|w|/|v|) = 1.8

(N,m) = 3,2
babbababbababbab
Seznam mezer:
(2,6)
|v| = 10, (k,l) = (2,0)
Minimální délka návratového slova = 10;
(k,l) = (2, 0)
1+(|w|/|v|) = 2.6

(N,m) = 4,0
babbababbababbababbab
Seznam mezer:
(2,6)
|v| = 128, (k,l) = (6,4)
|v| = 138, (k,l) = (6,6)
Minimální délka návratového slova = 128;
(k,l) = (6, 4)
1+(|w|/|v|) = 1.16406

(N,m) = 5,0
babbababbababbababbabbababbababbababbab
Seznam mezer:
(2,6)
|v| = 128, (k,l) = (4,2)
|v| = 266, (k,l) = (10,2)
Minimální délka návratového slova = 128;
(k,l) = (4, 2)
1+(|w|/|v|) = 1.30469
Maximum pro krátké faktory a jejich
bispeciály je: 2.6

Asymptotický kritický exponent:

(i,m) = ( 0 , 0)
1 0
2 3
z (k,l) = 1.89564 (2,0)
z (k,l) = 1.44782 (4,0)

(i,m) = ( 0 , 1)
1 0
5 3
z (k,l) = 2.39564 (2,0)

(i,m) = ( 0 , 2)
1 0
2 3
z (k,l) = 2.89564 (2,0)

(i,m) = ( 1 , 0)
0 1
3 5
z (k,l) = 1.17913 (6,4)
z (k,l) = 1.16667 (6,6)

(i,m) = ( 2 , 0)
1 1
5 2
z (k,l) = 1.32087 (4,2)
z (k,l) = 1.15465 (10,2)

(i,m) = ( 2 , 1)
0 1
1 2
z (k,l) = 1.5 (2,2)
z (k,l) = 1.24099 (8,2)
```

```
(i,m) = ( 2 , 2)
1 1
3 2
z (k,l) = 1.32733 (6,2)
z (k,l) = 1.22087 (6,4)

(i,m) = ( 3 , 0)
1 0
2 5
z (k,l) = 1.27913 (4,2)
z (k,l) = 1.16667 (6,6)

(i,m) = ( 4 , 0)
0 1
5 1
z (k,l) = 1.18693 (8,2)

(i,m) = ( 4 , 1)
1 1
0 1
z (k,l) = 1.29129 (6,2)

(i,m) = ( 4 , 2)
0 1
1 1
z (k,l) = 1.39564 (4,2)
z (k,l) = 1.19782 (8,4)

(i,m) = ( 5 , 0)
1 1
1 2
z (k,l) = 1.5 (2,2)
z (k,l) = 1.25 (4,4)
z (k,l) = 1.16667 (6,6)
[7 + 1 sqrt(21)]/4
Asymptotický kritický exponent: 2.89564,
(i,m) = (0, 2)

Dlouhé bispeciály:
A:
1 1
3 4
Lambda = 4.79129
i = 0, N0 = 0
i = 1, N0 = 0
i = 2, N0 = 0
i = 3, N0 = 0
i = 4, N0 = 0
i = 5, N0 = 0
Max = 0, (k,l) = (0,0)
Maximum pro dlouhé bispeciály bez
asymptotického kritického exponentu: 0

Parametry:
předperioda : 1,1,1
perioda : 3,1
y: 01
y': 234235
Kritický exponent je: 2.89564
```

Kapitola 8

Minimální kritický exponent

Kritický exponent odráží počet opakování faktorů hned po sobě ve slově. Přirozeně se tedy nabízí otázka, zda lze nalézt, popř. jak vypadá, slovo s nejmenším množstvím takovýchto opakování? Pro obecná nekonečná slova je tato otázka zodpovězena, obecnou mez udává následující věta.

Věta 8.1. *(Dejean's theorem) Nechť $d \in \mathbb{N}, d \geq 2$. Pak infimum z kritických exponentů nekonečných slov nad d-písmennou abecedou je rovno*

- *2 pro $d = 2$,*
- *$\frac{7}{4}$ pro $d = 3$,*
- *$\frac{7}{5}$ pro $d = 4$,*
- *$\frac{d}{d-1}$ pro $d \geq 5$.*

Množina balancovaných slov je podmnožinou všech nekonečných slov nad danou abecedou, minimální kritický exponent pro tuto skupinu slov tedy může být větší. Naše nejnovější poznatky opravdu ukazují, že minimální kritický exponent pro balancovaná slova se od toho obecného liší.

Pro $d = 1$ existuje pouze jedno nekonečné slovo, které je periodické. Pro $d = 2$ splývají aperiodická rekurentní balancovaná slova s těmi sturmovskými, která jsou již velmi dobře prozkoumaná. Další výzkum se proto věnoval balancovaným slovům nad abecedami o alespoň třech písmenech.

Touto tematikou se zabývali N. Rampersad, J. Shallit a É. Vandomme [17]. Speciálně vyslovili následující domněnku.

Domněnka. [17] Nechť $d \in \mathbb{N}, d \geq 3$. Minimální kritický exponent balancovaných slov nad abecedou o d písmenech je roven

- $2 + \frac{\sqrt{2}}{2}$ pro $d = 3$,
- $\frac{5+\sqrt{5}}{4}$ pro $d = 4$,
- $\frac{d-2}{d-3}$ pro $d \geq 5$.

Autoři dále dokázali, že pro $d \leq 10$ kritický exponent nemůže být menší a poskytli seznam balancovaných slov, na kterých by se minimální kritický exponent mohl nabývat. Jejich domněnku pro $5 \leq d \leq 8$ potvrdili A. R. Baranwal a J. Shallit [1, 2].

Autoři článků měli ke zkoumání kritického exponentu odlišný přístup, neměli totiž k dispozici algoritmus pro výpočet, a proto využívali pro hledání kandidátů extenzivní prohledávání konečných prefixů balancovaných slov. Tato technika však pro abecedu o 9 a více písmenech narazila na limity výpočetní techniky.

Pomocí našeho programu jsme nejdříve ověřili, že navrhovaná slova pro 9 a 10 písmen opravdu nabývají dolní meze $\frac{7}{6}$, resp. $\frac{8}{7}$. Další zkoumání však vyžadovalo trochu odlišný přístup, protože zkoumaná množina slov již byla rozsáhlá. Náš program totiž umí vypočítat kritický exponent jednoho daného balancovaného slova, možných tvarů řetězového zlomku θ a slov $\mathbf{y}$ a $\mathbf{y}'$ je ale mnohem více.

Program pro asymptotický kritický exponent nám však pomohl pro danou periodu θ a délky period slov $\mathbf{y}$ a $\mathbf{y}'$ omezit možné předperiody θ. S využitím znalostí, jak tvar slov $\mathbf{y}$ a $\mathbf{y}'$, především hodnoty $\beta(\mathbf{y})$ a $\beta(\mathbf{y}')$, ovlivňují krátké faktory, jsme dokázali omezit i tuto množinu. Toto vše dopomohlo tomu, že jsme dokázali najít slova nad 11- a 12-písmennými abecedami, která nabývají kritický exponent nižší, než je udávaná mez, a to $\frac{10}{9} < \frac{9}{8}$, resp. $\frac{11}{10} < \frac{10}{9}$. Domněnka tedy byla vyvrácena.

S využitím výpisu programu se nám dokonce podařilo najít obecný předpis pro slova nad abecedou o sudém počtu písmen (pro $d \geq 12$), která nabývají právě mez $\frac{d-1}{d-2} < \frac{d-2}{d-3}$. Domněnku se nám tedy podařilo nejen vyvrátit, ale dokonce vyslovit a z velké části dokázat novou.

Domněnka. Nechť $d \in \mathbb{N}, d \geq 11$. Minimální kritický exponent balancovaných slov nad d-písmennou abecedou je roven $\frac{d-1}{d-2}$.

V článku [9] se nám povedlo ve spolupráci s A. M. Shurem dokázat, že $\frac{d-1}{d-2}$ je dolní mez pro kritický exponent balancovaných slov nad d-písmennou abecedou pro $d \geq 11$. Dále jsme nalezli obecný předpis pro abecedu o sudém počtu písmen, $d \geq 14$, na kterých je dolní mez nabývána. Nyní zbývá ukázat, že je tato mez nabývána pro liché abecedy o více než 12 písmenech, nebo domněnku vyvrátit a najít nějakou vyšší mez.

Takto vypadá kompletní tabulka slov dosahujících minimálního kritického exponentu nalezených v [17] doplněná o naše výsledky (v tabulce červeně) pro $d = 11, 12$ a $d = 2\delta$, $\delta \geq 7$ a asymptotický kritický exponent pro $d \geq 5$.

Tabulka 8.1: Balancovaná slova nad d-písmennou abecedou pro $d \leq 12$, resp. $d = 2\delta$, kde $\delta \geq 7$, která nabývají minimálního kritického exponentu. Červeně jsou označeny naše výsledky. Dále $\tau = \frac{1+\sqrt{5}}{2}$.

d	θ	$\mathbf{y}$	$\mathbf{y}'$	$\mathrm{E}(\mathbf{v})$	$\mathrm{E}^*(\mathbf{v})$
3	$[0, 1, (2)^\omega]$	$(2)^\omega$	$(01)^\omega$	$2 + \frac{\sqrt{2}}{2}$	$2 + \frac{\sqrt{2}}{2}$
4	$[0, (1)^\omega]$	$(01)^\omega$	$(23)^\omega$	$\frac{5+\sqrt{5}}{4}$	$\frac{5+\sqrt{5}}{4}$
5	$[0, 1, (2)^\omega]$	$(01)^\omega$	$(3435)^\omega$	$\frac{3}{2}$	$\frac{3}{2}$
6	$[0, 2, 1, 1, (1, 1, 1, 2)^\omega]$	$(0)^\omega$	$(123415321435)^\omega$	$\frac{4}{3}$	$\frac{4}{3}$
7	$[0, 1, 3, (1, 2, 1)^\omega]$	$(01)^\omega$	$(234526432546)^\omega$	$\frac{5}{4}$	$\frac{5}{4}$
8	$[0, 3, 1, (2)^\omega]$	$(01)^\omega$	$(23452673254623752643257 6)^\omega$	$\frac{6}{5}$	$\frac{12+3\sqrt{2}}{14}$
9	$[0, 2, 3, (2)^\omega]$	$(01)^\omega$	$(2345672843652748632547 68)^\omega$	$\frac{7}{6}$	$\frac{13+2\sqrt{2}}{14}$
10	$[0, 4, 2, (3)^\omega]$	$(01)^\omega$	$(234567284963254768294365274869)^\omega$	$\frac{8}{7}$	$1 + \frac{\sqrt{13}}{26}$
11	$[0, 5, 1, (1, 1, 1, 2)^\omega]$	$(AB)^\omega$	$(123456781932547618395274163859721436 5879)^\omega$	$\frac{10}{9}$	$\frac{5(7+\sqrt{6})}{43}$
12	$[0, 1, 3, (2)^\omega]$	$(012345)^\omega$	$(6789AB)^\omega$	$\frac{11}{10}$	$\frac{8-\sqrt{2}}{6}$
2δ, $\delta \geq 7$	$[0, 1, \lfloor \delta/2 \rfloor, (1)^\omega]$	$(12345\ldots\delta)^\omega$	$(1'2'3'4'5'\ldots\delta')^\omega$	$\frac{2\delta-1}{2\delta-2}$	$1 + \frac{1}{\delta \cdot \tau^{N-1}}$, kde $\tau^{N+1} < \delta < \tau^{N+2}$

8.1 Minimální asymptotický kritický exponent

Ačkoliv otázkou minimálního kritického exponentu se zabývalo mnoho lidí, o asymptotickém kritickém exponentu toho není moc známo. Mimo jiné není známa ani obecná mez. V našem algoritmu pro výpočet kritického exponentu se však ukazuje, že znalost asymptotického kritického exponentu je klíčová. I při hledání slov s minimálním kritickým exponentem se velmi rychle ukázalo, že je potřeba napřed nalézt kandidáty s nízkým asymptotickým kritickým exponentem. Asymptotický kritický exponent se počítá jednodušeji než ten kritický, nepotřebujeme pro něj znát konkrétní tvar slov $\mathbf{y}$ a $\mathbf{y}'$ a již máme nástroje, jak procházet mnoho různých předperiod najednou a najít ty, které nám vyhovují.

Stávající program pro asymptotický kritický exponent nám byl tedy velkou oporou již při zkoumání minimálního kritického exponentu. Jak můžeme vidět v tabulce 8.2, pro abecedy o 8 a více písmenech je minimální asymptotický kritický exponent jistě nižší, než ten kritický. Program nám umožnil ukázat, že pro abecedy nad 3, 4 a 5 písmeny minimální kritický a minimální asymptotický kritický exponent splývají, ale už pro 6-písmennou abecedu je minimální asymptotický kritický exponent roven $\frac{5}{4}$, což je méně než minimální kritický exponent, který má hodnotu $\frac{4}{3}$.

Pro další hledání minimálního asymptotického kritického exponentu jsme zvolili odlišný přístup, a to skrze orientované grafy. Program zatím stále upravujeme a doplňujeme, abychom dokázali pracovat s čím dál většími abecedami, jeho popis bude proto předmětem budoucí práce. Výstupem tohoto programu jsou periody řetězového zlomku θ, které pro zadané délky period slov $\mathbf{y}$ a $\mathbf{y}'$ mohou nabývat asymptotický kritický exponent nižší, než je zadaná mez. Napřed je tedy nutno najít nějakou vhodnou teoretickou mez, aby výsledných period nebylo příliš mnoho, a dále musíme spustit náš program pro asymptotický kritický exponent se stejnými parametry, abychom nalezli předperiodu θ, na které je minimum opravdu nabýváno.

Touto tematikou se intenzivně zabýváme, program testujeme a dále rozvíjíme a do budoucna bychom rádi vyslovili a nejlépe i dokázali podobnou domněnku jako pro kritický exponent.

Tabulka 8.2: Balancovaná slova nad d-písmennou abecedou pro $d \leq 6$, která nabývají minimálního asymptotického kritického exponentu.

d	θ	$\mathrm{Per}(\mathbf{y})$	$\mathrm{Per}(\mathbf{y}')$	$\mathrm{E}^*(\mathbf{v})$
3	$[0, 1, (2)^\omega]$	1	2	$2 + \frac{\sqrt{2}}{2}$
4	$[0, (1)^\omega]$	2	2	$\frac{5+\sqrt{5}}{4}$
5	$[0, 1, (2)^\omega]$	2	4	$\frac{3}{2}$
6	$[0, 1, 3, (1, 2, 1, 1, 1)^\omega]$	1	16	$\frac{75+3\sqrt{65}}{80}$

Závěr

Cílem práce bylo popsat teoretické odvození algoritmů pro výpočet (asymptotického) kritického exponentu, ty implementovat a dále popsat z toho odvozené programy. Tento cíl byl naplněn, programy jsou dostupné na [15] a jsou intenzivně využívány pro zkoumání vlastností kritického exponentu a také dále rozšiřovány, aby co nejvíce pomohly při zkoumání minimálního asymptotického kritického exponentu. Jak mohl čtenář vidět i na numericky méně náročných příkladech 5.1 a 6.1, jde o zdlouhavé výpočty a je jednoduché při výpočtech v ruce nějaký faktor přehlédnout nebo udělat numerickou chybu. Dalším problémem výpočtu v ruce je jeho časová náročnost.

Důležité je tedy podtrhnout, že oba naše programy fungují a pro námi zkoumané velikosti abeced počítají exponenty v řádu sekund. Jsou proto pro nás nástrojem, který nám pomáhá posunout se dál při zkoumání otevřených problémů. Jedním z těchto otevřených problémů je zkoumání minimálního kritického exponentu balancovaných slov.

Tímto problémem se zabývali matematici již před námi, měli ale jiný přístup a narazili na problémy s obrovskou paměťovou a výpočetní náročností algoritmů. Například pro abecedu o 8 písmenech bylo potřeba 360 GB paměti a výpočet trval dvě hodiny. Pro abecedu o více než 8 písmenech pak již jejich výpočetní technika nestačila.

Náš přístup je jiný, protože pro zadané parametry umí naše programy (asymptotický) kritický exponent vyčíslit, proto pro 9- a 10-písmennou abecedu právě náš program ověřil, že předložení kandidáti nabývají kritického exponentu $\frac{7}{6}$, resp. $\frac{8}{7}$. Pro určení minimálního kritického exponentu nad většími abecedami již je potřeba napřed nacházet vhodné kandidáty, v čemž jsou nám existující programy také velkou oporou.

Systematickým procházením vhodných kombinací sturmovských slov a jejich obarvení pomocí šikovně zvolených slov s konstantními mezerami se napřed podařilo nalézt slova nad 11- a 12-písmennými abecedami, které mají kritický exponent nižší, než říkala původní domněnka. Pomocí výpisu programu se nám výsledek dokonce podařilo zobecnit i na libovolný sudý počet písmen. Dále jsme ve spolupráci s A. M. Shurem [9] dokázali, že nově nalezený minimální kritický exponent je opravdu dolní závorou. Nyní tedy stačí ukázat, jak je to pro abecedy o lichém počtu písmen.

Další zcela otevřenou otázkou je hledání minimálního asymptotického kritického exponentu balancovaných slov nad danou abecedou. Program nám umožnil ukázat, že pro abecedy nad 3, 4 a 5 písmeny minimální kritický a minimální asymptotický kritický exponent splývají, ale už pro 6-písmennou abecedu je minimální asymptotický kritický exponent roven $\frac{5}{4}$, což je méně než minimální kritický exponent, který má hodnotu $\frac{4}{3}$.

Pro zkoumání minimálního asymptotického kritického exponentu nad většími abecedami opět budeme muset změnit přístup a na novém programu intenzivně pracujeme. Cílem budoucí práce je pak vyslovit a dokázat domněnky o hodnotě minimálního asymptotického kritického exponentu i pro větší abecedy.

Literatura

[1] A. R. Baranwal, *Decision algorithms for Ostrowski-automatic sequences*, master thesis, University of Waterloo, http://hdl.handle.net/10012/15845 (2020).

[2] A. R. Baranwal, J. Shallit, *Critical exponent of infinite balanced words via the Pell number system*, Proceedings WORDS 2019, LNCS, 11682, Springer (2019), 80-92.

[3] L. Balková, M. Bucci, A. De Luca, J. Hladký, S. Puzynina, *Aperiodic pseudorandom number generators based on infinite words*, Theoret. Comput. Sci. 647 (2016), 85-100.

[4] Jean Berstel, *Axel Thue's papers on repetitions in words: a translation*, Publications du LaCIM, Département de mathématiques et d'informatique 20, Université du Québec à Montréal, 1995, 85 pages.

[5] F. Dolce, L'. Dvořáková, E. Pelantová, *On balanced sequences and their critical exponent*, arXiv: 2108.07503, 2021, submitted to Theoret. Comput. Sci.

[6] F. Dolce, L'. Dvořáková, E. Pelantová, *On balanced sequences and their asymptotic critical exponent*, Proceedings LATA 2021, LNCS vol. 12638, 293-304 (2021).

[7] F. Dolce, L'. Dvořáková, E. Pelantová, *Computation of critical exponent in balanced sequences*, In: Lecroq T., Puzynina S. (eds) Combinatorics on Words. WORDS 2021. Lecture Notes in Computer Science, vol 12847. Springer, Cham.

[8] L'. Dvořáková, K. Medková, E. Pelantová, *Complementary symmetric Rote sequences: the critical exponent and the recurrence function*, Discrete Math. Theoret. Comput. Sci. 20 (1) (2020) #20

[9] L'. Dvořáková, E. Pelantová, D. Opočenská, A. M. Shur, *On minimal critical exponent of balanced sequences*, Theoret. Comput. Sci. 922 (2022), 158-169, ISSN 0304-3975 , https://doi.org/10.1016/j.tcs.2022.04.021.

[10] G. A. Hedlund, M. Morse, *Symbolic dynamics II - Sturmian trajectories*, Amer. J. Math. 62 (1940), 1-42

[11] P. Hubert, *Suites équilibrées*, Theoret. Comput. Sci. 242 (2000), 91-108.

[12] J. Justin, G. Pirillo, *Episturmian words and episturmian morphisms*, Theoret. Comput. Sci. 276 (2002), 281-313.

[13] J. Justin and G. Pirillo, *Fractional powers in Sturmian words*, Theoret. Comput. Sci. 255 (2001), 363–376.

[14] M. Lothaire, *Algebraic combinatorics on Words, Encyclopaedia of mathematics and its applications*, Cambridge University Press (2002), doi:10.1017/CBO9781107326019.

[15] D. Opočenská, *Programy pro výpočet (asymptotického) kritického exponentu balancovaných slov*, dostupné na `https://github.com/Opocedan/Kriticky_exponent_balancovanych_slov`.

[16] Qt: `https://www.qt.io/`

[17] N. Rampersad, J. Shallit, É. Vandomme, *Critical exponents of infnite balanced words*, Theoret. Comput. Sci. 777 (2019), 454-463.

[18] L. Vuillon, *A characterization of Sturmian words by return words*, Eur. J. Combin. 22 (2001), 263-275.